

# post-carbon

Proceedings of the 27<sup>th</sup> International Conference on Computer-Aided  
Architectural Design Research in Asia (CAADRIA 2022)

**Volume 1**

*Edited by*

**Jeroen van Ameijde**

*The Chinese University of Hong Kong*

**Nicole Gardner**

*University of New South Wales*

**Kyung Hoon Hyun**

*Hanyang University*

**Dan Luo**

*University of Queensland*

**Urvi Sheth**

*CEPT University*

**Post-Carbon**

27<sup>th</sup> International Conference on Computer-Aided Architectural Design  
Research in Asia (CAADRIA 2022)

9-15 April 2022

Hosted By:

University of New South Wales

The University of Sydney

University of Technology Sydney

©2022 *All rights reserved and published by*

The Association for Computer-Aided Architectural Design Research in Asia  
(CAADRIA), Hong Kong

ISBN: 978-988-78917-7-2

ISSN: 2710-4257 (print)

2710-4265 (online)

*Printed in Sydney, Australia*

## Foreword

The annual Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) conference provides an international community of researchers and practitioners with a venue to exchange, to discuss and to publish their latest ideas and accomplishments. These proceedings, consisting of two volumes, contain the research papers that were accepted for presentation at the 27th International CAADRIA Conference, organised jointly by the University of New South Wales, The University of Sydney, and the University of Technology Sydney.

The work collected in these volumes has been produced during extraordinary circumstances, as the Covid-19 pandemic has continued to impact working and living conditions. Papers have been written, reviewed and edited throughout lockdowns, challenging family situations and health crises. The high number of submissions and quality of the papers is a testament to the resilience and strength of the CAADRIA community.

The papers in this publication have been selected through a two-stage, double-blind peer review process. All reviews and papers have been evaluated by the Paper Selection Committee. After receiving an initial 488 abstract submissions, a final set of 150 full papers has been selected for publication. We thank all authors for their contributions to the process and congratulate the accepted paper authors.

The papers in these proceedings are specifically selected for their contribution to this year's conference theme, following the conference organisers' call for authors to position their work in relation to the United Nations Sustainable Development Goals (SDGs). As the world is experiencing the increasing impacts of climate change, there is an urgent need to reflect on the potential of the latest research in architecture, urbanism and construction to address these global challenges.

Our special thanks go out to our more than 180 international reviewers, who volunteered their valuable time and effort. We also thank the organising team in Sydney for developing the 2022 CAADRIA Conference as a strong collaborative effort between academic institutions, industry and practice. We thank Gabriel Wurzer for his support during the proceedings' final production stage, and for facilitating their publication via the open access database CumInCAD.org. Finally, we sincerely thank the CAADRIA community for offering us the honour to serve as members of the paper selection committee for *Post-Carbon, the 27<sup>th</sup> International CAADRIA Conference 2022*.

<i>Jeroen van Ameijde</i>	<i>The Chinese University of Hong Kong (Chair)</i>
<i>Nicole Gardner</i>	<i>University of New South Wales</i>
<i>Kyung Hoon Hyun</i>	<i>Hanyang University</i>
<i>Dan Luo</i>	<i>University of Queensland</i>
<i>Urvi Sheth</i>	<i>CEPT University</i>

Theme: *‘Post-Carbon’*

An increase of the world population to up to 10 billion people by 2050, coupled with a current continued economic growth, digitalisation, and infrastructural expansions in many countries require us to reconsider the global carbon impact.

The consequences of the fourth industrial revolution and resulting climate changes can be directly and physically experienced as phenomena around the globe and by many of us; as coastal and river floods, cyclones, bushfires, air pollution, a decline of resources, and diminished natural habitats.

In this context, computational design, simulation, analysis, fabrication, and management allow us to evaluate, understand and forecast trends and consequences of connected impacts across multiple disciplines. In a post-carbon framework, computation can be applied for improving the quality, sustainability, and resilience of the architecture, infrastructures and public resources of the built environment.

Consequently, the conference theme ‘Post-Carbon’ seeks profoundly different approaches to identify closer dialogues, better collaboration, increased agency and effective ways to address a world in which climate change has become a reality. We specifically call for papers that address the UN Sustainable Development Goals as we are forced to live with extreme climate, live with limited resources and live with reduced biodiversity.

In 2022 we bring again together academics, researchers, and practitioners involved in contributing to applying computational design methods, tools, and processes towards achieving a post-carbon future of the architecture, engineering and urban design sector.

Yet to strengthen the focus on Post-Carbon issues, for the first time in the CAADRIA history we explicitly asked authors to nominate UN SDG’s (Sustainability Design Goals) as keywords for their paper submissions. This allowed a later search for SDGs in the CuminCad platform and linked technical topics in computational design research with SDGs. The conference discusses and presents including but not limited the following topics:

- Digital Representation and Visualisation
- Generative, algorithmic & evolutionary design
- Urban Analytics and Smart Cities
- Theory, Philosophy and Methodology of Computational Design
- Artificial Intelligence & Machine Learning for Urban Design
- Computational Design in Education
- Building Information Modelling
- Digital Heritage
- Robotics & Digital Fabrication in Construction and Fabrication
- Virtual and Augmented Reality
- Artificial Intelligence & Machine Learning

- Innovative material systems and 3D printing
- Environment performance based design and sustainability

A further novelty is that this conference is a joint conference between three Sydney universities. We, as General Chairs, have the belief that a post-carbon future and changing the contribution and impact of the Architecture, Engineering and Construction sector to climate change is challenge that need to bring academic and industry partners together and we have planned this CAADRIA conference as a starting point to work closer together and address climate change through a computational design lens.

*Conference Organisers & Hosts,*

*Dagmar Reinhardt      The University of Sydney*  
*Tim Schork              University of Technology Sydney*  
*M. Hank Haeusler      University of New South Wales*

{<https://caadria2022.org>}

## About CAADRIA

The *Association for Computer-Aided Architectural Design Research in Asia* (CAADRIA) promotes teaching and research in CAAD in the larger Austral-Asian and Pacific region supported by a global membership.

CAADRIA was founded in 1996 with the following objectives:

- To facilitate the dissemination of information about CAAD among Asian schools of architecture, planning, engineering, and building sciences.
- To encourage the exchange of staff, students, experience, courseware, and software among schools.
- To identify research and develop needs in CAAD education and to initiate collaboration to satisfy them.
- To promote research and teaching in CAAD that enhances creativity rather than production.

CAADRIA organizes among others an annual conference, the first of which was held in 1996 in Hong Kong. Since then, 26 conferences have been held in Australia, China, Hong Kong, India, Japan, Korea, Malaysia, New Zealand, Singapore, Taiwan, and Thailand. The annual CAADRIA conferences provide an opportunity to meet, to learn about the latest research, and to continue the discourse in the field. The 27<sup>th</sup> conference, in 2022, is hosted jointly by three universities located in Sydney, Australia: the University of New South Wales, the University of Sydney and the University of Technology Sydney. CAADRIA 2022 is held mainly as a virtual conference, continuing the Association's mission to bring together researchers, practitioners and schools of the Asia-pacific region at an extended time of global Covid-19 related travel restrictions.

CAADRIA is one of the four founding organizations of the *International Journal of Architectural Computing* (IJAC), and typically co-edits one issue each year. IJAC is published by SAGE in both paper and electronic versions.

*Christiane M. Herr*  
*President, CAADRIA*

## CAADRIA Officers

- President: Christiane M. Herr  
*Xi'an Jiaotong-Liverpool University, China*
- Secretary: Urvi Sheth  
*CEPT University, India*
- Treasurer: Hyoung-June Park  
*University of Hawaii, USA*
- Membership Officer: Immanuel Koh  
*Singapore University of Technology and Design, Singapore*
- Web Master: Kyung Hoon Hyun  
*Hanyang University, Korea*
- Administrative Officer: Marc Aurel Schnabel  
*Victoria University of Wellington, New Zealand*

## CAADRIA 2021 Conference Committees

### Hosting Institutions:

- University of New South Wales  
The University of Sydney  
University of Technology Sydney

### Organising Committee:

- |                  |  |                                    |
|------------------|--|------------------------------------|
| Dagmar Reinhardt | <i>The University of Sydney</i>        | <i>General Conference Chairs</i>   |
| Tim Schork       | <i>University of Technology Sydney</i> | <i>General Conference Chairs</i>   |
| M. Hank Haeusler | <i>University of New South Wales</i>   | <i>General Conference Chairs</i>   |
| Nicole Gardner   | <i>University of New South Wales</i>   | <i>Local Paper Selection Chair</i> |
| Mohammed Makki   | <i>University of Technology Sydney</i> | <i>Workshop Chairs</i>             |
| JuHyun Lee       | <i>University of New South Wales</i>   | <i>Workshop Chairs</i>             |
| Anastasia Globa  | <i>The University of Sydney</i>        | <i>Exhibition Chairs</i>           |

Linda Matthews  
*University of Technology Sydney*      *Exhibition Chairs*

K. Daniel Yu  
*University of New South Wales*      *Technical Coordinator and  
 Website Designer*

Tim Schork  
*University of Technology Sydney*      *Satellite Events*

Paper Selection Committee:

Jeroen van Ameijde  
*The Chinese University of Hong Kong*      *Chair*

Nicole Gardner  
*University of New South Wales*

Kyung Hoon Hyun  
*Hanyang University*

Dan Luo  
*University of Queensland*

Urvi Sheth  
*CEPT University*

Post Graduate Students Consortium:

Dagmar Reinhardt      *Co-Chair*  
 Sky Lo Tian Tian      *Co-Chair*  
 Christiane M. Herr      *Co-Chair*

Sasada Prize Committee:

Tom Kvan      *Chair*  
 Christiane M. Herr  
 Atsuko Karga

Young CAADRIA Award Committee:

Sky Lo Tian Tian      *Chair*  
 Yasushi Ikeda  
 Christiane M. Herr  
 Dagmar Reinhardt  
 Nic Bao Dingwen



## International Reviewing Committee

Anders Kruse Aagaard  
*Aarhus School of Architecture*

Aysegul Akcay Kavakoglu  
*Istanbul Technical University*

Mostafa Alani  
*Tuskegee University*

Suleiman Alhadidi  
*Harvard University*

Miktha Farid Alkadri  
*Universitas Indonesia*

Jayedí Aman  
*University of Missouri Columbia*

Gonçalo Araújo  
*IN+/ Técnico Lisboa*

Cem Ataman  
*Singapore University of Technology  
and Design*

Nan Bai  
*TU Delft*

Christopher Bamborough  
*University of Technology Sydney*

Ding Wen 'Nic' Bao  
*MIT University*

Shany Barath  
*Technion - Israel Institute of  
Technology*

Muge Belek Fialho Teixeira  
*The University of Adelaide*

Ardavan Bidgoli  
*Carnegie Mellon University*

Luis Borunda  
*Virginia Tech*

Johannes Braumann  
*Creative Robotics | Robots in  
Architecture*

Michael Budig  
*Singapore University of Technology  
and Design*

Inês Caetano  
*INESC-ID/IST, University of Lisbon*

Glenda Caldwell  
*Queensland University of Technology*

Renata Castelo-Branco  
*INESC-ID/IST, University of Lisbon*

Teng-Wen Chang  
*National Yunlin University of  
Science and Technology*

Aqil Cheddadi  
*Keio University*

Feihao Chen  
*The Chinese University of Hong Kong*

Jia-Yih Chen  
*CS Cheng & JY Chen Architects &  
Partners*

Shuo Chen  
*Beijing University of Technology*

Zi-Ru Chen  
*Southern Taiwan University of  
Science and Technology*

I-Ting Chuang  
*University of Auckland*

Verina Cristie  
*Singapore University of Technology  
and Design*

Kristof Crolla  
*The University of Hong Kong*

Camilo Cruz Gambardella  
*Monash University*

Weiwen Cui  
*Tsinghua University*

Pierre Cutellic  
*ETH Zurich*

Avishek Das  
*Aalborg University*

Mohammad Dastmalchi  
*University of Missouri*

Mohammad Reza Dastmalchi  
*University of Missouri*

Aurelie De Boissieu  
*University of Liège*

Danilo Di Mascio  
*The University of Huddersfield*

- Wowo Ding  
*Nanjing University*
- Thomas Dissaux  
*The University of Liège*
- Ben Doherty  
*University of New South Wales*
- Jonathan Dortheimer  
*Technion, Israel Institute of Technology*
- Theodoros Dounas  
*Robert Gordon University*
- Scott Drake  
*INDA Chulalongkorn University*
- Halil Erhan  
*Simon Fraser University*
- Alberto T. Estevez  
*iBAG-UIC Barcelona*
- Shayani Fernando  
*Digital Design Unit, TU Darmstadt*
- Paolo Fiamma  
*University of Pisa*
- Adam Fingrut  
*The Chinese University of Hong Kong*
- Antonio Fioravanti  
*Sapienza University of Rome*
- Vasiliki Fragkia  
*The Royal Danish Academy*
- Haruyuki Fujii  
*Tokyo Institute of Technology*
- Tomohiro Fukuda  
*Osaka University*
- Song Gao  
*Independent researcher*
- Fernando García Amen  
*The University of the Republic*
- Nicole Gardner  
*University of New South Wales*
- Nadja Gaudilliere-Jami  
*Digital Design Unit, TU Darmstadt*
- Mona Ghandi  
*Washington State University*
- Anastasia Globa  
*The University of Sydney*
- Zhuoxing Gu  
*Tongji University*  
*National University of Singapore*
- Qi Guo  
*Harbin Institute of Technology, Shenzhen*
- Benay Gursoy  
*Penn State University*
- Mahyar Hadighi  
*Texas Tech University*
- M. Hank Haeusler  
*University of New South Wales*
- Gilles Halin  
*MAP-CRAI laboratory, University of Lorraine*
- Yoojin Han  
*Yonsei University*
- Farahbod Heidari  
*Tarbiat Modares University*
- Christiane Herr  
*Southern University of Science and Technology*
- Pablo C Herrera  
*Peruvian University of Applied Sciences*
- Anca-Simona Horvath  
*Aalborg University*
- Wei Hu  
*Tongji University*
- Chien-Hua Huang  
*China Academy of Art*
- Xiaoran Huang  
*North China University of Technology*  
*Swinburne University of Technology*
- Markus Hudert  
*Aarhus University*
- Kyung Hoon Hyun  
*Hanyang University*
- Aswin Indraprastha  
*Institut Teknologi Bandung*
- Daichi Ishikawa  
*Fujitsu Limited*

- Patrick Janssen  
*National University of Singapore*
- Taysheng Jeng  
*National Cheng Kung University*
- Mads Brath Jensen  
*Aalborg University*
- Woonseong Jeong  
*Chungbuk National University*
- Guohua Ji  
*Nanjing University*
- Sam Joyce  
*Meta Design Lab, SUTD*
- Yuval Kahlon  
*Tokyo Institute of Technology*
- Ammar Kalo  
*American University of Sharjah*
- Nariddh Khean  
*City Intelligence Lab, Austrian  
Institute of Technology*
- Chin Koi Khoo  
*Deakin University*
- Daiki Kido  
*Kajima Corporation*
- Nayeon Kim  
*Yonsei University*
- Geoff Kimm  
*Swinburne University of Technology*
- Marirena Kladeftira  
*ETH Zurich*
- Immanuel Koh  
*Singapore University of Technology  
and Design*
- Seow Jin Koh  
*Defence Science & Technology  
Agency*
- Satakhun Kosavinta  
*King Mongkut's Institute of  
Technology Ladkrabang*
- Dennis Lagemann  
*ETH Zurich*
- Ih-Cheng Lai  
*Tamkang University*
- Christoph Langenhan  
*Technical University of Munich*
- António Leitão  
*Instituto Superior Tecnico*
- Surapong Lertsithichai  
*Chulalongkorn University*
- Matthias Leschok  
*ETH Zurich*
- Pok Yin Leung  
*ETH Zurich*
- Andrew Li  
*Independent researcher*
- Bin Li  
*South China University of  
Technology*
- Jie Liu  
*Tsinghua University*
- Yuezhong Liu  
*Nanyang Technological University*
- Tian Tian Lo  
*Harbin Institute of Technology,  
Shenzhen*
- Paul Loh  
*University of Melbourne*
- Davide Lombardi  
*Xi'an Jiaotong - Liverpool  
University*
- Thorsten Lomker  
*Zayed University*
- Werner Lonsing  
*Independent Researcher*
- Shuai Lu  
*The University of Sydney*
- Yao Lu  
*University of Pennsylvania*
- Dan Luo  
*University of Queensland*
- Bob Martens  
*TU Wien*
- Rodrigo Martin Iglesias  
*University of Buenos Aires*
- Samim Mehdizadeh  
*Digital Design Unit, TU Darmstadt*

- Deedee Min  
*Hanyang University*
- Hugo Mulder  
*University of Southern Denmark*
- Rizal Muslimin  
*The University of Sydney*
- Roberto Naboni  
*CREATE-University of Southern Denmark*
- Walaiporn Nakapan  
*Parabolab*
- Zuhair Nasar  
*University of Kufa*
- Worawan Natephra  
*Maharakham University*
- Provides Ng  
*The Chinese University of Hong Kong*
- Farzaneh Oghazian  
*The Pennsylvania State University*
- Mine Özkar  
*Istanbul Technical University*
- Hyoung-June Park  
*University of Hawaii at Manoa*
- Trevor Patt  
*Singapore University of Technology and Design*
- Nyoman Dewi Pebryani  
*Indonesian Institute of the Arts (ISI) Denpasar*
- Wanyu Pei  
*National University of Singapore*
- Inês Pereira  
*INESC-ID/IST, University of Lisbon*
- Peiman Pilechiha  
*Tarbiat Modares University*
- Mina Pouyanmehr  
*Iran University of Science and Technology*
- Hendro Trieddiantoro Putro  
*Universitas Teknologi Yogyakarta*
- Deyan Quan  
*Xi'an Jiaotong-liverpool University*
- Ahmad Rafi  
*Multimedia University*
- Balaji Rajasekaran  
*C.A.R.E School of Architecture*
- Dagmar Reinhardt  
*The University of Sydney*
- Jinmo Rhee  
*Carnegie Mellon University*
- Iasef Md Rian  
*Xi'an Jiaotong-Liverpool University*
- Manuel Rodriguez Ladron De Guevara  
*Carnegie Mellon University*
- Nicolas Rogeau  
*Ecole Polytechnique Fédérale de Lausanne*
- Luís Romão  
*University of Lisbon*
- Huda Salman  
*Robert Gordon University*
- Eike Schling  
*The University of Hong Kong*
- Marc Aurel Schnabel  
*Victoria University of Wellington*
- Gerhard Schubert  
*Technical University of Munich*
- Urvi Sheth  
*CEPT University*
- Sahar Soltani  
*University of South Australia*
- R. Spencer Steenblik  
*Wenzhou-Kean University*
- Djordje Stojanovic  
*University of Melbourne*
- Rudi Stouffs  
*National University of Singapore*
- Marcin Strzala  
*Warsaw University of Technology*
- Ying Yi Tan  
*Singapore University of Technology and Design*
- Wajiha Tariq  
*The Chinese University of Hong Kong*

- Teng Teng  
*Cornell University*
- Ziyu Tong  
*Nanjing University*
- Robert Trempe  
*Arkitektskolen Aarhus*
- Jeroen van Ameijde  
*The Chinese University of Hong Kong*
- Guzden Varinlioglu  
*Izmir University of Economics*
- Elena Vazquez  
*Penn State University*
- Tomas Vivanco  
*PUC Chile*  
*Tongji University*
- Likai Wang  
*Nanjing University*
- Shih-Yuan Wang  
*National Yang Ming Chiao Tung University*
- Sihan Wang  
*Singapore University of Technology and Design*
- Zichu Wang  
*EDG*
- Glen Wash  
*Xi'an Jiaotong-Liverpool University*
- Claudia Westermann  
*Xi'an Jiaotong-Liverpool University*
- Albert Wiltsche  
*Graz University of Technology*
- Thomas Wortmann  
*University of Stuttgart*
- Yen-Liang Wu  
*National Taichung University of Science and Technology*
- Chao Yan  
*Tongji University*
- Jiawei Yao  
*Tongji University*
- Hatzav Yoffe  
*Technion - Israel Institute of Technology*
- Rongrong Yu  
*Griffith University*
- Philip F. Yuan  
*Tongji University*
- Ata Zahedi  
*Technische Universität München*
- Yuqing Zhang  
*South China University of Technology*
- Hao Zheng  
*University of Pennsylvania*
- Jianjia Zhou  
*Tongji University*
- Guanqi Zhu  
*The University of Queensland*
- Xinwei Zhuang  
*University of California Berkeley*

## **Keynote Panels**

Whereas conferences are often stimulated and enriched by keynotes reporting on advancements in research, industry and applications, the Pandemic situation has enabled us to virtually and globally participate in such keynotes for the last two years.

CAADRIA in 2022 pursues a different trajectory in the format of four explicit roundtable conversations, to create direct exchanges and interactions between key agents from research, industry, councils, or planning that directly impact on the discussions for a post-carbon future. The aim is to bring stakeholders, operating at the front face of post-carbon challenges in contact with the computational design community to facilitate meaningful discussions leading to potential pathways of novel methods and tools to address climate change.

As tackling the UN's Sustainable Development Goals is a pressing issue that requires great resources and insights into vast amounts of information and data, we believe that computation and computing can play an important role to accelerate the speed of how these problems are addressed. Yet as Ostwald (2017) observed when studying research of the CuminCAD community published between 1995 - 2017 our research faces the “the dangers of self-referentiality and insularity, the possible loss of grounding in industrial or professional needs and applications, and the lack of consideration of a growing number of problems facing the modern world” - naming “themes include crime, homelessness, politics, poverty, gender, emotions, ethics and violence” as often absent as application of research findings.

Hence the proposed format of the panel discussions as a wake-up call for the community - to take action and to actively develop, discuss and share the advanced knowledge we have developed in the last decade towards direct applications for real life and pressing problems.

Topics as listed earlier that the conference discusses and presents must be brought into the direct social, cultural or philosophical context of which they can be applied to. Research in computational design must find its place as solvers for design, simulation and analysis, so that we can project and forecast pathways to a better future - across different scales and multiple threads. The following four panel formats and panel members were selected:

## **Keynote Panel 1 (Computational to Sustainable Computational)**

### **When We Delivered Digital Architecture in Australia and What We Would Deliver Now**

While Digital Architecture has successfully entered architecture and the built environment worldwide, how can we now approach a digital sustainable architecture? Twenty years ago, an Australian Research Council funded research into ‘Delivering Digital Architecture’ in Australia. led by Prof. Mark Burry and Prof. Mike Xie. The research focus at the time centred around the observation that “design, documentation, and off-site prefabrication of numerous recent major international architectural projects, were achieved with evolving integrated 3D digital methods”. It pointed out that, at the time “local understanding and uptake of the opportunities is still very low [in Australia]”, and that joining these opportunities could result in “a huge cultural dividend to Australia from greater affordable freedom of architectural expression”. While it is arguable if digital architecture has been already fully delivered in Australia and the cultural dividend could have been cashed in, this discussion wants to focus on the future as a post-carbon future. Thus, it discusses with experts who aimed to deliver digital architecture at the time on how a digital sustainable architecture can be delivered over the next decade to meet 2030 carbon reduction goals.

This discussion will set the scene for the conference as it reflects and unpacks, after a short overview of “we delivered digital architecture in Australia”, current motivations, barriers, and opportunities to a digital sustainable architecture. Focus on this online forum will holistically discuss research to education to practice with a national and international perspective.

#### **Panel:**

1. Prof Mark Burry (Swinburne, Melbourne, Australia)
2. Prof Mike Xie (RMIT, Melbourne, Australia)
3. Prof Mette Ramsgaard Thomsen (CITA, Copenhagen, Denmark)
4. Prof Philip Yuan (Tongji, Shanghai, China)
5. Prof Areti Markopoulou (IAAC, Barcelona, Spain)
6. Louise Wotton (Regional Computational Design Lead NZ Aurecon Wellington, New Zealand)

#### **Moderator:**

Prof Michael Ostwald (UNSW, Sydney, Australia)

## **Keynote Panel 2 (Urban / Systems)**

### **The Role of Digital Disruptors in Assisting the Design of Resilient Cities**

Are digital tools and platforms in the AEC industry only for developers to increase profits or do they have potential in creating an equal and just city? Over recent years AEC platforms have gained global prominence and raised millions of dollars of capital. But not only overseas, local Sydney AEC platforms have positioned themselves successfully on a global market and started interrupting the AEC sector. Focus of these platforms are in: “revolutionises the way property professionals find, assess and design sites. Transform your months’ long process into a ten-minute analysis” (Archistar, 2022), “conducting site analysis in real time with architectural design” (Giraffe Technology, 2022), or to provide “an end-to-end digital platform that provides the property and construction industry with insights and clarity never experienced before” (Podium). But do they address the UN’s SDGs and help with a post-carbon future? If they do so, what tools exist for improving the urban condition? Where are future directions and initiatives in urban design and what are current and future obstacles?

The discussion in this face-to-face panel aims to unpack the opportunities within digital transformation and disruption of the AEC sector towards a sustainable digital approach by inviting a panel of PropTech and RealTech CEOs, sustainable experts, and community activists. We want to discuss innovation potentials through synergies between computational creativity, environmental science and digital technology and disruption. Outcomes of this discussion could contribute to pressing but interrelated issues of property ownership & public space; thermal comfort, floods, and/or bush fires; and data and/or machine learning.

#### **Panel:**

1. Rob Asher (CEO Giraffe Technology, Sydney, Australia)
2. Abbie Galvin (Government Architect NSW, Sydney, Australia)
3. Steve Fox (Principle Architectus, Sydney, Australia)
4. Dr. Ben Coorey (CEO ArchiStar, Sydney, Australia)
5. Kathlyn Looseby (CEO AACA, Sydney, Australia)
6. Amanda Wyzenbeek (Technical Director, Mott MacDonald, Sydney, Australia)

#### **Moderator:**

Prof Flora Salim (UNSW, Sydney, Australia)



### **Keynote Panel 3 (Materials Futures)**

#### **Resilience, Resources and Consumption – Circular Economy, New Materials and Construction Methods**

The Architecture, Engineering, Construction (AEC) sector has a devastating effect on the environment and contributes significantly to climate change. It is responsible for almost 50% of the world's total energy consumption and 40% of global carbon emissions. It consumes 50% of the world's raw materials and produces 35% of its waste. Yet with rising population we need to build even more to provide housing and infrastructure for all, but with the existing impacts we cannot afford to continue building in the way we do now. Thus, there is a real and urgent need to rethink our current building culture and practices and to find solutions that allow to build more with less material and using renewable resources that create less greenhouse gases while at the same time creating high-quality outcomes.

This panel will focus on what actions are required to create a decarbonised building culture. It will discuss opportunities and challenges for responsible consumption and production and examine the potentials of new material systems, transformative technology and alternative economic business models that can reduce the environmental footprint of our industry.

#### **Panel:**

1. Dr Rumana Hossain (SMART Center, UNSW, Sydney, Australia)
2. Dr Josephine Vaughan (MECLA, Newcastle, Australia)
3. Enrico Zara (Decarbonisation Lead, Arup, Sydney, Australia)
4. Prof Peter Ralph (Director C3, UTS, Sydney, Australia)
5. Chris Bickerton (BVN, Sydney, Australia)
6. Dr Rebecca Huntley (Climate change researcher and strategist, Sydney, Australia)
7. Dr Kar Mei Tang (Chief Circular Economist, Circular NSW, Sydney, Australia)

#### **Moderator:**

Prof Anna Cristina Pertierra (UTS, Sydney, Australia)

## **Keynote Panel 4 (Knowledge Systems and Culture)**

### **Knowledge Systems and Cultures for a New Ecology Approach**

To conclude this conference, we ask: In which way can computational design, analysis and frameworks support our understanding of natural conditions in our cities, inform the planning, design, and delivery of built environment and sustainable projects – all to help humans to develop meaningful connections with Ecology, Country and Living Species?

There is evidence that increasingly, multi-expert teams collaborate to plan for events of Climate Change including fire, drought, and flooding through sustainable land and water use practices. Here recent projects range from the contribution of tree canopies to generate better climatic conditions in dense urban zones to developing guidelines and measurement systems that enable us to better understand the complex interactions and quality affordances across connected and interacting SDGs. One can witness new approaches to cultural landscapes beyond human-centred design that consider ethical and just multispecies cohabitation and interaction. Indigenous knowledge guided by people and communities is drawn upon for valuing, respecting and working with the country, supporting health and wellbeing is one example for change. Similarly, Artificial Intelligence can provide frameworks for knowledge systems to contribute to resilient future cities. How can these topics merge into a new knowledge system to create cultures for a new ecology approach?

The discussion in this face-to-face panel aims to unpack the opportunities, opening crucial aspects and demands that could be addressed and further developed through current computational design methods, thinking, and tools. As the last panel in the Roundtable series, we seek to generate here discussion on and pathways for the future, and establish a platform for the CAADRIA community with researchers, experts and thinkers. We invite speakers to share their work, experiences, provocations, thought frames and reports across the fields of Indigenous Knowledge, Artificial Intelligence, Urban Heat Islands, Multispecies Justice and Resilient Communities. By addressing the many dimensions of Climate Change for a different Culture of Ecology, beyond a mere human driven or technology focused worldview.

#### **Panel:**

1. Danielle Celermajer (SEI Sydney Environmental Institute, Sydney, Australia)
2. Dr Sebastian Pfautsch (WSU, Sydney, Australia)
3. Jorge Chapa (GBCA Green Building Council Australia, Sydney, Australia)
4. Prof Toby Walsh (Scientia Professor AI UNSW, Sydney, Australia)
5. Amanda Sturgeon (Regenerative Design Lead, Mott Macdonald, Sydney, Australia)
6. Michael Mossman (USyd, Associate Dean Indigenous, Sydney, Australia)

**Moderator:**

Prof Robyn Dowling (Usyd, Sydney, Australia)

REFERENCES

- Ostwald, M. (2017). Digital research in architecture: Reflecting on the past, analysing the trends, and considering the future. *Architectural Research Quarterly*, 21(4), 351-358. doi:10.1017/S135913551800009X



## TABLE OF CONTENTS

<b>Virtual and Augmented Reality</b>	<b>9</b>
Kurashiki Viewer: Qualitative Evaluations of Architectural Spaces inside Virtual Reality <i>Taro Narahara</i>	11
Virtual Reality Collaborative Platform for E-learning: Analysis of Student Engagement and Perceptions <i>Joowon Jeong, Qinchuan Chen, Nayeon Kim, Hyunsoo Lee</i>	19
Collaborative Design Review Sessions in Virtual Reality: Multi-Scale and Multi-User <i>Shahin Sateei, Mattias Roupé, Mikael Johansson</i>	29
Dynamic Projection <i>James Nanasca, Aaron G. Beebe</i>	39
Automatic Flythrough Authoring in VR <i>Robin Schmidiger, Romana Rust, Roi Poranne</i>	49
Understanding Design Experience in Virtual Reality for Interior Design Process <i>Hwan Kim, Kyung Hoon Hyun</i>	59
GIS-Based Educational Game Through Low-Cost Virtual Tour Experience-Khan Game <i>Guzden Varinlioglu, Sepehr Vaez Afshar, Sarvin Eshaghi, Özgün Balaban, Takehiko Nagakura</i>	69
A Method of VR Enhanced POE for Wayfinding Efficiency in Mega Terminals of Airport <i>Shuyang Li, Chengyu Sun, Yinshan Lin</i>	79
Developing an Augmented Reality System with Real-Time Reflection for Landscape Design Visualization, Using Real-Time Ray Tracing Technique <i>Hao Chen, Tomohiro Fukuda, Nobuyoshi Yabuki</i>	89
A Natural Human-Drone Interface For Better Spatial Presence Experiences <i>Chun-Yu Lo, June-Hao Hou</i>	99
<b>Generative, Algorithmic &amp; Evolutionary Design</b>	<b>109</b>
Calibrating a Formfinding Algorithm for Simulation of Tensioned Knitted Textile Architectural Models <i>Farzaneh Oghazian, Nathan Brown, Felecia Davis</i>	111

OptiGAN: Topological Optimization in Design Form-Finding With Conditional GANs <i>Xuyou Yang, Ding Wen Bao, Xin Yan, Yucheng Zhao</i>	121
Visual Character Analysis within Algorithmic Design: Quantifying Aesthetics Relative to Structural and Geometric Design Criteria <i>Robert Stuart-Smith, Patrick Danahy</i>	131
Machine Learning Modeling and Genetic Optimization of Adaptive Building Façade Towards the Light Environment <i>Yuanyuan Li, Chenyu Huang, Gengjia Zhang, Jiawei Yao</i>	141
Reinforcement Learning-Based Generative Design Methodology for Kinetic Façade <i>Sida Dai, Michael Kleiss, Mostafa Alani, Nyoman Pebryani</i>	151
Parasite City: Retaining the Industrial District of Alexandria, Sydney as an Integral Part of Urban Regeneration <i>Guoyi Chen, Seungcheol Choi, Mohammed Makki, Jordan Mathers</i>	161
A Generative Design Approach to Urban Sustainability Rating Systems During Early-Stage Urban Development <i>Or Moscovitz, Shany Barath</i>	171
An Alternative Model for Urban Renewal: A Generative Approach to the (Re)-Development of Xian Village <i>Ling Kit Cheung, Zhitao Xu, Pei Chen, Mohammed Makki</i>	181
Optimizing Design Circularity: Managing Complexity in Design for Circular Economy Through Single and Multi-Objective Optimisation <i>F. Peter Ortner, Jing Zhi Tay</i>	191
PL-System: Visual Representation of Pattern Language using L-System <i>Junah Yu, Deedee Min</i>	201
<b>Artificial Intelligence and Machine Learning in Design</b>	<b>211</b>
Where Will Romance Occur, A New Prediction Method of Urban Love Map through Deep Learning <i>Zhiyong Dong, Jinru Lin, Siqi Wang, Yijia Xu, Jiaqi Xu, Xiao Liu</i>	213
A Machine-Learning Approach to Urban Design Interventions In Non-Planned Settlements <i>Anna Boim, Jonathan Dortheimer, Aaron Sprecher</i>	223
Energy-driven Intelligent Generative Urban Design, Based on Deep Reinforcement Learning Method With a Nested Deep Q-R Network <i>Chenyu Huang, Gengjia Zhang, Minggang Yin, Jiawei Yao</i>	233
Data-Driven Evaluation of Streets to Plan for Bicycle Friendly Environments: A Case Study of Brisbane Suburbs <i>Gabrielle Toohey, Tommy Bao Nghi Nguyen, Ritva Vilppola, Waishan Qiu, Wenjing Li, Dan Luo</i>	243

Transit Oriented Development Assistive Interface (TODAI) - A Machine Learning Powered Computational Urban Design Tool for TOD <i>Garry Hangge Zhang, Leo Lin Meng, Nicole Gardner, Daniel Yu, Matthias Hank Haeusler</i>	253
Perceiving Fabric Immersed in Time, an Exploration of Urban Cognitive Capabilities of Neural Networks <i>Zhiyong Dong</i>	263
An Assessment of Tool Interoperability and its Effect on Technological Uptake for Urban Microclimate Prediction with Deep Learning Models <i>Nariddh Khean, Serjoscha Düring, Angelos Chronis, Reinhard König, Matthias Hank Haeusler</i>	273
A Machine Learning Evaluation Method for Sustainability Evaluation: The Case of 'Neighbourhoods' Design <i>Noam Raanan, Hatzav Yoffe, Jacob Grobman</i>	283
UAV-based People Location Tracking and Analysis for the Data-Driven Assessment of Social Activities in Public Spaces <i>Jeroen van Ameijde, Carson Ka Shut Leung</i>	293
Urban Scale 3 Dimensional CFD Approximation Based on Deep Learning A Quick Air Flow Prediction for Volume Study in Architecture Early Design Stage <i>Yahan Xiao, Akito Hotta, Takaaki Fuji, Naoto Kikuzato, Kensuke Hotta</i>	303
Synthetic Machine Learning for Real-time Architectural Daylighting Prediction <i>Rutvik Deshpande, Maciej Nisztuk, Cesar Cheng, Ramanathan Subramanian, Tejas Chavan, Camiel Weijnenberg, Sayjel Vijay Patel</i>	313
Autocompletion of Floor Plans for the Early Design Phase in Architecture: Foundations, Existing Methods, and Research Outlook <i>Viktor Eisenstadt, Jessica Bielski, Burak Mete, Christoph Langenhan, Klaus-Dieter Althoff, Andreas Dengel</i>	323
Design Intents Disentanglement: A Multimodal Approach for Grounding Design Attributes in Objects <i>Manuel Ladron de Guevara, Alexander Schneidman, Daragh Byrne, Ramesh Krishnamurti</i>	333
Rhetoric, Writing, and Anexact Architecture: The Experiment of Natural Language Processing (NLP) and Computer Vision (CV) in Architectural Design <i>Yuxin Lin</i>	343
Is Language All We Need? A Query Into Architectural Semantics Using a Multimodal Generative Workflow <i>Daniel Bolojan, Emmanouil Vermissou, Shermeen Yousif</i>	353
Deep Learning-Based Surrogate Modeling for Performance-Driven Generative Design Systems <i>Shermeen Yousif, Daniel Bolojan</i>	363

Bubble2Floor: A Pedagogical Experience With Deep Learning for Floor Plan Generation <i>Pedro Veloso, Jinmo Rhee, Ardavan Bidgoli, Manuel Ladron de Guevara</i>	373
Asynchronous Digital Participation in Urban Design Processes: Qualitative Data Exploration and Analysis With Natural Language Processing <i>Cem Ataman, Bige Tunçer, Simon Perrault</i>	383
A Data-Driven Workflow for Modelling Self-Shaping Wood Bilayer, Utilizing Natural Material Variations with Machine Vision and Machine Learning <i>Zuardin Akbar, Dylan Wood, Laura Kiesewetter, Achim Menges, Thomas Wortmann</i>	393
Quantifying the Intangible, A Tool for Retrospective Protocol Studies of Sketching During the Early Conceptual Design of Architecture <i>Jessica Bielski, Christoph Langenhan, Christoph Ziegler, Viktor Eisenstadt, Andreas Dengel, Klaus-Dieter Althoff</i>	403
<b>Urban Analytics, Urban Modelling and Smart Cities</b>	<b>413</b>
Evolutionary Design of Residential Precincts, A Skeletal Modeling Approach for Generating Building Layout Configurations <i>Likai Wang, Patrick Janssen, Kian Wee Chen</i>	415
POI Data Versus Land Use Data, Which Are Most Effective in Modelling Theft Crimes? <i>Jiajia Feng, Yuebing Liang, Qi Hao, Ke Xu, Waishan Qiu</i>	425
Realtime Urban Insights for Bottom-up 15-minute City Design <i>Cesar Cheng, Yuke Li, Rutvik Deshpande, Rishan Antonio, Tejas Chavan, Maciej Nisztuk, Ramanathan Subramanian, Camiel Weijenberg, Sayjel Vijay Patel</i>	435
Presence Stickers: A Seamlessly Integrated Smart Living System at a Solitary Elderly Home <i>Chor-Kheng Lim</i>	445
Quantifying the Imbalance of Spatial Distribution of Elderly Service with Multi-source Data <i>Pixin Gong, Xiaoran Huang, Chenyu Huang, Marcus White</i>	455
Machine-Reading Places & Spaces: Generative Probabilistic Modelling of Urban Thematic Zones & Contexts <i>Nuozhi Liu, Immanuel Koh</i>	465
Morphological Regeneration of the Industrial Waterfront Based on Machine Learning <i>Shuyi Huang, Hao Zheng</i>	475



Data-driven Research on Street Environmental Qualities and Vitality Using GIS Mapping and Machine Learning, a Case Study of Ma On Shan, Hong Kong <i>Xinyu Liu, Jeroen van Ameijde</i>	485
Evaluating the Accessibility of Amenities toward Walkable Neighbourhoods: an Integrated Method for Testing Alternatives in a Generative Urban Design Process <i>Sifan Cheng, Carson Ka Shut Leung, Jeroen van Ameijde</i>	495
Elemental Motion in Spatial Interaction (EMSI): A Framework for Understanding Space through Movement and Computer Vision <i>Margaret Z. Zhou, Shi Yu Chen, Jose Luis García del Castillo y López</i>	505
Retail Commercial Space Clustering Based on Post-carbon Era Context: A Case Study of Shanghai <i>Qinyu Cui, Shuyu Zhang, Yiting Huang</i>	515
Digital Twin-Based Resilience Evaluation of District-Scale Archetypes <i>Pradeep Alva, Martin Mosteiro-Romero, Clayton Miller, Rudi Stouffs</i>	525
An Integrated Parametric Generation and Computational Workflow to Support Sustainable City Planning <i>Hang Xu, Tsung-Hsien Wang</i>	535
Environmental Performance Assessment - The Optimisation of High-Rises in Vienna <i>Serjoscha Duering, Theresa Fink, Angelos Chronis, Reinhard König</i>	545
Automated Semantic SWOT Analysis for City Planning Targets: Data-driven Solar Energy Potential Evaluations for Building Plots in Singapore <i>Ayda Grisiute, Zhongming Shi, Arkadiusz Chadzynski, Heidi Silvennoinen, Aurel von Richthofen, Pieter Herthogs</i>	555
Quantifying the Coherence and Divergence of Planned, Visual and Perceived Streets Greening to Inform Ecological Urban Planning <i>Qing Yang, Chufan Cao, Haimiao Li, Waishan Qiu, Wenjing Li, Dan Luo</i>	565
The Effect of Path Environment on Pedestrians' Route Selection: A Case Study of University of Cincinnati <i>Jing Tian, Ming Tang, Julian Wang</i>	575
A Framework for a Gameful Collective Urbanism Based on Tokenized Location Data and Liquid Democracy: Early Prototyping of a Case Study Using E-bikes <i>Salma Tabi, Yasushi Sakai, Nguyen Tung, Masahiro Taima, Aqil Cheddadi, Yasushi Ikeda</i>	585
Sensing the City: Leveraging Geotagged Social Media Posts and Street View Imagery to Model Urban Streetscapes Using Deep Neural Networks <i>Jayedi Aman, Timothy C Matisziw, Jong Bum Kim, Dan Luo</i>	595

Impact of Covid-19 on Associations between Land Use and Bike-Sharing Usage <i>Yudi An</i>	605
Measuring Resilient Communities: an Analytical and Predictive Tool <i>Silvio Carta, Tommaso Turchi, Luigi Pintacuda</i>	615
A Web-based Interactive Tool for Urban Fabric Generation: A Case Study of Chinese Rural Context <i>Qiyang Zhang, Biao Li, Yichen Mo, Yulong Chen, Peng Tang</i>	625
Rebugging the Smart City. Design Explorations of Digital Urban Infrastructure <i>Nick Förster, Gerhard Schubert, Frank Petzold</i>	635
Machine Learning-Based Walkability Modeling in Urban Life Circle <i>Pixin Gong, Xiaoran Huang, Chenyu Huang, Marcus White</i>	645
Addressing Flood Resilience In Jakarta's Kampung Through The Use Of Sequential Evolutionary Simulations <i>Kim Ricafort, Ethan Koch, Mohammed Makki</i>	655
<b>Digital Heritage</b>	<b>665</b>
Exploring the Topological System of Dougong <i>Han-Ting Lin, June-Hao Hou</i>	667
Geelong Digital Outdoor Museum (GDOM) - Photogrammetry as the Surface for a Portable Museum <i>Domenico Mazza, Tuba Kocaturk, Sofija Kaljevic</i>	677
Revisiting Shoei Yoh: Developing a Workflow for a Browser-Based 3D Model Environment to Create an Immersive Digital Archive <i>Nicole Gardner, Matthias Hank Haeusler, Daniel Yu, Jack Barton, Kate Dunn, Tracy Huang</i>	687
Learning from Hale: An Educational Augmented Reality Application for an Indigenous Hawaiian Architecture <i>Richard Robinson, Hyoung-June Park</i>	697
Conflict and Reconciliation Between Architectural Heritage Values and Energy Sustainability, A Case Study of Xidi Village Anhui Province <i>Zhixian Li, Xiaoran Huang, Szymon Ruszczewski</i>	707
Use of Object Recognition AI in Community and Heritage Mapping for the Drafting of Sustainable Development Strategies Suitable for Individual Communities, With Case Studies in China, Albania and Italy <i>Francesco Grugni, Marco Voltolina, Tiziano Cattaneo</i>	717
Deep Architectural Archiving (DAA), Towards a Machine Understanding of Architectural Form <i>Frederick Chando Kim, Jeffrey Huang</i>	727

- Web-Based Three-Dimensional Augmented Reality of Digital Heritage for  
Nighttime Experience 737  
*Tomohiro Fukuda, Shiho Nagamachi, Hoki Nakamura, Yuji Yamauchi,  
Nao Ito, Shunta Shimizu*
- Robotic Carving Craft, Research on the Application of Robotic Carving  
Technology in the Inheritance of Tradition-AI Carving Craft 747  
*Zhe Guo, Zihuan Zhang, Ce Li*
- Remembering Urban Village: Using CloudXR Technology as an Enhanced  
Alternative to Better Disseminate Heritage 757  
*Xinyi Zhou, Yao Chen, Fukai Chen, Kan Li, Tian Tian Lo, Rufeng Xiang,  
Liquan Liu*



# **Virtual and Augmented Reality**



## **KURASHIKI VIEWER:**

### *Qualitative Evaluations of Architectural Spaces inside Virtual Reality*

TARO NARAHARA<sup>1</sup>

<sup>1</sup>*New Jersey Institute of Technology.*

<sup>1</sup>*narahara@njit.edu, 0000-0002-6972-8898*

**Abstract.** This paper discusses how virtual reality (VR) environments can be employed as a data collection tool beyond visualization and representation tools through a simple experiment in a VR space and speculates about its potential applications. Using a VR model that runs on a web browser based on an existing historic town in Japan called Kurashiki, the experiment asked 30 recruited participants to freely walk around and leave ratings on a 5-point scale on any buildings or objects appealing to them. The proposed system in this paper can display points of interest of multiple participants using heatmaps superimposed on a map that can help users visually understand statistical preferences among them. The project's goal is to provide a quantitative means for qualitative values of architectural and urban spaces, making such data more shareable. We intended to show that such a platform could help multiple stakeholders reach better consensus and possibly collect training datasets for machine learning models that could extract features related to the attractiveness in architecture and urban spaces.

**Keywords.** Virtual Reality; Crowdsourcing; SDG 10; SDG 11.

## **1. Introduction**

In recent years, Virtual Reality (VR) is increasingly more accessible, and its visual quality and experience in virtual 3-D environments have become more realistic with development in technology. As Ivan Sutherland pointed out in his article (1970), the challenge in computer graphics has been to make the virtual world look more real, move, and respond to interaction in real-time, and even feel real.

Since the 1990s, the adaptation of VR technologies has been increased through applications in vehicle simulation, entertainment, architectural design, and spatial arrangement to increase productivity, improve team communication, and reduce costs. For example, NASA has used VR for over 30 years for its astronaut training (McGreevy, 1993), and biomedical engineers have used games in VR for rehabilitations and therapies (Holden, 2005). Over the last 30 years, the graphics rendering system has improved with the innovations in both software and hardware. For example, NVIDIA corporation introduced a real-time ray-tracing SDK in 2019,

widely applied to VR game engine tools and minimizing visual gaps between real and virtual worlds.

Today, in the architecture and real estate industries, VR technology has been widely applied for communications among designers, various engineers, occupants, and clients. Notably, the use of VR in real estate is primarily employed by buyers and clients making initial qualitative selections and judgments, including some subjective evaluations by them. Some leading real estate companies, such as Redfin (2021) in the U.S. and AtHome (2021) in Japan, have listed some of their properties with 3-D walkthrough home tours that allow potential clients to make initial decisions about their selections without physically being at the properties. Another leading company in the U.S., Zillow (2021), introduced their app for free to sellers, including homeowners, to capture their properties using their mobile phones, enabling sellers to post their properties' 3-D walkthroughs on their portal sites. In architecture, engineering, and construction (AEC) practices, a wide range of applications of VR have been reported (Johansson, M., Roupé, M., 2019), including design review sessions (Zaker and Coloma, 2018; Wolfartsberger, 2019). A recent online conference organized by VR game engine developers, Build Architecture 2021 (Epic Games, Inc. 2021), featured many use cases by major architectural corporations such as HOK and Zaha Hadid Architects. Several research works have focused on the visibility of occupants in architectural spaces, such as (Schwartz, 2021) and their perceptual needs in densified urban environments (Fisher-Gewirtzman and Polak. 2019).

Several previous works have studied subjective evaluations in architecture using 2-D images, such as property images, photos, satellite images, and floor plan images (Wang, et al., 2019, Narahara and Yamasaki, 2019). In terms of urban city-scale applications, there are some standardized quantified measures for urban living conditions, including some qualitative, subjective criteria such as safety and ease of access (Walk Score, Zillow Inc. 2021).

The strength of this paper's proposed approach is to use immersive 3-D environments that allow users to leave their feedback more intuitively through the proposed user interfaces. Moreover, users' feedback is associated with local 3-D locations where they are situated within VR, enabling more three-dimensional spatial characteristics analysis than using 2-D images. In addition, more open 3-D data for cities and buildings have been released, for example, Google Map, OpenStreetMap, and Project PLATEAU by the Ministry of Land, Infrastructure, Transport, and Tourism in Japan (2021).

With the recent exponential development in hardware and software for real-time 3-D visualization technologies, it is expected that VR environments will be further used for collecting data through human interactions with their near photo-realistic visual qualities beyond mere representation tasks. Such data can link spatial representations, such as rendered perspective images or spatial geometries, to certain characteristics of spaces quantified based on user feedback. The data could also be used as training datasets for machine learning models to predict these characteristics, including perceptive values found in architecture and urban spaces in VR environments.

Recently, there are some reliable existing network models for tasks, including object detections and Bounding Box Regression, such as Faster R-CNN (Ren et al., 2016) and YOLOX (Ge et al., 2021), that can estimate certain values associated with





Figure 1. A view of Kurashiki viewer app that allows users to post location-based ratings (left); a virtual model of Kurashiki town (middle) and the real-life photo of the town (right). (VR environments produced in VR Design Studio and F8VPS by Forum8 Co. Ltd. (2022))



Figure 2. (a) Participants' paths in red dashed lines; (b) heatmap of cumulative target scores where participants added heart marks and magnitude are normalized from 0 to 1 in scale with red being high and blue being low; (c) participants' positions and directions of views superimposed on b; (d) heatmap of cumulative scores based on where participants situated when they added heart marks.

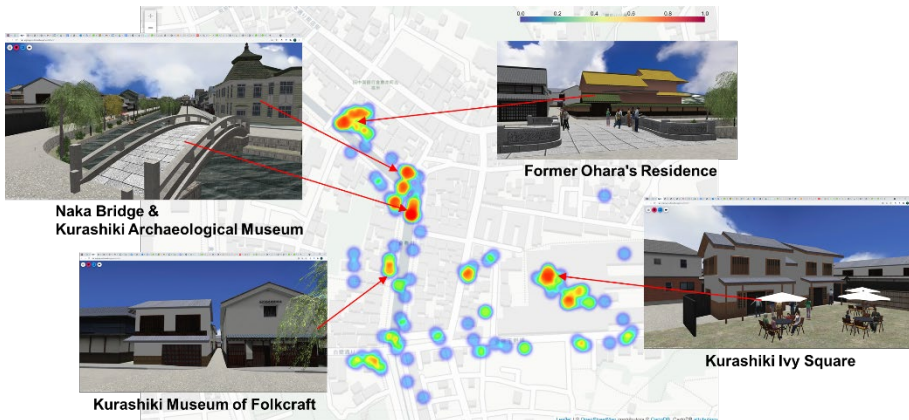


Figure 3. Target locations highlighted red in Figure 2b, where participants added higher scores with more frequencies, coincide with the locations of common tourist destinations in real-life. (VR environments produced in VR Design Studio and F8VPS by Forum8 Co. Ltd. (2022))

objects on images. In future practice, the use of the data retrieved from VR environments for training such models could be considered as a possibility. The paper further discusses how VR environments can be employed as a data collection tool for such applications in Section 3.

## 2. Methods

The proposed tool was built using an existing 3-dimensional real-time VR software package, VR Design Studio (2022) and its SDK offered by a Japanese software development company, Forum8 Co., Ltd. (2022), who supported and sponsored this project. With the support and help from Forum8's software developers, Forum8 Virtual Platform System (F8VPS, 2022) was also used to run the proposed application (app) on a web browser to record user feedback in addition to VR Design Studio. The software, VR Design Studio provides a series of VR models of existing cities in its library, and we used one of the VR models based on an existing preserved historic town on a canal in Japan called Kurashiki. The 3-D buildings on the site from its library were modeled and textured using street view photos of the real-life buildings to capture characteristics of the original buildings. To balance the speed of performance to run the app on a web browser, the real-time rendering is not at the highest level of photo-realistic quality compared to today's state of the arts for this experiment (Figure 1).

In total, 30 participants were recruited, and the following overall reasonably balanced attributes of participants were obtained. Eighteen participants identified themselves as male and 12 as female. Twenty-three participants are students in their 20s, of which 11 of them are in architecture major, and the others are not. Seven participants are in their 30s or over. In total, 18 participants are in the areas related to architecture, and 12 are in the areas non-related to architecture.

First, participants were provided with the link to the proposed application that runs on the Chrome browser on Windows 10 computers with a mouse and keyboard as user interfaces. Participants were asked to freely walk around the VR city as if they were visiting the city for sightseeing (None of the participants have ever visited the modeled site of Kurashiki city in their real life.) Then, they were asked to put heart marks on a scale of 1 to 5 on target objects on a screen whenever they found anything they liked, as many times as they preferred. The start position of the trip is fixed in the middle of the environment near the bridge for all participants in the app. There is no time limit for the experiment, they can finish anytime when they feel that they have completed their tour, and all participants finished within 15 minutes.

We added a function to create logs of the user position periodically each time the user has moved after a certain distance (over 1 meter) or the user has put the rating on a certain target position. The log files of the users were saved as JSON files after the users ended their tours and were collected. The logs of the user were recorded in the following format:

- Timestamp (elapsed time in millisecond since the start of the app)
- User position in (x, y, z)
- Target position in (x, y, z)
- Rating (Number of heart marks)

Extracted data from 30 participants were processed using Folium (2022), an open-source library for visualizing geospatial data built on Python and mapping capabilities of leaflet.js (a JavaScript library) and mapped on the OpenStreetMap using the latitude and longitude estimated from the 3-D model. Figure 2a shows 30 participants' paths in

red dashed lines superimposed in 50% semi-transparent color cumulatively. The result shows that participants' paths concentrated on the same walkways and bridges along the canal, which indicated that their travel paths overlapped quite a lot and went through similar paths even though they were not instructed to go any specific route. Figure 2b shows the heatmap of cumulative target scores where participants added heart marks and magnitude are normalized from 0 to 1 in scale with red being high and blue being low. Figure 2c shows participants' positions and directions of views when they added heart marks superimposed on 2b. Figure 2d shows the heatmap of normalized cumulative scores based on where participants were situated when they added heart marks.

### 3. Results and Discussion

Figure 2b shows that certain buildings in target locations selected by participants are more popular than other locations in the VR model as the heatmap shows concentrated areas. We can infer that there are common preferences among participants. In Figure 2b, participants added higher scores with more frequencies for three to four target locations. Those locations coincide with the locations of common tourist destinations in real-life, including Kurashiki Archaeological Museum, Former Ohara's Residence, Kurashiki Ivy Square, and Kurashiki Museum of Folkcraft (Figure 3).

The proposed app runs on any device supporting web browsers. The proposed app was developed using the cloud-based VR software package's platform and its SDK, allowing multiple users to participate and retrieve data automatically (Figure 4). Easy retrieval of big data from multiple locations and users can benefit the communication among different parties through quick derivations of visual results that reflect opinions from multiple users. Therefore, the platform could support consensus building in urban or architectural environments among different stakeholders. Since multiple users can participate online and virtually observe spatial conditions in architectural or urban environments under consideration, the app allows users to communicate their opinions after their tours through the proposed rating system and visualized results that statistically reflect collective opinions.

There are perceptual gaps between VR and real spaces, and the purpose of this experiment is not to validate that the results from VR spaces alone can deliver reliable results in real physical spaces. Further, inspection in the real spaces is required to validate any results based on subjective evaluations inside VR spaces. However, schematic experiments in VR spaces could be used to develop possible hypotheses for further validations in real spaces without constructing and preparing new physical spatial conditions or events, which could save costs for many practitioners. The software platform supports physical-based rendering technology, which could further minimize the gaps by using more sophisticated user interfaces for future experiments.

This time, the experiment used only one 3-D model of a city from Kurashiki and collected data from only 30 participants. However, this platform can be used with multiple city models and can acquire thousands of data into the cloud server with the appropriate consent from users and approval from IRB reviews (we have approval from the author's organization for this experiment on this paper.) The proposed platform can create a dataset that includes images (street views) labelled with

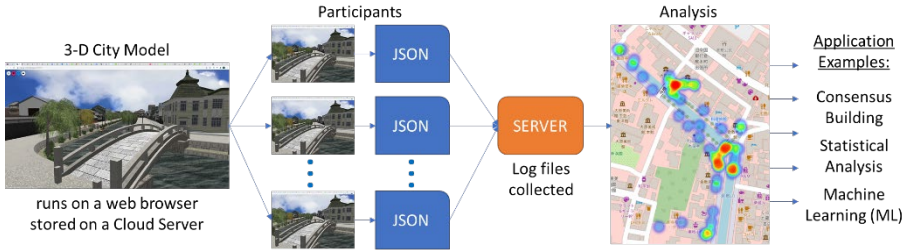


Figure 4. From log files of users and saved camera views with ratings of objects that are appealing to participants, annotated datasets can be created for training networks such as Faster R-CNN to predict estimated popularity from any new images. (VR environments produced in VR Design Studio and F8VPS by Forum8 Co. Ltd. (2022))

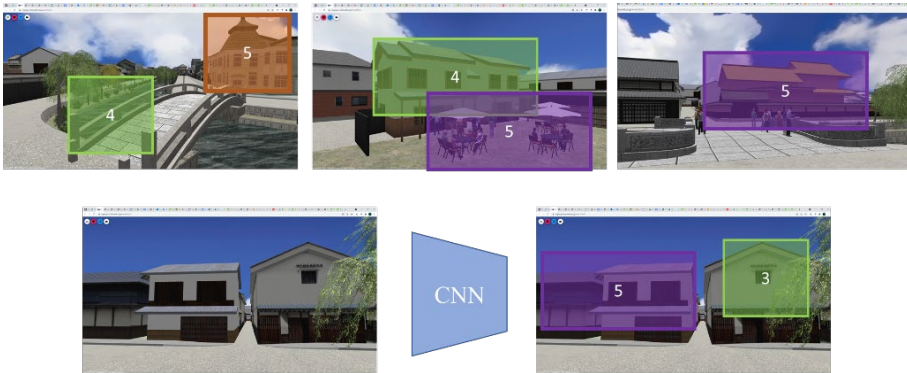


Figure 5. Images representing potential applications: From log files of users and saved camera views with ratings of objects that are appealing to participants, annotated datasets can be created for training networks using Bounding Box Regressors such as Fast R-CNN to predict estimated popularity from any new images. (VR environments produced in VR Design Studio & F8VPS by Forum8 Co. Ltd. (2022))

participants' review scores, which will provide training data for a Bounding Box Regression model using networks, such as Faster R-CNN (Girshick et al., 2014; Ren et al., 2016; Lee et al., 2019) and YOLOX (Ge et al., 2021), to predict estimated scores of objects that are more appealing to people from street views (Figure 5). The images can be retrieved from a recorded user and target positions in (x, y, z) with ratings provided by users. With a sufficient number of variations in users and city models, such datasets can train deep neural networks to predict estimated popularity and values for objects detected inside images for future applications to help planners estimate the values of their proposed spatial conditions yet to be constructed in the real world. With the improvements in real-time photorealistic rendering technology in the future, the gaps between real and virtual worlds are anticipated to be further minimized, contributing to opening new possibilities.

#### 4. Conclusions

This paper discussed how virtual reality environments could be employed as a data collection tool beyond the visualization and representation tools through a simple experiment in a VR space.

The project's goal is to provide a quantitative means for qualitative values of architectural and urban spaces, making such data more transparent and shareable. The app can display and visualize multiple users' statistical preferences to certain locations instantaneously from data collected from the app running on a web browser. We intended to show that such a platform could help multiple stakeholders or people from different backgrounds reach better consensus.

A platform such as ours could also be used to automatically collect annotated training datasets for machine learning models through interactions from multiple online participants. Understanding relationships between spaces and their implicit qualities in a quantitative data format would be essential to obtain machine learning models that could extract features related to their attractiveness found in architecture and urban spaces. In future studies, I plan to use the dataset to develop prediction tools for qualitative values from given spatial conditions.

Our cities are becoming more efficient using new hardware proposed by engineers, as we can see from recent developments in many proposed Smart Cities. Innovations in new energy harvesting technologies such as highly efficient solar cells and AI-inspired hardware technologies, including autonomous driving and smart home electronics, could make our cities and human settlements more sustainable and efficient.

Beyond simply providing efficiency and functionality, I aim to contribute the proposed methods to serve as the step to provide quantitative measures to assess qualitative values of cities. I hope that such tools can help incorporate more purpose and joy beyond efficiency into our living conditions.

#### Acknowledgements

The author wishes to acknowledge the full support received from Forum8 Co., Ltd., including the opportunity to use their software and guidance from their developers. The author would like to sincerely thank the software engineer at Forum8, Philippe de Quillacq, and Chief Manager of the VR Development Group, Yoann Pencreach, for their technical support on this project. The author would also like to thank Dr. Yoshihiro Kobayashi at Arizona State University for organizing the online virtual reality workshop event in summer 2021, which provided the opportunity to develop this project.

#### References

- AtHome Inc. (2021). *Lists of properties using VR (in Japanese)*. Retrieved November 1, 2021, from <https://www.athome.co.jp/chintai/theme/vr/>
- Epic Games, Inc. (2021). *Build: Architecture 2021*. Retrieved November 1, 2021, from <https://www.unrealengine.com/en-US/events/build-architecture-2021>.

- Fisher-Gewirtzman, D., & Polak, N. (2019). A learning automated 3D architecture synthesis model: demonstrating a computer governed design of minimal apartment units based on human perceptual and physical needs, *Architectural Science Review*, 62(4), 301-312.
- Forum8 Co., Ltd. (2022). *Forum8, a leader in photorealistic, immersive 3D VR and engineering software*. Retrieved January 31, 2022, from <https://www.forum8.com/>
- Forum8 Co., Ltd. (2022). *VR Design Studio*. Retrieved January 31, 2022, from <https://www.forum8.com/technology/vr-design-studio/>
- Forum8 Co., Ltd. (2022). *Forum8 Virtual Platform System (F8VPS)*. Retrieved January 31, 2022, from <https://www.forum8.co.jp/forum8/f8vps/>
- Ge, Z., Liu, S., Wang, F., Li, Z., & Sun, J. (2021). Yolox: Exceeding yolo series in 2021. *arXiv preprint arXiv:2107.08430*.
- Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. *In Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 580-587).
- HOK, Inc. (2018). *Introducing the HOK VR App*. Retrieved November 1, 2021, from <https://www.hok.com/news/2018-03/introducing-the-hok-vr-app/>
- Holden, M. K. (2005). Virtual environments for motor rehabilitation. *Cyberpsychology & behavior*, 8(3), 187-211.
- Johansson, M., & Roupé, M. (2019). BIM and Virtual Reality (VR) at the construction site, *CONVR 2019 Enabling digital technologies to sustain construction growth and efficiency* (pp. 1-10), Chalmers University of Technology.
- Lee S., Kwak S., & Cho M. (2019). Universal Bounding Box Regression and Its Applications. *In: Jawahar C., Li H., Mori G., Schindler K. (eds) Computer Vision – ACCV 2018*.
- McGreevy, M. W. (1993). *Virtual reality and planetary exploration*. *In Virtual Reality* (pp. 163-197). Academic Press.
- Ministry of Land, Infrastructure, Transport and Tourism, Japan (2021). *Project PLATEAU*. Retrieved November 1, 2021, from <https://www.mlit.go.jp/plateau/>
- Narahara, T., & Yamasaki, T. (2019). A Preliminary Study on Attractiveness Analysis of Real Estate Floor Plans, *GCCE, IEEE*, (pp. 445-446).
- Redfin Inc. (2021). *What is a 3D Walkthrough Home Tour? Virtually tour homes for sale anytime, anywhere*. Retrieved November 1, 2021, from <https://www.redfin.com/services/redfin-virtual-brokerage>
- Ren, S., He, K., Girshick, R., & Sun, J. (2016). Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence*, 39(6), 1137-1149.
- Schwartz, M. (2021). Human-centric accessibility graph for environment analysis. *Automation in Construction*, 127, 103557.
- Story, R. (2021). *Folium*. Retrieved November 1, 2021, from <http://python-visualization.github.io/folium/>
- Sutherland, I.E. (1970). Computer Displays, *Scientific American*, 222(6), June, pp. 57-81.
- Wang, X., Takada, Y., Kado, Y., & Yamasaki, T., (2019). Predicting the Attractiveness of Real-Estate Images by Pairwise Comparison using Deep Learning, *ICMEW, IEEE* (pp. 84-89).
- Wolfartsberger, J. (2019). Analyzing the potential of Virtual Reality for engineering design review. *Automation in Construction*, 104, 27-37.
- Zaker, R., & Coloma, E. (2018). Virtual reality-integrated workflow in BIM-enabled projects collaboration and design review: a case study. *Visualization in Engineering*, 6(1), 1-15.
- Zillow Inc. (2021). *Make your listing pop with Zillow 3D Home® tours*. Retrieved November 1, 2021, from <https://www.zillow.com/z/3d-home/>
- Zillow Inc. (2021). *Walk Score, Where to Live: City, Suburbs or Beyond?* Retrieved November 1, 2021, from <https://www.zillow.com/home-buying-guide/where-should-i-live/>

# VIRTUAL REALITY COLLABORATIVE PLATFORM FOR E-LEARNING: ANALYSIS OF STUDENT ENGAGEMENT AND PERCEPTIONS

JOOWON JEONG<sup>1</sup>, QINCHUAN CHEN<sup>2</sup>, NAYEON KIM<sup>3</sup> and HYUNSOO LEE<sup>4</sup>

<sup>1,2,3,4</sup>*Yonsei University, Seoul, South Korea*

<sup>1</sup>*jjeong20@yonsei.ac.kr; 0000-0002-4358-8815*

<sup>2</sup>*chenqinchuan@yonsei.ac.kr; 0000-0002-6221-5926*

<sup>3</sup>*ny.kim@yonsei.ac.kr; 0000-0002-3567-5708*

<sup>4</sup>*hyunsl@yonsei.ac.kr; 0000-0003-0023-88270*

**Abstract.** In this paper, we discuss the potential of using virtual reality collaborative platforms for e-learning to improve the quality of online education. First, we explore the characteristics of existing online platforms that can be used for e-learning. Second, we present a method for creating a Virtual Reality Collaborative Environment (VRCE) for e-learning using an online platform, namely FrameVR. Third, an experiment is conducted to investigate participants' behavioural and emotional engagement when using Zoom and the VRCE for online learning. Valid survey data from twenty-two participants are analysed. Then, participants are interviewed about their perceptions of using a VRCE for e-learning. The results of the experiment confirm that using a VRCE can increase student engagement, especially emotional engagement compared to Zoom. However, the findings also suggest that there is still room for improvement in the use of VRCE for e-learning. Therefore, further suggestions are made on the drawbacks of VRCE to improve the user experience. This paper provides insight into incorporating VRCE to enhance the e-learning experience and contribute to the development of online education.

**Keywords.** Virtual Reality Collaborative Environment; E-Learning; FrameVR; Online Education; SDG 4.

## 1. Introduction

“New normal” is a term that has been used frequently throughout the COVID-19 pandemic. In education, this “new normal” includes a sharp increase in the use of online learning platforms, as the pandemic has led to new ways of teaching and learning. In order to continue educating students during the pandemic, educational institutions around the world have had to rely on online learning platforms. One of the most popular platforms for online education, Zoom, has seen a dramatic increase in usage, from 10 million daily meeting participants in December 2019 to more than 300 million by mid-2020 (Statista, 2021). Similar platforms such as Skype have also seen

a significant increase in usage, with calls and meetings between Skype users increasing by 220% in recent months (Sherr, 2020). In today's world, e-learning has become a necessary resource for students and schools, and the demand for distance learning will continue to grow in the future (Gautam, 2020). While online learning has been a critical part of the educational process during the pandemic, some academics, as well as users, argue that it may have disadvantages. Most e-learning platforms offer only basic features such as video calls and text chats. Although such systems can satisfy basic communication needs, students often lack a sense of immersion and physical interaction with other users, leading to low student engagement, both behavioural and emotional (Rodgers, 2008). Studies have found that low student engagement can result in boredom, frustration, and low achievement (Lee, 2013). Since low student engagement is closely associated with low academic performance, it is important to find new strategies to improve engagement in e-learning.

In recent years, Virtual Reality (VR) technology has been extensively promoted and accepted as an important development in facilitating lifelong education (Alexiou et al., 2004). There is ample evidence that VR technology is an effective visualization tool that provides users with a more immersive experience, enables them to interact with other users in the virtual environment with a high sense of realism, enhances their creativity, and improves their work efficiency (Bucea-Manea-Țoniș et al., 2018; Shen et al., 2019). As a result, VR technology has great potential to support online learning and improve students' experience. Various cloud-based platforms have been developed that allow users to manage online activities in the virtual environment. However, these VR platforms are mostly used for work and commercial purposes (Jaehning, 2021). The value of adopting these platforms for e-learning has not been considered as frequently. This paper, therefore, aims to explore the potential of using VR collaborative platforms for online learning, following three steps: First, we examine and compare the key features of existing VR platforms that can be used for e-learning; second, we select a VR platform, FrameVR, and use it to create a VR collaborative environment (VRCE), with which we conduct an experiment to investigate whether the adoption of VRCE for online education can improve student engagement compared to 2D video-based platforms such as Zoom; third, we carry out interviews after the experiment in order to understand students' views on using VRCE for e-learning.

## **2. VR Collaborative Online Platforms**

### **2.1. FEATURES OF VR COLLABORATIVE PLATFORMS**

We examined eight online VR collaborative platforms which are acknowledged by mainstream media, such as Forbes (Forbes, 2020). These platforms were reviewed and compared based on two main criteria: (1) general features and (2) VR feature availability. In addition to general features such as screen sharing, a notable feature of these platforms is the availability of avatars. Unlike 2D video-based platforms such as Zoom; some VR platforms allow users to create their own 3D avatars before joining a VR online meeting. These avatars represent the individual participant and can be controlled to interact with other users and the virtual environment. Not only does the use of avatars provide an entertaining experience for users, but there is also clear evidence that avatar-based virtual events provide other advantages. For example,



research by Casanueva and Blake (2001) has shown that using avatars in virtual environments can create a high degree of co-presence, that is, the feeling that one is in the same place as the other participants and interacting with real people. In addition, another study has found that the use of avatars can help users who are reluctant to express themselves in physical spaces to present their opinions in virtual spaces (Blake and Moseley, 2010). Furthermore, platforms such as FrameVR and Mozilla Hubs allow users to customize their virtual environment by uploading a model created with 3D programs such as SketchUp and Blender. Users can create 3D objects and import them into the virtual environment. This feature can be helpful for lectures in subjects such as architecture. Here, a building model can be presented to students virtually so that they can understand the building structure more easily and directly (Arslan and Dazkir, 2017). The availability of the features of the various online platforms is summarized in Table 1.

Table 1. Existing VR collaborative online platforms

Platform	General Features	VR Features Availability		
	Main Content Sharing Methods	3D Object		Environment
		Avatar	Custom	Custom
Shapspark	Hyperlinks, uploading files (PDFs, video), material configurator	Limited	-	Limited
FrameVR	Whiteboard, files sharing (PDFs, 2D images, video, audio), text/ audio chat, screen sharing	√	√	√
Mozilla Hubs	Whiteboard, 2D images, video, screen sharing, text/ audio chat	√	√	√
MootUP	Images, video, audio sharing, desktop sharing	√	√	Bespoke
Spatial	Handwriting, post-it notes, screen sharing, uploading files (PDFs, images, video)	√	√	Bespoke
MeetinVR	Whiteboard, post-it notes, uploading files (subscription required), text/ audio chat	√	√	-
Glue	Uploading files (images, video), text/ audio chat	√	√	√
Engage	Video, website streaming, text/ audio chat	√	√	Bespoke

## 2.2. CREATION OF A VRCE FOR E-LEARNING

While most VR collaborative platforms have built-in environments that allow users to easily and effortlessly launch online events, these built-in settings are often limited. Users can only select a few types of environments from the built-in sources (e.g., museum, gallery, and office). Finding an appropriate environment for an activity with a specific purpose (e.g., an online course) might be difficult for users. In addition, these built-in settings often offer only basic features such as screen sharing, which may not be sufficient for managing more complex activities. Due to the drawbacks of built-in settings, we designed a VRCE specifically for educational purpose using a platform that supports custom environments. In this case, we used FrameVR because it is more accessible compared to other platforms. It does not require user registration and can be

used for free. The procedure for creating a VRCE using FrameVR is shown in Figure 1: First, we created a 3D model of the lecture hall using Blender, a modeling tool that best supports the GLB format required by FrameVR. Next, the model in GLB format was uploaded to the FrameVR cloud and converted to a virtual environment. Then, we set up the environment with built-in features that can be used for online courses (e.g., image sharing and whiteboards).

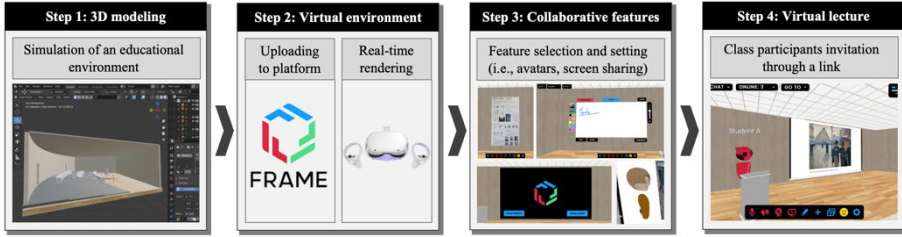


Figure 1. Process of creating a VR collaborative environment for e-learning

### 3. Methods

An experiment was conducted to investigate whether the use of a VRCE in online education can improve student engagement compared to a 2D video-based platform such as Zoom. We simulated the scenario of an online course under two different conditions (1) a VRCE and (2) Zoom. Each participant experienced both conditions in random order. A questionnaire survey was distributed to measure their engagement. Then, we interviewed participants about their perceptions of the advantages and disadvantages of using a VRCE for online courses and how they would like to see the VRCE evolve to improve the user experience.

#### 3.1. MATERIALS AND EXPERIMENTAL SETUP

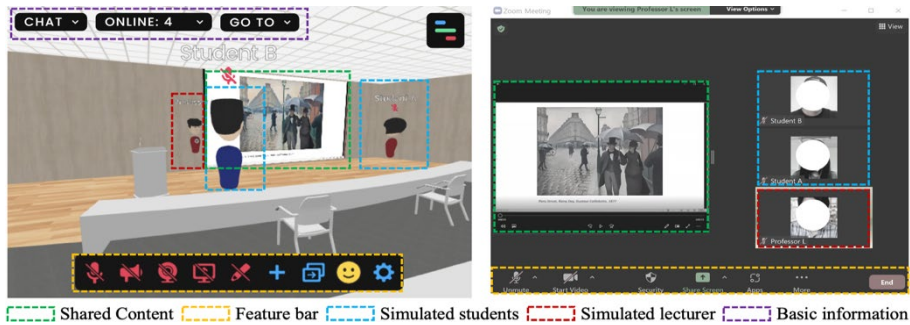


Figure 2. Experimental conditions and user interface, VRCE (left) and Zoom (right)

To ensure realistic experience, two simulated students and one simulated instructor were placed in each condition. The user interfaces of the two experimental conditions are shown in Figure 2. The VR tool used for the experiment was Oculus Quest 2. The Oculus Quest 2 system consists of a head-mounted device and two wireless controllers.

The computer used to access FrameVR and share/control VR in real-time was a Gigabyte RP75 with Intel Core i7 processor and a NVIDIA Geforce RTX 3070 graphics card. The Zoom version was 5.8.3.

### 3.2. QUESTIONNAIRE

Student engagement in both experimental conditions was measured using a self-report questionnaire. Existing literature on student engagement views it as a "meta-construct" that includes *behavioural* (e.g., participation, effort, and attitude), *emotional* (e.g., positive reactions to school and a sense of belonging), and *cognitive* (e.g., willingness to continue working on an assignment) engagement (Finn, 1989). However, Skinner et al. (2008) found that in-class student engagement is only reflected in their behaviour and emotions. Since this paper relates to student engagement during the e-learning process, only behavioural and emotional engagement are discussed here. To ensure validity and reliability, all the measurement items were taken from previous literature (Skinner et al, 2009). A total of ten items were used, five of which measured participants' behavioural engagement, while the remaining five measured emotional engagement. The items were measured on a seven-point Likert scale. The details of these items can be found in Table 2.

Table 2. Student engagement questionnaire items

Factors	Item No.	Item Description
Behavioural Engagement	BE1	I could freely communicate with others.
	BE2	During the lecture, I wanted to express my opinions to others.
	BE3	During the lecture, I wanted to ask questions to an instructor.
	BE4	During the lecture, I listened to the lecture very carefully.
	BE5	Overall, I could fully concentrate on the lecture.
Emotional Engagement	EE1	I felt enjoyment in this environment.
	EE2	I felt interested during the lecture.
	EE3	I felt comfortable during the lecture.
	EE4	I got involved during the lecture.
	EE5	Overall, listening to the lecture was fun.

### 3.3. PARTICIPANTS

This experiment targeted students who have used and are familiar with e-learning systems. Participants were recruited through posts in university group chats. A total of twenty-two participants were recruited for the experiment. The personal details of the recruited participants are as follows (1) age: from 25 to 42 (M = 29); (2) degree: four of them are undertaking undergraduate degree and eighteen of them are studying a master's degree or Ph.D.; (3) major: twenty of them are from departments related to interior design, two participants majored in architectural engineering; (4) gender: thirteen participants are female, nine participants are male.

### 3.4. EXPERIMENT PROCEDURE

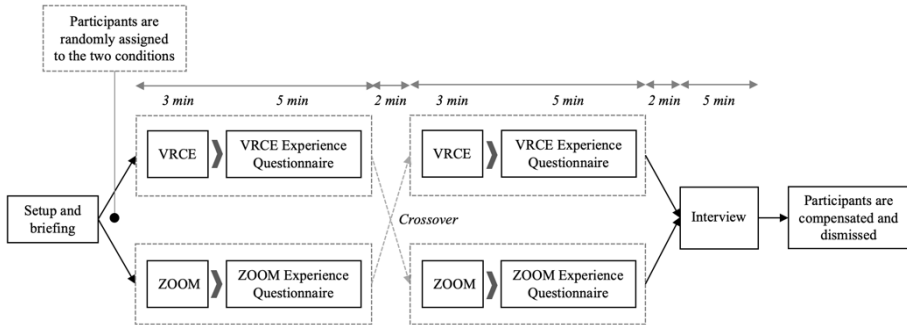


Figure 3. Experiment procedure

As shown in Figure 3, participants were randomly assigned to one of the experimental situations (VRCE then Zoom or Zoom then VRCE). This procedure was conducted following similar experimental studies by Azmi et al. (2020) and Hong et al. (2019) in order to eliminate any sequence effects. In previous studies, a two-minute pause between environmental situations was recommended to reduce the carryover effect. During the experiment, participants were asked to pretend to attend an online lecture. A three-minute pre-recorded lecture video was shared and played to simulate the course content. To encourage interaction, we included a quiz in the video. The quiz was suggested based on the video content, and participants were asked to answer the quiz verbally if they wished. It should be noted that participants wore a head-mounted device during the VRCE (Figure 4). After participants experienced each situation, they were given a questionnaire to measure their sense of engagement. After they completed the questionnaire, we conducted a short interview in which we asked them about their perceptions of the advantages and disadvantages of both methods for online courses and how they would like to see the VRCE evolve to improve the user experience. Upon completion of the interview, participants were rewarded and dismissed.



Figure 4. Participant experiencing two experimental conditions, VRCE (left) and Zoom (right)

## 4. Results

### 4.1. STUDENT ENGAGEMENT

Valid survey data from twenty-two participants were analysed using Jamovi (ver.1.6). We used descriptive statistics, reliability analysis, and the t-test to assess if there were significant differences in student engagement between VRCE and Zoom. A p-value < 0.05 was considered statistically significant. Internal consistency of the measurement scale was checked by Cronbach's alpha values. The results of the reliability test show that the Cronbach's alpha of all variables was above 7.0: VRCE-BE (0.7793), VRCE-EE (0.7979), Zoom- BE (0.8777), Zoom- EE (0.7861). Therefore, all scales used in the study were internally consistent and reliable. The results of the t-test comparing behavioural engagement between VRCE and Zoom showed that BE1 and BE2 were significantly different: BE1 ( $t = .8742$ ,  $p < 0.05$ ), BE2 ( $t = 2.5578$ ,  $p < 0.005$ ), stating that participants showed more positive attitudes toward communicating and expressing their ideas when using a VRCE for online courses. It was also found that all emotional engagement variables were significantly different in the two settings: EE1( $t = 7.5450$ ,  $p < 0.001$ ), EE2( $t = 4.9271$ ,  $p < 0.001$ ), EE3( $t = -2.3090$ ,  $p < 0.05$ ), EE4( $t = 2.4081$ ,  $p < 0.05$ ), EE5 ( $t = 4.0005$ ,  $p < 0.005$ ), indicating that participants showed a higher sense of enjoyment and involvement in e-learning in the VRCE.

### 4.2. STUDENT PERCEPTION

We conducted short interviews to explore participants' views on using (1) Zoom and (2) a VRCE for online learning. Although the majority of participants indicated that Zoom is an effective tool to support their learning process during the pandemic, some of them highlighted drawbacks. One of the most common responses from participants was related to the difficulty of interaction between instructors and students. For example, one participant reported, "when I use Zoom, I cannot physically interact with my instructors and fellow students. It is even more difficult to have a real conversation, so I feel isolated from the others" (P14). In addition, one interviewee noted, "I find it very difficult to concentrate in online lectures via Zoom...not being in a designated space like a classroom, I get distracted easily, especially when taking an online course from home" (P4).

While some participants identified that there were some barriers to using Zoom for online learning, a large percentage (81%) stated that using a VRCE enhanced their learning experience compared to Zoom. The advantages of VRCE were described with reference to several aspects. First, avatars played an essential role in creating social experiences for users. Some participants noted that avatar-based platforms helped them foster their sense of connectedness. They also found that they were able to express themselves more freely than when using Zoom for an online learning. One participant reported, "when I have questions, I find it difficult to ask for help when I use Zoom for online learning. But after the experience of attending lectures using VRCE today, I think avatar-based platforms help me to enhance the sense of connectedness and promote interaction with other avatars." (P8). Another participant pointed out that "VRCEs would make the learning experience richer, ... engaging ... When I see an avatar with a professor's sign, I could feel a higher sense of motivation and enjoyment

in learning” (P2). In addition, one of the reasons that nearly half of the participants (48%) gave regarding the benefits of employing VR in online learning was that a VRCE increases concentration, “puts learners in an immersive environment and provides an all-around realistic experience of a lecture. While [they] are easily distracted by the surroundings when using Zoom, the headset eliminates all external distractions and helps [them] concentrate” (P1). Despite these benefits of using a VRCE for e-learning, participants further made other suggestions to improve the learning experience. Suggestions mentioned by them were “interpersonal” (P10), “the inability to take notes” (P8), and “the bulkiness of the VR headset” (P3) as examples.

## 5. Discussion and Conclusion

### 5.1. DISCUSSION

The results show that students are more engaged in e-learning when using a VRCE setting compared to Zoom. Specifically, students are found to be more willing to communicate with others and express their opinions in a VRCE setting. In addition, using a VRCE for e-learning significantly improves students' emotional engagement. For online lectures, students perceive the VRCE setting to be more comfortable and enjoyable than Zoom. Although there is great potential for the use of VRCEs in online education, there are some areas of VR, which could be further improved by IT specialists. Based on our analysis of student feedback on the use of VRCEs for e-learning, there are three main ways to improve student experience. First, non-verbal cues could be added for better interaction between simulated users. Behavioural techniques such as gestures, eye contact, and facial expressions play an important role in communication and effective class management as they convey additional information to others (Petrie et al., 1998). Besides the possibility of using custom avatars, providing natural and realistic-looking avatars would be key to improving the user experience in VRCEs.

Second, it is suggested to provide note-taking features to help students attend lectures more efficiently. A study by Chen et al. (2019) investigated the use of an “interactive note-taking interface (iVRNote)” to address the challenge of taking notes while wearing a head-mounted display. The introduction of iVRNote in VRCEs would further enhance the learning experience for students. In addition to taking notes, other functions such as voice recording or screenshots could also be created in VRCEs.

Third, students would be more likely to accept VRCEs as a learning tool if the head-mounted display was lighter and smaller. However, this should not be seen as a hindrance as technology is constantly evolving. For example, VR glasses, such as the Huawei VR glasses, are being launched. Unlike the Oculus Quest, which combines a computing unit and a battery in the headset, the Huawei VR glasses do not contain such components (Skarredghost, 2021). Hence, it is important to highlight that technological development may contribute to the improvement of the online learning system, as e-learning takes into account both the technology and the user's perspective.

### 5.2. CONCLUSION

According to the United Nations, nearly 1.6 billion children were out of school by April

2020 due to the COVID -19 pandemic. Never before have so many children been out of school at the same time. Therefore, as part of the goal of quality education (SDG 4), it is important to help countries implement innovative solutions for delivering education remotely, leveraging high-tech, low-tech, and no-tech approaches to ensure that education never stops. In this context, e-learning is a promising solution to help students continue their studies during the pandemic. Currently, most e-learning systems offer only basic features such as video calls and text chats. Although such systems can meet basic communication needs, they often lack immersion and physical contact with other users, which can lead to low student engagement. In this paper, we present a method for creating a VRCE for e-learning using an online platform, namely FrameVR, and confirm that using a VRCE can make online courses more engaging. There is a great potential for using VR technology to enhance the e-learning experience.

Some limitations should be discussed. First, this study only examined a small number of participants. A more comprehensive experiment can be conducted in the future with a larger group of participants in different demographic contexts (e.g., age groups, degree levels). Second, participants experienced the pretend lecture through a three-minute video, which is a relatively short period of time compared to a real course. Participants' perceptions and attention problems might be less likely to manifest in such a short period of time. Future research could extend the duration of the simulated sessions to more accurately capture participants' responses. Third, as a pilot study to explore the potential of VR technology for e-learning, only a small number of participants (two simulated students per class) were included in the experiment. In a later experiment, more students can be included to simulate a realistic teaching scenario. We could also investigate how architecture-related students would perceive the use of other features, such as custom 3D objects, to enhance their learning experience. A more holistic approach to exploring the views of users of VRCEs would provide an opportunity to promote quality education alongside immersive learning experiences.

### **Acknowledgements**

This research was supported by the BK21 and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2020R1I1A1A01073447).

### **References**

- Alexiou, A., Bouras, C., Giannaka, E., Kapoulas, V., Nani, M., & Tsiatsos, T. (2004, March). Using VR technology to support e-learning: the 3D virtual radiopharmacy laboratory. In *24th International Conference on Distributed Computing Systems Workshops, 2004. Proceedings.* (pp. 268-273). IEEE.
- Arslan, A. R., & Dazkir, S. S. (2017). Technical Drafting and Mental Visualization in Interior Architecture Education. *International Journal for the Scholarship of Teaching and Learning*, 11(2), 2. <https://doi.org/10.20429/ijstl.2017.110215>.
- Azmi, A., Ibrahim, R., Abd Ghafar, M., & Rashidi, A. (2020). Improving empathic evaluations in virtual reality for marketing of housing projects. In *54th International*

- Conference of the Architectural Science Association, ANZAScA 2020* (pp. 600-609). The Architectural Science Association (ANZAScA).
- Bjekić, D., Obradović, S., Vučetić, M., & Bojović, M. (2014). E-teacher in inclusive e-education for students with specific learning disabilities. *Procedia-Social and behavioral sciences*, 128, 128-133.  
<https://doi.org/10.1016/j.sbspro.2014.03.131>.
- Blake, A. M., & Moseley, J. L. (2010). The emerging technology of avatars: some educational considerations. *Educational Technology*, 13-20.
- Bucea-Manea-Țoniș, R., Andronie, M., & Iatagan, M. (2018). e-Learning in the Era of Virtual Reality. *eLearning & Software for Education*, 1.
- Casanueva, J., & Blake, E. (2001). *The effects of avatars on co-presence in a collaborative virtual environment*.
- Chen, Y. T., Hsu, C. H., Chung, C. H., Wang, Y. S., & Babu, S. V. (2019, March). ivrnote: Design, creation and evaluation of an interactive note-taking interface for study and reflection in VR learning environments. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (pp. 172-180). IEEE.
- Fink, C. (2020, May 29). *Remote collaboration and Virtual Conferences, the future of work*. Forbes. Retrieved November 1, 2021, from <https://www.forbes.com/sites/charliefink/2020/05/27/remote-collaboration-and-virtual-conferences-the-future-of-work/?sh=651ebf7a6b9c>.
- Finn, J. D. (1989). Withdrawing from school. *Review of educational research*, 59(2), 117-142. <https://doi.org/10.3102/00346543059002117>.
- Fredricks, J. A., Blumenfeld, P. B., & Paris, A. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Education Research*, 74, 59-109.  
<https://doi.org/10.3102/00346543074001059>.
- Gautam, P. (2021, May 12). *Advantages and disadvantages of online learning*. *eLearning Industry*. Retrieved November 9, 2021, from <https://elearningindustry.com/advantages-and-disadvantages-online-learning>.
- Hong, T., Lee, M., Yeom, S., & Jeong, K. (2019). Occupant responses on satisfaction with window size in physical and virtual built environments. *Building and Environment*, 166, 106409. <https://doi.org/10.1016/j.buildenv.2019.106409>.
- Laurell, C., Sandström, C., Berthold, A., & Larsson, D. (2019). Exploring barriers to adoption of Virtual Reality through Social Media Analytics and Machine Learning—An assessment of technology, network, price and trialability. *Journal of Business Research*, 100, 469-474. <https://doi.org/10.1016/j.jbusres.2019.01.017>.
- Petrie, G., Lindauer, P., Bennett, B., & Gibson, S. (1998). Nonverbal Cues: The Key to Classroom Management. *Principal*, 77(3), 34-36.
- Rodgers, T. (2008). Student engagement in the e-learning process and the impact on their grades. *International Journal of Cyber Society and Education*, 1(2), 143-156.
- Shen, H., Zhang, J., Yang, B., & Jia, B. (2019). Development of an educational virtual reality training system for marine engineers. *Computer Applications in Engineering Education*, 27(3), 580-602. <https://doi.org/10.1002/cae.22099>.
- Sherr, I. (2020, March 30). *Microsoft's Skype sees massive increase in usage as coronavirus spreads...* CNET. Retrieved November 1, 2021, from <https://www.cnet.com/tech/mobile/microsofts-skype-sees-massive-increase-in-usage-as-coronavirus-spreads/>.
- Skinner, E., Furrer, C., Marchand, G., & Kindermann, T. (2008). Engagement and disaffection in the classroom: Part of a larger motivational dynamic. *Journal of educational psychology*, 100(4), 765. <https://doi.org/10.1037/a0012840>.



# COLLABORATIVE DESIGN REVIEW SESSIONS IN VIRTUAL REALITY: MULTI-SCALE AND MULTI-USER

SHAHIN SATEEI<sup>1</sup>, MATTIAS ROUPÉ<sup>2</sup> and MIKAEL JOHANSSON<sup>3</sup>

<sup>1,2,3</sup> *Construction Management, Chalmers University of Technology*

<sup>1</sup> *shahin.sateei@chalmers.com, 0000-0002-5606-9707*

<sup>2</sup> *mattias.roupe@chalmers.com, 0000-0002-3706-8485*

<sup>3</sup> *jomi@chalmers.com, 0000-0002-6108-8662*

**Abstract.** The use of Virtual Reality (VR) for design reviews in projects is becoming more common in construction. However, the use of VR in these processes has been limited to been used more as a complementary reviewing tool alongside information medias such as 2D drawings and 3D models. Furthermore, immersive VR has been argued to have limitations when it comes to orientation and understanding and reasoning about functional links between physical layouts in a facility. This paper presents a case study of a VR system used during design reviews that support end-users to switch between different representations and scale i.e., miniature model/bird-eye view, and a 1:1 scale experience of the facility. The data gathered, consisted of recorded observation of the VR based design review process and study what type of discussion and design errors that was found during two VR-workshops connected to a new elementary school. The result shows, that by supporting switching between miniature model and 1:1 scale VR experience facilitated spatial orientation and understanding and collaboration across disciplines in the project. The study also show how collaborative immersive VR can be used as an efficient communication-tool during the design process in a real-world project.

**Keywords.** Virtual Reality; VR; Immersive Virtual Environments; Collaboration; Levels of detail; SDG 4; SDG 9; SDG 11; SDG 12.

## 1. Introduction

A common challenge during the design process of a building is to provide a representational medium that facilitates understanding and communication among all involved stakeholders. For instance, clients and building end-user often have difficulties to fully understand and comprehend traditional design mediums such as 2D-plans, elevations, sketches, and 3D-models, yet their input and feedback can be crucial in order to obtain a high-quality outcome.

This is especially true for educational facilities where teaching methods and student

performance is linked to and affected by design choices and physical layout of the school (Frelin and Grannäs, 2020; Byers, Imms, and Hartnell-Young, 2018).

However, in this context, immersive Virtual Reality (VR) has been shown to offer an alternative or complementary representational medium (Mastrolembo Ventura et al., 2019; Sateei et al., 2021). In contrast to traditional design mediums, immersive VR allow all stakeholders to experience and comprehend future buildings from a self-centered, egocentric perspective, which better mimics a real-life experience. Especially for laypersons, such as building end-users, this has been shown to provide another level of understanding and perception of space (Abouelkhier, 2021; Johansson & Roupé, 2019; Roupé et al., 2019; Roupé et al., 2020).

Still, factors such as immature technology, client-contractor dynamics, requirements for implementation and structure and an overall lack of experience have prevented VR from being fully established (Delgado et al 2020). As such, it is important to provide real-world examples on how to successfully integrate VR beyond that of pilot testing (Teixeira et al. 2021). In addition, previous studies have identified limitations with VR compared to 2D drawings in that a user is more likely to get disoriented in a large VR-model, and that it might be less suitable for an easy overview and understanding of the big picture (Liu et al., 2020). With educational facilities being characterized by the importance of functional logistics and communication and flows between spaces, supporting better orientation and easy overview of large projects thus becomes an essential aspect to solve for immersive VR. For the specific case of head-mounted displays (HMDs) it is also important to consider how to support collaboration. With HMDs being primarily a tool for the individual, it makes it less suitable for active collaboration during design review sessions.

In this paper we investigate how collaborative immersive VR can be used as an efficient communication-tool during the design process. The setting is a real-world case where architects and client representatives used VR as the only medium during two design review sessions for a new elementary school. During these sessions empirical data was collected by means of video-recorded observations as well as interviews. In order to support a better overview and aid navigation between different locations in the VR-model we also explored the concept of having easy access to a miniature model within the VR environment. In particular, we are interested in the interplay and relationship between allocentric, and egocentric frames of reference offered by combining full-scale and miniature-scale representations within the same medium. Although previous research has investigated – and advocate – a combination of allocentric representations (e.g., 2D-plans or bird's eye view) and egocentric representations (e.g., immersive VR) throughout the design process (Roupé et al. 2019), far less has been explored in terms of offering these two representations at the same time within the same medium.

## 2. Method

The case study was to study and investigate a VR based design review process and study what type of discussion and design issues that were found during two VR-workshops connected to a new elementary school. The involved participants were involved and responsible for different parts of the design process e.g., a project manager, development managers (for the school), and an architect responsible for fixed

furnishing as well as an interior architect from the municipality of Gothenburg.

The idea of using VR was recognized when 2D and 3D models could not provide a sufficient level of spatial understanding among the project group.

The study was done on two separate occasions via VR-workshops (n=8 in the first session and n=6 on the second session) at Chalmers University of Technology. Given that the workshops were conducted shortly before construction was about to start, the participants wanted apart from the above-mentioned objective, also investigate how VR could facilitate the planning processes of new school premises from an end-user perspective. In this context, they wanted to investigate how it could mitigate the risk of costly reconstruction close after commissioning.

The workshops were conducted during the detail-design phase of the project with the key objective to use VR as a design review tool to discover any design error that would otherwise be difficult to discover and address with 2D drawings that were the main reviewing tool in the project. Before starting the first workshop, participants were briefed on the aim of the research and subsequently introduced to a VR environment (learning how to navigate and interact within the VR-model). Following this, participants were asked to review operational requirements related to the functionality and workflow. The second workshop was oriented around a more detailed design and room layout, in terms of having furnished spaces to review in the provided VR-models.

During these two workshops, at least one person from the research team was available for supervision and providing help in terms of showcasing how to navigate via the provided VR-hardware. The VR system that was used were three Oculus Rift S kits, together with the software BIMXplorer. The software supported direct import of IFC-files from the design process, without any need for further optimization. Additionally, there were functions such as a measuring tool, taking screenshots, markups, and support for multi-user collaboration. During the VR-workshops, the participants used mentioned hardware kits coupled with a big screen display and two laptop displays, see Fig 1.



*Figure 1. The set-up during the first workshop (left). Different participants viewing and discussing potential design issue areas during the second workshop (right).*

The methods used in this study were observations (video recorded) during the workshops and a follow up open discussion at the end of each VR-workshop where the participants reflected on their experience. The design error found by the participants have been categorised and mapped to the allocentric and ego-centric display-options enabled in BIMxplorer; a virtual miniature model of the building that enabled

participants to crosscut the virtual model to view the building from a bird-eye perspective (i.e., a digital version of a 2D view) and the 1:1 scale option typically associated with VR-models.

The result was analysed within the theoretical framework of egocentric and allocentric design space representation. This was done to discover patterns in terms of the design issues discovered and discussed among the participants during the workshops as well as to gain a better understanding for how a combination of these two design space representations help facilitate spatial understanding among the participants.



*Figure 2. Miniature model showcasing the building in its entirety as well as sectioning of the various floors reviewed by the participants.*

### 3. Results

The following section presents the design errors the participants found in each respective workshop session as well as in what view-modes in BIMxplorer they were found in (i.e., the miniature model view or the 1:1 scale). These tables will then in the subsequent analyzation chapter be further discussed. This is done to see how potential patterns among the discovered design faults can be found and how these relate to the allocentric, and egocentric design space representations described in the earlier sections.

As illustrated in table 1, most of the discussion and design changes were related to changing of the space and width and height and size of the rooms and corridors in the building due to better understanding and perception of the designed space. Other changes were sightlines pertaining to the visibility between the entrance of the school and the restaurant, given that it was important for the project members to understand the logistical flow between points in the building. The participants switched between the two different view-modes (miniature model and 1:1 scale) to better understand how design issues related to logistical workflow could be addressed. The logistical workflow was critical for the project, given that the different floors and parts of the building covered different school grades and it was important to keep classes separated from each other.

Further, the project leader participating in the workshop acknowledged during the open discussion taking place after the workshop, how the sightlines between the building entrance and restaurant as well as the placement of glass portion in the various

educational rooms facing the adjacent areas were previously in 2D not apparent to them:

"The design issues we discovered and discussed today were difficult to see let alone discuss about in the 2D drawings we have used up until now."

### 3.1. WORKSHOP 1

Table 1: showcasing types of design feedback and errors found during the first workshop.

Types of design feedback and errors found	Miniature model	Scale 1:1
<b>Change related to perception of room size and space</b>		
Disorderly placed furnishment restaurant		X
Too narrow of a serving space in the restaurant	X	X
Logistical flow in staircase between floor plans (e.g., width and landing size)	X	X
<b>Change related to functions in studied space</b>		
Accessibility and logistic flow connected to the counter in café/restaurant		X
Door placement and swing direction to the restaurant to be adjusted to enhance logistical flow to restaurant		X
Design error: Security glassed curtainwalls and door were missing between floor 3 and 4		X
<b>Change related to views &amp; sightlines</b>		
Ensured unstructured sightline between kitchen and restaurant	X	X
Glass in door between restaurant serving and tray line		X
Glass partition is missing in the teacher's room		X
Lower the sill-height of window in kitchen/office		X

A development manager also made a point about being able to discover bigger design issues at an earlier stage in the design process:

"Rather than trying to interpret and understand design issues by myself, I would preferably be able to communicate and resolve design issues together with different actors in different phases of the design process."

Lastly, the observations showed how the first session entailed a more exploratively driven approach by the participants, where the focus was more oriented towards learning how to navigate and use the interface. Consequently, the design issues detected and addressed in the first workshop were done in a non-systematic way with three participants being in the model at the same time and using for example the interface functionality of coordinating other participants to their own location once they detected something of interest, they deemed worth discussing (left picture in figure 1).

### 3.2. WORKSHOP 2

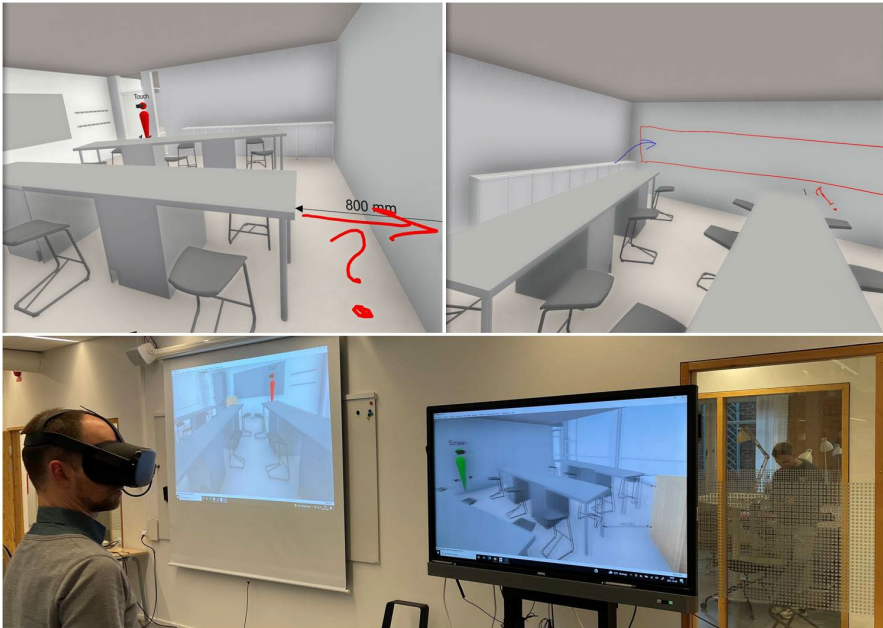
Types of design feedback and errors found	Miniature model	Scale 1:1
<b>Change related to perception of room size and space</b>		
Window recess		X
Radiators		X
Window recess done with stone slab on top of concrete wall		X
Can windows recess be used as seating area?		X
<b>Change related to functions in given space</b>		
Alternative placement of water dispenser	X	X
Teachers' workplace without daylight not a viable option		X
Placement and amount of storage cabinets in teaching room	X	X
Furnishment in majority of rooms is good	X	
Worktop alongside corridor may have to be done deeper		X
Shelves on wall in the school nurse's room must be stripped away		X
Potential inversion of furnishment in the school nurse's room	X	
<b>Change related to sightlines</b>		
Glass partition towards waiting area must be stripped away	X	X
Glass partition towards staff room must be stripped away → door is done with glass	X	X

Table 2: showcasing types of design feedback and errors found during the second workshop.

The VR-model used in the second workshop included furnishment, done by the interior architect. This led to an increase in number of questions regarding logistical flow, as well as participants switching more noticeably between the different view-modes, more so than the preceding workshop.

Further, design questions pertaining to functionalities of certain rooms were more tangible in comparison to the first workshop. Specifically, questions were oriented around how functionality of classrooms could be affected by the fixed furnishment (e.g., tables and lockers). The discussion that followed revolved around the cost of maintaining the current arrangement of the fixed furniture and how it would affect workflow as well as how it could lead to increase cost over time. One example of this is when one of the development managers suggested rearranging tables in such way that it would open up space from the teacher desk, seeing as the project and facility manager experienced it as too cramped. This was an important discovery as these fixed furnishings also included sanitary piping (e.g., laboratory equipment desks), which would have been very costly to change after construction, when the building was in operation.

"We can gain more space if we push the tables in the centre of the room against the wall and thereby making it less crowded for those sitting at the edge of the room. This should lead to a better logistical flow and avoiding potential risk of conflict."



*Figure 3. The development manager mentioned above, reviewing the placement of furnishment, illustrated in 1:1 and visible on the projector screen for other participants to discuss.*

Lastly, it was observed how the second workshop, in comparison to the first session, entailed a more systematic approach to detecting, discussing, and addressing design issues. With the first workshop revolving around three participants simultaneously interacting and collaborating in the same VR-model, the multi-user factor shifted over to discussion revolving around one user being in the model and another participant sporadically joining in the same discussion taking place or detecting issues in a different space on the same floor, as seen in figure 2. The participants not wearing the VR-goggles were instead discussing in front of the project screen, as seen on the right picture in figure 1.

## 4. Discussion

### 4.1. DESIGN CHANGES DUE TO VR DESIGN REVIEW

The observations made in both the workshops, showed participants quickly learning how to interact with each other via the multiuser functionality as well as interacting with the model via the various tools (e.g., using the measuring tool, taking screenshots, markups etc.). The participants expressed how these features enabled them to

collaborate and understand each other's point of view and spatial reasoning in a more participatory way (Van den Berg et al., 2017) in comparison to the 2D drawings used in the project up until the point of the workshops. This common frame of reference became tangible when design changes related to the furnishment of the facility were discussed and addressed during the second workshop and could be explained by the less demanding cognitive workload apparent when using VR over 2D drawings (Dadi et al., 2014; Liu et al., 2020; Roupé et al., 2014).

Moreover, it became clear that the addition of furnishment in the models resulted in a deeper and more focused discussion around the design errors that were discovered. This also reinforces the argument presented in similar studies where furnishment and functionality related details were deemed more important than photorealism (e.g., lightning, texture of material, colour scheme in the model) in terms of enhancing spatial understanding among end-users (Abouelkhier, 2021; Mastrolembo Ventura et al., 2019; Sateei, 2021) Furthermore, our result confirms previous research in that gradual increase in level-of-detail (LOD) level should be advocated instead of photorealism. Lastly, it could be argued that the amount of design changes addressed in the first workshop in comparison to the second workshop was mainly influenced by the addition of furnishment, which is also reflected in the higher amount of design issues discovered, discussed, and addressed in the second workshop. In addition, furnishment influences the context and meaning to the space and this in turn affect the operations of the educational facility (Frelin and Grannäs (2020).

#### 4.2. VR SUPPORTING ALLOCENTRIC AND EGOCENTRIC FRAME OF REFERENCE

Previous studies have shown that users can get disorientated in VR and that it could be difficult to understand spatial and zoning relationships (Castronovo, 2019; Liu, 2020). However, we did not see this behaviour in our study. Instead, participants mainly relied on the miniature model as a map to gain a better understanding, orientation, and overview of the facility. The instant switching between view-modes also allowed participants to explore the same design issues from both self-centered egocentric perspectives as well as environment-centred allocentric perspectives within the same medium.

This switching between the two view-modes could then argued to be supported by cues of both frames of reference in each respective view-mode. An example of this is the green-coloured arrow in the miniature model-view that supported positioning and orientation in the facility. The multi-user representation, seen as avatars in both view-modes (figure 2), also facilitated positioning and orientation which aided coordination and collaboration among the participants.

Furnishment also acted as another cue. This can be seen for example in figure 3, where discussions about furnishment and how it affected the functionality of studies space came about as a result of the interplay between the two view-modes. Moreover, it was clear that the additional furnishment influenced the participants' frame of reference in the different view-modes. Furnishment seen in the allocentric based miniature model helped participants understand what type of room it was whereas furnishment seen in the egocentric oriented 1:1 scale helped participants understand how the room worked. In other words, furnishment could be seen as an important



allocentric cue in both view-modes.

Previous research highlights and advocates the need of multiple modes of representation to support individual work as well as to support cross discipline communication during the design process (Mastrolembo Ventura et al., 2019; Roupé at al., 2019). For instance, architects sometimes have a preference to start from a more allocentric point of view (i.e., 2D drawings) before switching over to 3D and other information mediums. At the same time, building end-users often state that they “do not know how to read 2D-drawings” (Mastrolembo Ventura et al., 2019). In a similar way, the use of immersive full-scale VR during our study came as a result of the client’s representatives expressing that it was difficult to get a full understanding of and review the project from 2D-drawings only. However, during the workshops the miniature model was used to a surprisingly high degree for both discussions and individual inspection. In this context, the miniature model can perhaps be seen as sort of a hybrid between a 2D-plan representation and a physical scale model, especially with the use of a section plane (Figure 2). It is then interesting to see that the representation that they didn’t understand before the workshop then became the centre of attention to such a high degree during the workshop (e.g., see Figure 1, right). Still, it is somewhat unclear if it was the use of immersive, full-scale VR that made the participants better understand and relate to a (hybrid) 2D-plan representation of the project, or if the miniature model itself is a better representation in terms of project overview compared to a traditional 2D-plan. What became clear, though, was that the two view-modes were used to discuss and address different design issues. For example, logistical flow that affected the facility at large, was discussed and addressed mostly in the miniature model whereas room functionality (e.g., placement of furnishment) was primarily discussed in the 1:1 scale view-mode.

## 5. Conclusion

The result of this paper shows that by supporting easy switching between both egocentric and allocentric frame of reference in the same VR medium, it is possible to achieve an enhanced spatial understanding for orientation and overview of the design. Furthermore, by allowing multiple users in the same model the participants could simultaneously explore same and different view-modes, which facilitates collaboration and leaves less space for misinterpretation of the reviewed design issues. In this context it was found that the miniature-model was used more than expected considering that the main purpose with the workshops was to review the design in scale 1:1 in immersive VR. It was also found that for both view-modes the addition of furnishment greatly enhanced understanding, collaboration, and discussions among users. Furnitures bring purpose and context to the facility and therefore it becomes easier for the participants to reason about different locations and the logistical flow and also how it affected the immediate area. This observation thus highlights the importance of having a sufficient LOD for the purpose of the review. Lastly, through the integration in education, the project contributes to SDG 4. Furthermore, it also supports SDG 9, SDG 11, SDG 12 as it contributes to a more effective and sustainable design and construction process.

For future research, it would be interesting to study how LOD can be mapped to different phases of the design process and how the different phases affect cross-

discipline collaboration and the level of design issues discussed and addressed collaboratively.

## References

- Abouelkhier, N., Shawky, D. & Marzouk, M. (2021). Evaluating distance perception for architecture design alternatives in immersive virtual environment: a comparative study. *Construction Innovation*. <https://doi.org/10.1108/CI-11-2020-0188>.
- Byers, T., W. Imms, & E. Hartnell-Young. (2018). Evaluating Teacher and Student Spatial Transition from a Traditional Classroom to an Innovative Learning Environment. *Studies in Educational Evaluation*. 58, 156–166. <https://doi.org/10.1016/j.stueduc.2018.07.004>.
- Dadi, G., Goodrum, P., Taylor, T., & Carswell, C. (2014). Cognitive Workload Demands Using 2D and 3D Spatial Engineering Information Formats. *Journal of Construction Engineering and Management*. 140, 04014001. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000827](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000827).
- Davila Delgado, M., Oyedele, L., Beach, T., & Demian, P. (2020). Augmented and Virtual Reality in Construction: Drivers and Limitations for Industry Adoption. *Journal of Construction Engineering and Management*. 146, 04020079. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001844](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001844).
- Frelin, A. & Grannäs, J., (2020). Teachers' pre-occupancy evaluation of affordances in a multi-zone flexible learning environment – introducing an analytical model. *Pedagogy, Culture & Society*. <https://doi.org/10.1080/14681366.2020.1797859>.
- Johansson, M., & Roupé, M. (2019). BIM and Virtual Reality (VR) at the construction site. *International Conference on Construction Applications of Virtual Reality*. CONVR 2019. Enabling digital technologies to sustain construction growth and efficiency, 19, 1-10.
- Liu, Y., Castronovo, F., Messner, J., & Leicht, R. (2020). Evaluating the Impact of Virtual Reality on Design Review Meetings. *Journal of Computing in Civil Engineering*, 34(1), 04019045. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000856](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000856).
- Mastrolembo Ventura, S., Castronovo, F., Nikolic, D., & Ciribini, A. (2019). A framework of procedural considerations for implementing virtual reality in design review. In *European Conference on Computing in Construction*. (pp. 442-451). European Conference on Computing in Construction, Greece, 2019. <https://doi.org/10.35490/EC3.2019.160>.
- Roupé, M., Johansson, M., Elke, M., Karlsson, S., Tan, L., Lindahl, G., & Hammarling, C., (2019). Exploring different design spaces – VR as a tool during building design. *19th International Conference on Construction Applications of Virtual Reality*. CONVR 2019.
- Roupé, M., Johansson, M., Maftei, L., Lundstedt, R., Viklund, & Tallgren, M. (2020). Virtual Collaborative Design Environment: Supporting Seamless Integration of Multi-touch Table and Immersive VR. *Journal of Construction Engineering and Management*. 146, 0001935. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862](https://doi.org/10.1061/(ASCE)CO.1943-7862).
- Sateei, S., Eriksson, J., Roupé, M., Johansson, M., & Lindahl, G. (2021). How Virtual Reality is used when involving healthcare staff in the design process. *Proceedings of the 38th International Conference of CIB W78*, Luxembourg, 13-15 October, (pp. 419-428).
- Teixeira, M., Pham, K., Caldwell, G., Seevinck, J., Swann, L., Rittenbruch, M., & Volz, K. (2021). A user-centred focus on augmented reality and virtual reality in AEC: Opportunities and barriers identified by industry professionals. In *26th International Conference on Computer-Aided Architectural Design Research in Asia: Intelligent and Informed, CAADRIA 2021*, Vol. 2, (pp. 273-283).
- Van den Berg, M., Hartmann, T., & de Graaf, R. (2017). Supporting design reviews with pre-meeting virtual reality environments. *Journal of Information Technology in Construction (ITcon)*, 22(16), 305-321.

# DYNAMIC PROJECTION

JAMES NANASCA<sup>1</sup> and AARON G. BEEBE<sup>2</sup>

<sup>1</sup> Pratt Institute, <sup>2</sup> Consortium for Research and Robotics

<sup>1</sup>[jnanasca@pratt.edu](mailto:jnanasca@pratt.edu), 0000-0003-0481-1719

<sup>2</sup>[abeebe39@pratt.edu](mailto:abeebe39@pratt.edu), 0000-0003-2361-1499

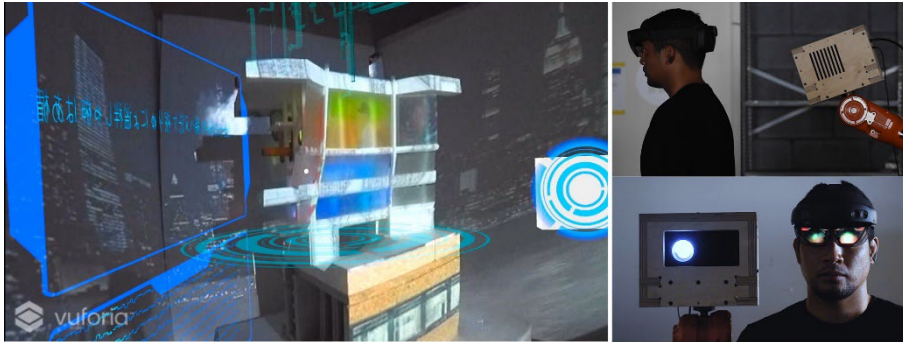
**Abstract.** Rarely are technologies of projection mapping (PM) and mixed reality (MR) used together with an architectural agenda. Dynamic Projection imagines the confluence of accessible PM and MR technologies and asks “How might we leverage the strengths of both technologies while obviating their weaknesses?” And then “How might this technology be of use in making architecture from within the Climate Movement?” First, we will examine the dormant potential of Projected MR by augmenting a physical model in an exhibition setting. The exhibition set-up deploys Unity and Vuforia to generate MR, and Mad Mapper to generate a projection mapped background space. Using this set-up reveals strengths in both technologies, which we can evaluate with a Cybernetically Enhanced Mixed Reality Framework. We can leverage this Projected MR as a suite of tools to make architecture a more active participant in the Climate Movement: for example, by augmenting buildings with statistics that could help reduce energy consumption or through the augmentation of the construction process, helping facilitate waste reduction through efficient construction. Our initial research is being expanded through development of a more versatile Projected MR platform with Dynamic Projection 02, in which we are utilizing better MR tools, more responsive PM tools, and an industrial robot to simulate various dynamic feedback systems. This expanded research design speculates on a 3-part exhibition that can respond with low latency via Projected MR controls during a public and private interactive experience.

**Keywords.** Projection Mapping; Augmented Reality; Projected Augmented Reality; Cybernetics; Mixed Reality; Responsible Consumption and Production; Climate Action; SDG 12; SDG 13.

## 1. Introduction

Mixed Reality (MR) (often used interchangeably with Augmented Reality) and Projection Mapping (PM) are related but markedly different technologies used to overlay digital information onto real-world objects. Using different strategies, both mix the real and the virtual to varying degrees. Projection Mapping casts light onto objects to augment them with virtual effect, while MR typically relies on a handheld, or head mounted prosthetic to align virtuality with reality. This paper investigates the latent

affect of the layering of Projection Mapping with Augmented Reality. Using recent expansions and reframing of two important conceptual frameworks, Cybernetics (Wiener 1948), and the Reality-Virtuality Continuum (Milgram et al. 1995) allows us to reassess previous Mixed Reality research in architecture, and better position future research moving forward.



*Figure 1: Augmented Reality and Projection Mapping Overlayed on Exhibition Object, and Dynamic Projection Elevations*

### 1.1. MIXED REALITY FRAMEWORK AND CYBERNETICS

In a formative study, Paul Milgram and a team in Toronto developed a continuum of Reality-Virtuality that has since become a keystone theory in the field of Mixed Reality and Human-Computer Interface (Milgram et al. 1995). This work was recently reappraised through a comprehensive analysis of 68 peer reviewed papers about Mixed Reality (MR), and via interviews with ten Augmented Reality/Virtual Reality authorities in Academia and industry, providing an updated understanding of how these experts define MR (Speicher et al. 2019). The recent analysis adds depth to the MR Continuum by proposing 6 definitions of MR (Table 1, Column 1) used by experts in the field, and then constructing a conceptual framework of MR features that allows us to unambiguously describe MR experiences (Table 1). This framework describes specific variations of MR and provides readily usable attributes for categorizing MR experiences.

The key parts of a communication theory of Cybernetics can be aligned with our query into PM and MR by examining them through the lens of Speicher et al.'s work. Wiener's original Cybernetic theory has been summarized as being

“...concerned with system models in which some sort of monitor sends information about what is happening within or around said system at a given time to a controller. This controller then initiates whatever changes are necessary to keep the system operating within normal parameters.” (Lasky 2020).

We here propose an additional dimension (grey columns in Table 1) to identify a relationship between the Cybernetic Dimensions of input and output, or the Monitor and the Controller. Using this framework, we can now further investigate how PM and

MR behave differently at the communication and control level – allowing for better research design.

Dimension	No. of Environments		Level of Immersion			Level of Virtuality			Interaction		Monitor/ Feedback	Response	Feedback Relationship		
	A	B	A	B	C	D	E	C	D	E	F	G	H	H	H
1. Continuum	✓		✓		✓	✓	✓	✓	✓		✓		✓	✓	✓
2. Synonym	✓		✓		✓	✓		✓			✓		✓	✓	✓
3. Collaboration	✓	✓		✓	✓	✓		✓	✓		✓	✓	✓	✓	✓
4. Combination	✓		✓		✓	✓		✓	✓		✓		✓	✓	✓
5. Alignment		✓	✓		✓	✓	✓	✓	✓	✓	✓		✓	✓	✓
6. Strong AR	✓		✓		✓			✓			✓	✓	✓	✓	✓

Table 1: Cybernetically Enhanced Conceptual MR Framework (Speicher et al. 2019, Wiener 1948)

*A. One; B. Many; C. Not; D. Partly; E. Fully; F. Implicit; G. Explicit; H. Any, Grey: Cybernetic Enhancement*

## 2. Objectives

By aligning definitions of MR from the field of Human-Computer Interfaces with our own architectural research, this paper seeks to counter the growing uniformity of research projects around the field of architectural MR. Further, by using this method to unambiguously classify various types of MR, we can help guide future architectural MR research, and begin to develop uses and modes of action that can have immediate real-world consequences.

Advances in BIM, Facebook’s Metaverse, and technology of the like; will lend themselves to hybrid information systems, Architectural design and construction will necessarily tap into these definitions of MR. Our research aims to provoke an awareness of the utility of hybrid information technologies for sustainable construction methods, efficient design, and climate-change related management of construction, among others.

Finally, by aiming to pull AR/MR functionality back from the interior, private world of prosthetically-enhanced interactions, we hope to privilege and address the actual, physical world of climate change, hunger, waste, and politics. We challenge

future researchers to use this framework to judge their own work and its impact on sustainable development.

### 3. Background

#### 3.1. INTERACTIVE PROJECTION MAPPING

For many years, interactive PM has been the subject of academic research framed for the entertainment industry. This has included a low-latency, room-filling multi-projector set up to play handheld controller video games (Ryu et al. 2006); a Dyadic Projected Spatial Augmented Reality that allowed a pair of users at fixed perspectives to interact with 3D spatial projections without the need of hand-held or head-worn prosthesis by using three video projector–Microsoft Kinect Rigs (Benko et al. 2014); and on a smaller scale, a hand gesture based holographic 3D modeling experiment that leveraged a single projector, semi-reflective screen and Leap Motion sensor (Johnson and Teng, 2014). When evaluated using the Cybernetically Enhanced Conceptual MR Framework (Table 1) these projects show use-cases of an interactive mixed reality that are: heavy with implicit interaction; single environment; single and multi-user; partially-fully immersive; partially-fully virtual, both implicitly and explicitly interactive, and have a cybernetic loop with feedback monitors and response. These dimensions are characteristics of interactive projects that could lead to richer experiences for users of MR devices and are the ones we are pursuing with our Dynamic Projection research.

#### 3.2. MIXED REALITY/AUGMENTED REALITY

Typically, AR/MR research is focused on enhancing fabrication processes that liberate the construction process from 2D drawings, allowing for the fabrication of complex 3D forms. Fologram – a software plug-in for Rhino 3D and Grasshopper applications that facilitates the building instructions and geometry streaming in mixed reality to precisely track the bending steel, and steam bending of wood into curvilinear shapes, is a common tool (Jahn et al. 2019). Others have leveraged custom-built apps using Vuforia, a plug-in for Unity, to visualize step-by-step instructions through a HoloLens and iPhone, aiding the construction of his space frame structure and panelization for his thesis project (Gopel 2019). Still others have applied AR using real-time motion tracking cameras like OptiTrack and Kinect to track progress of diverse construction methods (Hahm 2019).

These types of deployments of MR/AR technologies are exemplary of current research, revealing a uniformity of methodology and objectives (MR/AR used as a set of 3D instructions to assemble atypical forms). Examining these MR/AR projects using the Cybernetically Enhanced Conceptual MR Framework (Table 1), we can see that all these projects are: single environment; single-user (one person controlling); partially-fully immersive; partially-fully virtual, both implicitly and explicitly interactive, and have a cybernetic loop with feedback monitors and response.

These projects rely on interactions between the MR interface and physical reality – and the AR artifacts only exist in the MR/AR interface unless acted upon. Thus, interactions in the MR/AR digital world remain in the MR Prosthetic until an operator

acts on the objects in the physical world. Understanding this limitation offers a ripe opportunity for architectural MR/AR researchers to create systems that utilize the MR interactions of multiple users, facilitating a control-monitor-output loop via Projection Mapping, and building tighter ties between virtuality and reality.

## 4. Methods

### 4.1. MODEL DESIGN AND EXHIBITION SET UP

Our initial research centred on the design of an exhibition object for MR/AR experimentation. We began with a 1/16" = 1'-0" scale architectural chunk model of a waste-to-energy Facility in Brooklyn NY. This object let us study different applications for projection mapping, with some applications directly scalable to real buildings. The chunk model was fabricated with concrete, PLA, and frosted acrylic to emulate a photo-receptive media facade that allows for advertisements, art, videos, and waste-to-energy statistics to be projected onto the facade (Figure 4).

The exhibition was installed with our model and a plain backdrop. A Lightform LF2 projector end effector was attached to an ABB IRB1600 robotic arm housed in a plywood enclosure. Using a 6-axis arm to hold the projector allowed us to accurately engage different projection angles, memorized by our Software Control Computer. More than one operator could experience the additional layer of MR on the exhibition model using their own MR prosthetic - in this case a HoloLens or iPhone (Figure 3).

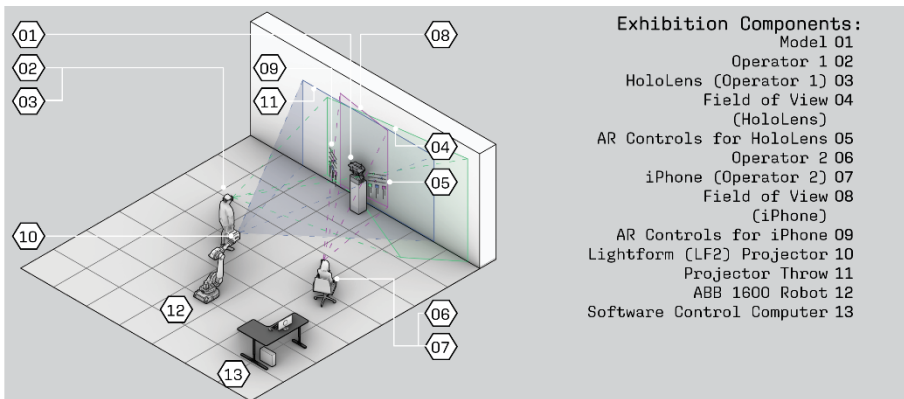


Figure 3: Exhibition Set Up for Dynamic Projection

## 4.2. PROJECTION MAPPING

Our initial setup provided a visual, but not interactive experience. We then added interaction between the software and the users. Looped videos and still textures were projected onto the model to set ambiance and animate the exhibition object. We used Mad Mapper 3.7.4 to manage the textures and animations. Figure 4 shows the plain object and the projection mapped object. Metal texture was mapped onto the PLA of the model (Figure 4.05), allowing us to see how different facade systems may present themselves. The media facade shows beverage advertisements (Figure 4.08), waste-to-energy infographics (Figure 4.07), tech advertisements (Figure 4.10), and a baseball game (Figure 4.09), among other things.

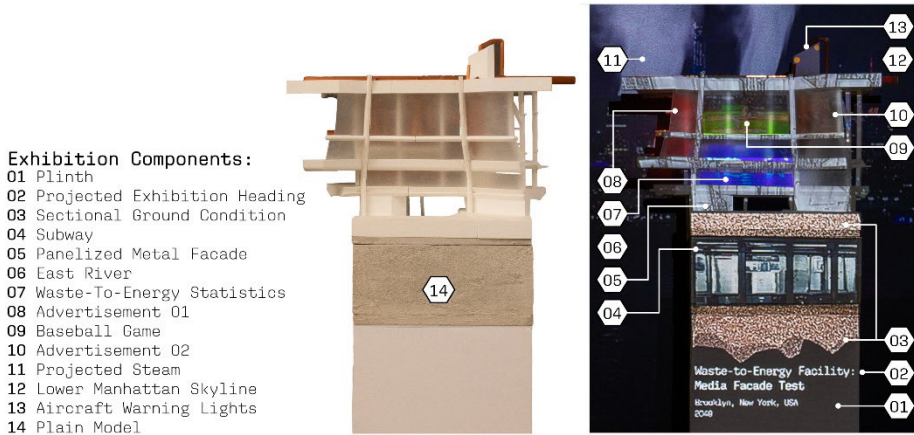


Figure 4: (Left) Plain Model, (Right) Projection Mapped

## 4.3. MIXED REALITY

The MR set up in Dynamic Projection was discrete from the PM being deployed. We used Vuforia 9.8.8 and Unity 2019.3.6.1 to make a custom application for the MR devices (iPhoneX and Hololens2), allowing a single user to pull additional visualization from a digital model in Unity. These custom applications use a 3D target marker generated by Vuforia to align the digital media with the real world, allowing interactions between the operator and digital model to take place. We populated the MR with a series of animated placeholder texts filled with information about the exhibition and a button (Figure 5) that toggles between exploded views of the Mechanical and Structural assembly holograms.



## 5. Discussion

### 5.1. DYNAMIC PROJECTION

The first iteration of Dynamic Projection uncovered latent potential by combining the mixed reality method of PM and MR Prosthetics. Using the Cybernetically Enhanced Mixed Reality Framework (Table 1) to analyse our exhibition set up revealed the following aspects of MR: multi-environment (discrete Projection Mapping and MR digital environments); single and multi-user (multiple users each using their own MR Prosthesis but sharing the same Projection); partially-fully immersive; partially-fully virtual, both implicitly and explicitly interactive, and have a weak cybernetic loop with feedback monitors and response.

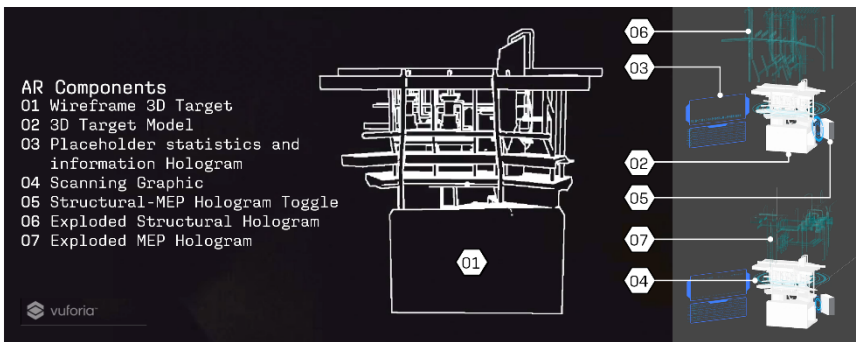


Figure 5: - MR Schematic

Dissecting this set up, we see a loose-to-no connection of PM to MR. MR interactions are only visible via prosthetic, and PM scenes, while visible without prosthetic, cannot change once Dynamic Projection is in action. We see potential for improving multi-user experience through multiple MR devices interacting with the same object in a shared environment tied together with PM. The combination of PM and MR in an interactive Projected Mixed Reality platform could be deployed at an architectural scale.

Scaled up to architecture, we posit that Dynamic Projection could make architecture a more active participant in addressing the systemic injustices of climate change. At the construction stage of architecture, we can see a Dynamic Projection set-up as an interactive, re-orientable version of previous projector assisted fabrication experiments (Ahn et al. 2019). This could minimize construction waste by allowing laborers to more efficiently and responsibly cut down and dispense material, for example. Existing architecture could be transformed into a site for interactive media. Building facades could be outfitted with Dynamic Projection set-ups to make public facing and interactable displays of energy usage and other related information to help regulate energy consumption and even educate people interacting with the building.

## 5.2. DYNAMIC PROJECTION 2

### 5.2.1. Exhibition Set Up and Hardware Changes

For our current and future research, we utilize a similar exhibition set up. Dynamic Projection 2 focuses on the development of the interactive Projection Mapping and MR connection, imbuing positional tracking of the MR Prosthetics to inform our mobile robotics platform (Figure 6).

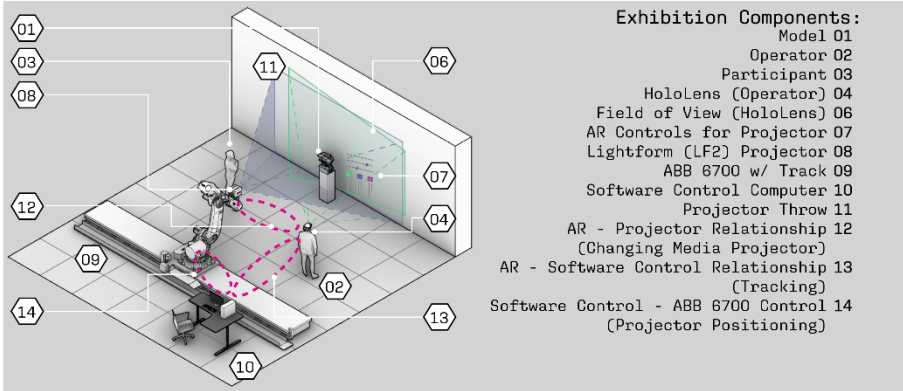


Figure 6: Proposed Layout for Dynamic Projection 02

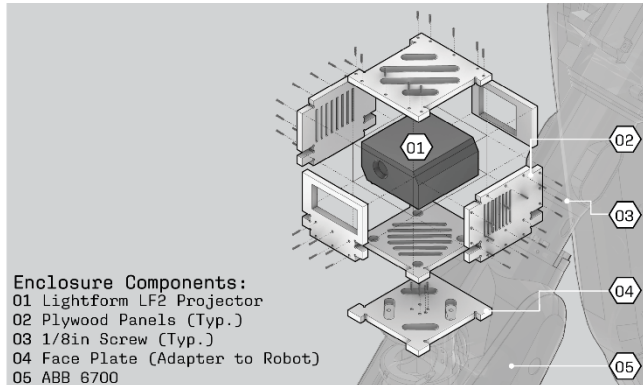
The most prominent design-hardware change is the use of a much larger ABB IRB6700 on a 30-foot track to emulate the wider range of motion inherent in a variety of potential future setups. This development of AR controlled projection mapping will engender discussions for the private (HoloLens or phone) and public PM realms of this proposed system and will likely prompt investigations into this research that are both privately and publicly framed.

The new end effector we have designed is both big enough to enclose the projector and vented to keep air circulating where required. Vibrations could be further minimized by using a lighter material and rubber gaskets to dampen the shaking. We imagine the current box design of the end effector will remain largely the same for the next iteration (Figure 7).

### 5.2.2. Software Changes

Figure 6 shows a rudimentary layout for our intended MR interface. Built on Fologram: sliders and buttons will be used to encourage explicit interaction. For example, an MR slider interaction could move our robot in predetermined positions for alternate projection mapped scenes. An MR button could be used to switch between materials being projected onto artifacts, or toggle different information displayed in the background.

To aid the interactive PM, TouchDesigner will be used in lieu of Mad Mapper for its capability to communicate with Rhino3D/ Grasshopper via a communication plug-in called gHowl. Grasshopper will serve as the main platform that facilitates the communication between TouchDesigner and the MR interactions on Fologram.



*Figure 7: End Effector Enclosure*

## 6. Conclusion

This paper presents a use case for the confluence of PM and MR technologies. Previous technologies were examined by leveraging Weiner’s Cybernetics theory of feedback loops, and Speicher et al.’s expansion on the reality-virtuality continuum. These frameworks allowed us to explicitly categorize aspects of Mixed reality in the projects studied and let us tailor specific aspects of what we’d like to imbue and tease out into our own Dynamic Projection experiments. This early investigation into the confluence of PM and MR will allow us to intelligently design more interactive Projected Mixed Reality platforms in the future. The first iteration of Dynamic Projection revealed some limitations in current MR technologies, pointed out the latent potentials for the confluence of MR, and allowed us to speculate on how Dynamic Projection can help architecture be a more active participant in the Climate Movement. Our current research is expanding on those potentials and beginning to create real-world applications for use in the fields of architectural design and construction.

## Acknowledgements

Thank you to the research assistants and staff at The Consortium for Research and Robotics during the Summer and Fall of 2021 in assisting with fabrication and documentation of Dynamic Projection (Rohan Saklecha, Rob Woods, Emily Gibson, and Phoebe Degroot). Thank you to Jeffery Anderson and Olivia Vien at Pratt Institute for encouraging me to produce an academic research paper after attending their class. Many thanks to Mark Parsons for supporting this research.

## References

- Goepel, G. (2019). Augmented Construction. ACADIA 19: UBIQUITY AND AUTONOMY, *Proceedings of the 39th Annual Conference of the Association for Computer Aided Design in Architecture (ACADIA)* (pp. 430-437).
- Gwyllim, J., Andrew, W., and James, P. (2019). [BENT]: Holographic handcraft in large-scale steam-bent timber structures. ACADIA 19: UBIQUITY AND AUTONOMY *Proceedings of the 39th Annual Conference of the Association for Computer Aided Design in Architecture (ACADIA)* (pp. 438-447).
- Hahm, S., (2019). Augmented Craftsmanship. ACADIA 19: UBIQUITY AND AUTONOMY, *Proceedings of the 39th Annual Conference of the Association for Computer Aided Design in Architecture (ACADIA)* (pp. 448-457).
- Hrvoje, B., Wilson, A. D., and Zannier, F., (2014). Dyadic Projected Spatial Augmented Reality. *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*.
- Jaeho, R., Naoki, H., Makoto, S., Masashi, S., and Ryuzo, O. (2006). A Game Engine Based Architectural Simulator on Multi-Projector Displays. CAADRIA 2006, *Proceedings of the 11th International Conference on Computer Aided Architectural Design Research in Asia* (pp. 613-615).
- Jones, B., Rajinder S., Murdock, M., Mehra, R., Benko, H., Wilson, A., Ofek, E., Macintyre, B., Raghuvanshi, N., and Shapira, L, (2014). RoomAlive. *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*.
- Lasky, J., (2020) *Cybernetics*. Salem Press Encyclopedia of Science. Retrieved from <https://search-ebscohostcom.ezproxy.pratt.edu/login.aspx?direct=true&db=ers&AN=87321730&site=eds-live&scope=site>.
- Milgram, P., Takemura, H., Utsumi, A., and Kishino, F., (1995). Augmented Reality: A class of displays on the reality-virtuality continuum. *Telemanipulator and Telepresence Technologies*.
- Sangjun, A., Han, S., and Al-Hussein, M. (2019). 2D Drawing Visualization Framework for Applying Projection-Based Augmented Reality in a Panelized Construction Manufacturing Facility: Proof of Concept. *Journal of Computing in Civil Engineering*, 33 (5).
- Speicher, M., Brian D. Hall, and Nebeling, M. (2019). What Is Mixed Reality. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3290605.3300767>.
- Teng, T., and Johnson, B. (2014) "Inspire: Integrated spatial gesture-based direct 3D modeling and display." ACADIA 14: Design Agency, *Proceedings of the 34th Annual Conference of the Association for Computer Aided Design in Architecture (ACADIA)* (pp. 445-452).
- Wiener, N. (1948). *Cybernetics: Control and Communication in the Animal and the Machine*. Wiley.

# AUTOMATIC FLYTHROUGH AUTHORIZING IN VR

ROBIN SCHMIDIGER<sup>1</sup>, ROMANA RUST<sup>2</sup> and ROI PORANNE<sup>3</sup>

<sup>1,2,3</sup>*ETH Zurich*

<sup>1</sup>*[schmrobi@student.ethz.ch](mailto:schmrobi@student.ethz.ch), 0000-0002-3110-1061*

<sup>2</sup>*[rust@arch.ethz.ch](mailto:rust@arch.ethz.ch), 0000-0003-3722-8132*

<sup>3</sup>*[roi.poranne@gmail.com](mailto:roi.poranne@gmail.com), 0000-0002-1749-2596*

**Abstract.** The presentation of architectural models in virtual reality (VR) enables architects to evaluate their designs and present them to stakeholders, improving fidelity of architectural communication in architectural design. Virtual walkthroughs are a powerful tool to convey architectural concepts and unique features of the planned building. However, creating walkthroughs is a rather tedious process involving manual setting of keypoints at points of interest and in between, which are then interpolated to create the trajectory. This often results in awkward camera poses, sharp turns, intersection with geometry and clipping problems, all of which must be resolved by tweaking or adding new keypoints. Furthermore, the process requires constantly switching between an inside view, for local refinements, and an outside overview, for more global adjustments. To address this challenge, we present an interactive design tool for walkthroughs in VR. Our tool automatically provides a set of ideal keypoints, based on a customizable set of objectives, and instantly generates a complete and smooth trajectory that passes through them. It allows the user to tweak positions and update the walkthrough in real time, in VR. In addition, it allows the designer to have both an inside view and an overview simultaneously.

**Keywords.** Automatic Walkthroughs; Virtual Reality; Optimal Trajectory; User Interface; SDG 11.

## 1. Introduction

Walkthroughs, flythroughs, and virtual tours--also known as trajectories--are some of the most effective techniques for showcasing and highlighting a building and its unique features. Advances in real-time 3D rendering allow walkthroughs to reach new levels of realism, and have become necessary in many practices. As VR devices become more ubiquitous, so will immersive walkthroughs. There is a twist however: users are no longer limited to walking and looking in the direction of the path prescribed by the designer but are free to look around and stray from the path. Walkthrough designers should take this into consideration and carefully examine their walkthroughs in VR, using a system that facilitates tweaking and fine-tuning.

Designing a walkthrough can be a rather tedious process. It involves manually specifying a path via key points, which are then interpolated in various ways. Standard interpolation methods, e.g. splines, often result in paths that intersect with the geometry, leading to distracting clipping problems. These must be resolved by tweaking the existing key points, or by adding new ones. With this iterative process, the designer inevitably ends up with many redundant keypoints that are difficult to modify. Automatic trajectory generation algorithms have been proposed and are discussed in Section 2, but were not intended for an interactive VR interface.

Our goal is to develop a computer-assisted, immersive, and interactive tool for walkthrough design. The tool provides an initial walkthrough, optimized for a variety of metrics, and allows the user to tweak and simultaneously optimize it in an intuitive fashion in VR. The challenge lies in the optimization itself, which involves finding an optimal set of optimal keypoints based on metrics such as spatial coverage and visibility. Next, we address the question of how to specify the best visiting order, and finally, how to create a smooth and intersection-free trajectory. To enable interactivity, these computations must be performed in real time. Our contributions in this paper are threefold:

- An optimal keypoint placement method.
- A collision-free optimal trajectory that optimizes keypoint ordering.
- A VR design interface.

The early design phase has the highest impact on a building project and hence greatly influences material consumption and environmental concerns. Therefore, it is paramount that we include both implicit (aesthetic, cultural or emotional) as well as explicit (e.g. functional, environmental, economic) criteria early on to mitigate design problems. With our project we respond to goal 11 of SGS - (Make cities and human settlements inclusive, safe, resilient and sustainable), improving the communication of design ideas to stakeholders, that are involved at later stages of the design-to-construction processes, and therefore foster better informed design decisions.

## 2. Related work

Many applications provide tools to manually generate walkthroughs. A common approach is to prescribe a sequence of keypoints, which are then interpolated to create a path. Keypoint specification can be done by placement in the plan or the model of the building, and is the traditional approach when working with CAD/BIM/3d modelers. Alternatively, game engines allow the user to walk the scene in a first- or third-person perspective, and to assign keypoints at their avatar's position. Previous work on automatic walkthrough generation in the context of architecture is rather limited. To the best of our knowledge, the only academic work that directly addresses this problem is by Graf and Yan (2008). They proposed to create a graph where each node represents a room and each edge represents a doorway. This graph is then used to create a traversal of the building.

In this work, we rely on path-planning and trajectory-optimization techniques commonly employed in robotics. Path-planning refers to task of finding any path between two points of an environment, while avoiding obstacles. See Yang et al.

(2016) for a recent review. The path does not have to be optimal in any sense, and hence, path-planning is often followed by a trajectory-optimization phase (Kelly 2017). Lin et al. (2013) proposed a path-planning approach for 3D spaces using Fast Marching Methods. They provide insights on how, in the context of walkthrough planning, common 2D path-planning can be used instead of 3D path-planning, effectively reducing the complexity of the problem.

Creating a walkthrough of an architectural model is in particular similar to generating paths for drone cinematography (Galvane et al. 2018). Although virtual cameras are not constrained by the laws of physics, drone and virtual camera trajectory optimization share many goals. These include, but are not limited to, collision avoidance, cost optimization, smoothness of the trajectory, and coverage planning. Gebhardt et al. (2016) for example, developed a computational design tool that allows its users to easily create quad-rotor trajectories. They used an optimization-based method to ensure that the trajectories are feasible in the real world. Another method proposed by Nægeli et al. (2017) also tackles automated aerial videography. This time, online optimization is utilized, and the system is extended to work with multiple drones. Interactively-defined aesthetic framing objectives are also incorporated.

### 3. Automatic walkthrough generation

The computational aspect of our approach consists of two steps. First, a set of keypoints is determined. These keypoints are then sequenced by (approximately) solving the Travelling Salesmen Problem, and a path is created. In the second step, the path is further optimized to create a smooth and collision-free trajectory, while optionally considering other objectives as well. This optimization process can continuously run in the background while the user edits the keypoints, constantly improving the trajectory.

#### 3.1. DETERMINING KEYPOINTS

To determine keypoint locations, we first construct an importance map, where higher values signify greater importance. The map is based on metrics such as distance to walls and visibility (Section 3.1.1). The keypoints are iteratively extracted based on their importance (Section 3.1.2).

##### 3.1.1. Computing the importance map

The importance map is sampled on a voxel grid and consists of three components: obstacle, distance, and visibility (Fig. 1). The obstacle map is a binary map where a voxel contains '1' if it intersects with the structure's geometry. Voxels of the distance map contain the distance to the closest obstacle (floors and ceilings excluded) and is computed as a distance transform of the obstacle map. The goal of the visibility map is

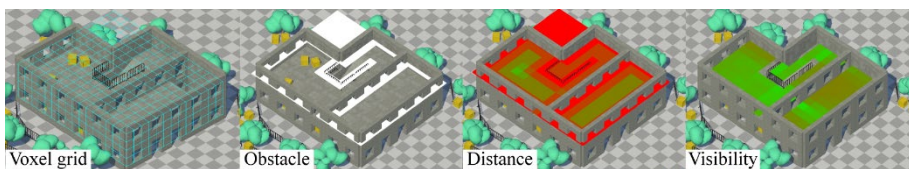


Figure 1. The different maps used for locating the keypoints

to measure how much of the structure is seen from each voxel. To compute the visibility at a certain voxel, we cast rays from that voxel to all other voxels and count the number of successful rays. The result is divided by the total number of cells and stored in the visibility map. Once the maps are computed, they are combined together as a weighted sum

$$\text{Importance}(P) = \lambda_{\text{obs}}(1 - M_{\text{obs}}) + \lambda_{\text{dist}}M_{\text{dist}} + \lambda_{\text{vis}}M_{\text{vis}},$$

Where  $M_{\text{obs}}$ ,  $M_{\text{dist}}$ ,  $M_{\text{vis}}$  are the obstacle, distance and visibility maps respectively, and the  $\lambda$ 's are user defined weights. Here,  $\lambda_{\text{obs}}$  is chosen as an arbitrarily high number, so voxels that collide with obstacles have low importance value. As can be inferred, voxels that are distant from obstacles have higher importance values, as well as voxels that are visible from many other voxels.

### 3.1.2. Extracting keypoints based on importance

With the importance function computed, our next goal is to place keypoints. The desiderata are to place them at locations with high importance values, but evenly distributed. A naïve approach would be to place a point in each voxel that exceeds a certain importance threshold. However, this has the risk of clumping many points closely together in regions of high importance, or neglecting peaks of high importance that are below the threshold. Another alternative is to pick local maxima of the importance map, but we found them to be too few, in general. Our approach is a mix of the two approaches. We iteratively pick the center of the voxel with the highest importance value as the next keypoint. Then, we reduce the importance of this voxel and all the voxels around it by a value proportional to the distance to the voxel. This creates a "cone" of low values around the selected voxel, preventing any close voxels from being selected in subsequent iterations. The radius and proportion of this cone can be adjusted by the user. Smaller radii will have higher chances of generating denser clusters of keypoints, while larger radii might produce too few keypoints. The process continues until the desired number of keypoints has been reached.

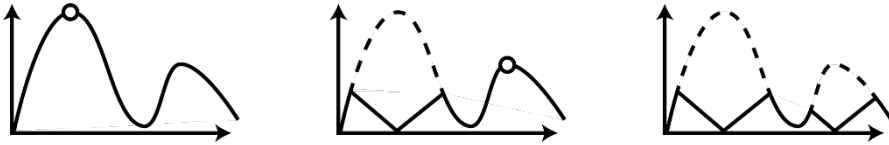


Figure 2. An illustration of the keypoint extraction process. From the left, the point with the maximal importance value is selected. It is then reduced to zero, with its vicinity reducing proportionally, creating a cone. The next maximum is the selected as the next keypoint, and the process repeats.

### 3.2. COMPUTING A TRAJECTORY

Once the points of interest have been determined, the next task is to compute an optimal trajectory that passes through them. We divide the task into two stages too. First, we determine the optimal sequence in which these points should be visited. We then create an initial rough trajectory and follow-up by a full trajectory-optimization routine.



### 3.2.1. Sequencing the keypoints

The optimal sequence is, arguably, the one that produces the shortest path that visits all of the keypoints. The problem of finding such a path is known as the Traveling Salesmen Problem. We begin by computing all shortest paths between every pair of keypoints. To do so, we leverage Unity's Navigation System, which works by automatically creating Navigation Meshes, and use them for path-finding. Ideally, we would then be able to solve the Traveling Salesmen Problem, but this problem is known to be intractable for more than a few points, since it scales factorially. Instead, we use a common approximation and compute the minimal spanning tree first, then proceeding with a depth first search. This creates a path that is guaranteed to be shorter than twice the length of the Traveling Salesmen path, and can be executed in real time. In any case, we also allow the user to modify the sequence, in order to address subjective considerations.

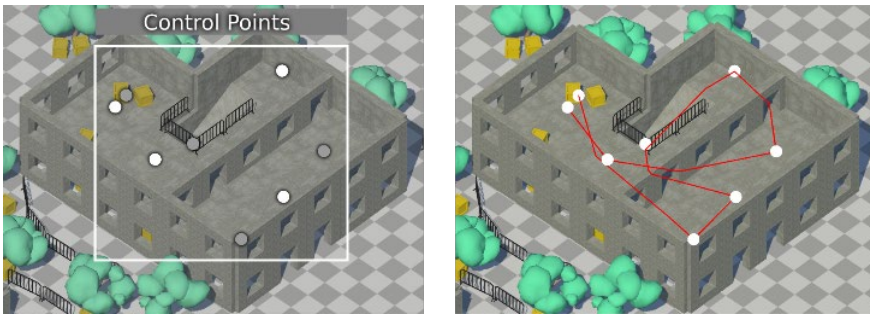


Figure 3. An image showing the keypoints on the left, and the optimal sequence shown on the right. The greyed out keypoints are points on the lower level that would normally be occluded.

### 3.2.2. Optimizing the trajectory

As can be seen in Figure 3, the path created is a polyline that has jagged edges and sharp corners, which might be visually disturbing. A common way to overcome that is to replace the straight edges by a smooth spline, e.g. a cubic b-spline. However, as mentioned, splines will sometime overshoot and collide with the structure geometry. As a remedy, users can add more keypoints, but this quickly becomes unmanageable. Instead, we opt to use a trajectory optimization approach. We discretize the polyline using a user defined number of points  $n$  (typically  $n = 100$  points) and represent them by a vector  $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ , where  $\mathbf{x}_i = (x_i, y_i, z_i)$ . Trajectory optimization is then formulated as a minimization problem

$$\min_{\mathbf{x}_1, \dots, \mathbf{x}_n} \mathcal{O}(\mathbf{x}_1, \dots, \mathbf{x}_n),$$

Where  $\mathcal{O}(\mathbf{x}_1, \dots, \mathbf{x}_n)$  is an objective that measures the quality of the trajectory. We solve this minimization problem using the quasi-Newton method L-BFGS (Nocedal and Wright 2006). The objective itself is a weighted sum of different subobjectives we specify in the following.

In our formulation, we used five different subobjectives. The first subobjective penalizes deviation from the initial polyline trajectory. Letting  $(\mathbf{x}_1^0, \dots, \mathbf{x}_n^0)$  represent

that trajectory, this subobjective is expressed by

$$\mathcal{O}_d = \sum \|\mathbf{x}_i - \mathbf{x}_i^0\|^2$$

The next two subobjectives aim to minimize the velocity and acceleration of the trajectory. Minimizing the velocity serves to reduce the length of the trajectory, at the cost of slightly straying away from the keypoints. Minimizing the acceleration encourages a smooth trajectory, as well as constant velocity. They are expressed using first-order finite differencing, i.e.,

$$\mathcal{O}_v = \sum \|\mathbf{x}_{i+1} - \mathbf{x}_i\|^2, \quad \mathcal{O}_a = \sum \|\mathbf{x}_{i+1} - 2\mathbf{x}_i + \mathbf{x}_{i-1}\|^2.$$

These three subobjectives result in a smooth trajectory, but one that might pass through walls and obstacles. We encode obstacle avoidance as an additional objective which penalizes proximity to obstacles. We let the user define the desired distance from obstacles  $D$ . Then we find, for each  $\mathbf{x}_i$ , the closest point on an obstacle  $\mathbf{x}_i^c$ . If the distance between  $\mathbf{x}_i$  and  $\mathbf{x}_i^c$  is greater than  $D$ , the objective contributes 0. Otherwise we define

$$\mathcal{O}_c = \sum (\|\mathbf{x}_i - \mathbf{x}_i^c\| - D)^2,$$

to encourage  $\mathbf{x}_i$  to move to a distance  $D$  away from  $\mathbf{x}_i^c$ . Note that  $\mathbf{x}_i^c$  is updated in each iteration and scenes crowded with obstacles can slow down the procedure. We note that we have also experimented using the distance map as a means to formulate this objective, but we found the results to be inferior, since the resolution of this map is limited. Finally, we included a flying height subobjective that ensures that the elevation of the trajectory over the floor remains as constant as possible. To this end, for each  $\mathbf{x}_i$  we compute the distance to the floor  $f(\mathbf{x}_i)$  and let

$$\mathcal{O}_h = \sum (f(\mathbf{x}_i)^2 - H)^2$$

be the height objective, where  $H$  is a user-specified desired height. This objective is useful in treating issues that occur around slopes, as can be seen in Figure 4. These are related to inconsistent height and lack of smoothness occurring between the keypoints at the bottom and top of the slope.

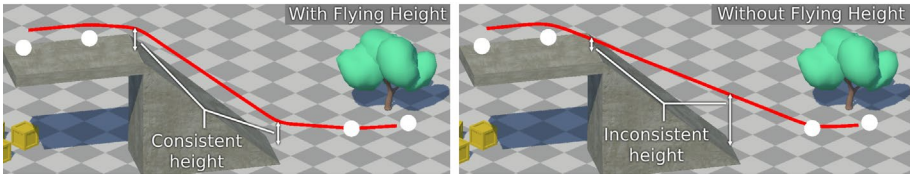


Figure 5. Results with and without the flying height objective. As can be seen, without this objective the height of the trajectory above the slope is only determined by the top and bottom keypoints, and thus is inconsistent along the slope. The issue is not present with this objective included.

The subobjectives are combined to form the total objective

$$O = \lambda_d O_d + \lambda_v O_v + \lambda_a O_a + \lambda_c O_c + \lambda_h O_h$$

Where  $\lambda$  are user-defined weights. The result of optimizing the path (shown in Figure 3) is shown in Figure 5. The significance of each subobjective is shown in Figure 6.

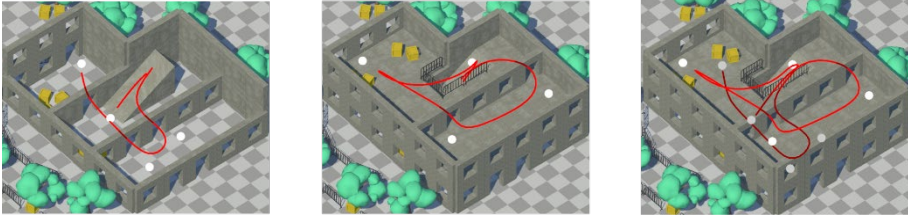


Figure 5. an optimal walkthrough trajectory of a two story building. Left and middle images show the trajectory in the first and second floor, and the right image shows the entire trajectory

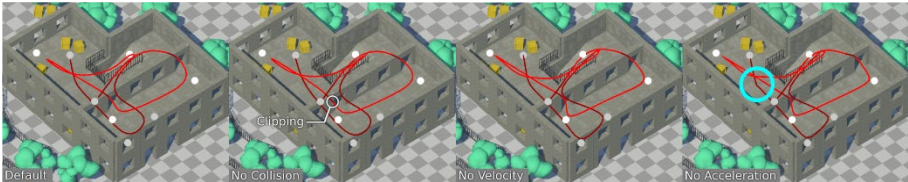


Figure 6. Demonstrating the effects of dropping any of the subobjectives

#### 4. User Interface

While the trajectory generated with our method is optimal in a sense, users may wish to customize and manually modify keypoint locations, as well as certain characteristics of the optimal trajectory. Our algorithm enables this with little-to-no waiting time, which allows for real-time interaction. Our interface allows users to explore the interior of the building in VR in a familiar way, but also allows them to carry a miniature version of the building and observe the exterior. This creates a unique experience, whereby users can be inside and outside of the building simultaneously. A few images demonstrating this interface with the DFAB house (Graser et al. 2020) are shown in Figure 7.



Figure 7. Images showing the miniature inside the building. The user can see the insides by manipulating a clipping plane (shown in the left and middle images). The right image shows the keypoints and trajectory as well.

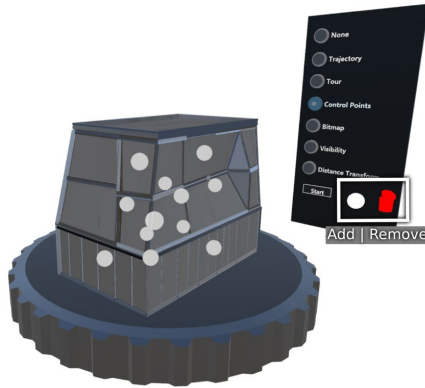


Figure 8. The miniature with its side panel. The user can easily move keypoints around by grabbing them, or add and remove point using the menu. Different visualization modes are shown in the menu as well

The miniature and its side panel are shown (without texture for improved clarity in the paper) in Figure 8. A specific keypoint can be grabbed and relocated using the controller. New keypoints can easily be added by grabbing them from the Add-section on the side panel and dragging them into the miniature. In order to remove a control point, it has to be dragged and released into the red recycling bin. The menu can be used to switch between visualization modes, as can be seen in Figure 9, and to start and stop playing the walkthrough using a button.

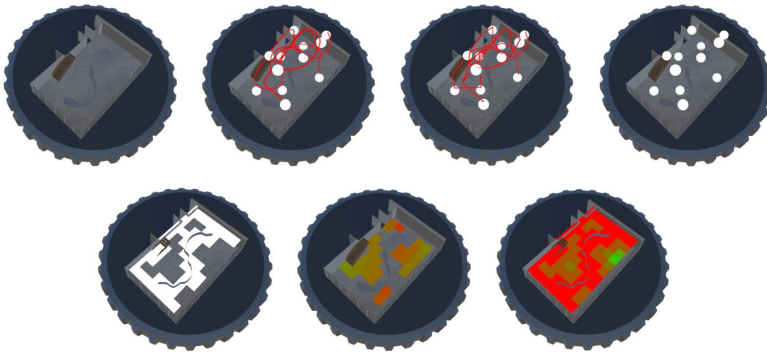


Figure 9. Different visualization modes seen on the miniature

To evaluate the user interface, we ran a small user study with 5 non-expert participants. Users were asked to design a walkthrough by placing keypoints sequentially, using our VR interface, but without the optimization. They were then asked to fill in a System Usability Scale (SUS) questionnaire (Sauro & Lewis 2012). Then, they were asked to design a walkthrough using our optimization and fill in the questionnaire again. In addition, participants were asked to comment on the differences between the two approaches. The SUS score is on a 1-100 scale, where a score above 68 is considered to be above average. The scores for the VR interface alone were 80, 77.5, 100, 87.5, 85 with an average of 86, while the scores for the optimization were

75, 75, 85, 82.5, 82.5 with an average of 80. Both were considered to be within the B range of usability. Participants noted that the initial automatic walkthrough gives them a good starting point and would save a lot of time. Participants also noted that without optimization the process was felt much longer, but at the same time, they felt more in control. Some participants identified boundary cases, which we must address in future work, which likely explain the drop in the score.

## 5. Conclusion

We presented an approach for automatic walkthrough generation based on detection of important keypoints and trajectory optimization. In addition, we implemented a VR user interface for planning walkthroughs, which leverages our optimization and allows real-time editing. The interface shows favourable results among participants, but further work is required in order to bring it beyond the experimental stage and tighten the experience in the UX/UI sense. While SUS gives initial indication regarding the usability, the fact that it is one-dimensional poses a limitation. Indeed, based on participant comments, it is clear that control and workload, as measured by, e.g. the NASA-TLX, play a critical role in system evaluation. A future user study will examine this with different populations and also include assessment of the resulting walkthroughs.

## Acknowledgements

Special thanks to Konrad Graser for providing the DFAB HOUSE 3D model and to Foteini Salveridou for modelling the Unity scene. This work was supported in part by the Swiss National Science Foundation through the National Centre of Competence in Digital Fabrication (NCCR dfab).

## References

- Delgado, J. M. D., Oyedele, L., Demian, P., & Beach, T. (2020). A research agenda for augmented and virtual reality in architecture, engineering and construction. *Advanced Engineering Informatics*, 45, 101122.
- Galvane, Q., Lino, C., Christie, M., Fleureau, J., Servant, F., Tariolle, F.-L., and Guillotel, P. (2018). Directing Cinematographic Drones. *ACM Transactions on Graphics* 37, 3, 18. <https://doi.org/10.1145/3181975>
- Gebhardt, C., Hepp, B., Naegeli, T., Stevšić, S., & Hilliges, O. (2016). Airways: Optimization-Based Planning of Quadrotor Trajectories according to High-Level User Goals. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*.
- Graf, R., & Yan, W. (2008). Automatic Walkthrough Utilizing Building Information Modeling to Facilitate Architectural Visualization. *eCAADe* 26 (pp.555). Education and research in Computer Aided Architectural Design in Europe (eCAADe).
- Graser, K., Baur, M., Apolinarska, A., Dörfler, K., Hack, N., Jipa, A., Lloret, E., Sandy, T., Pont, D., Hall, D., & Kohler, M. (2020). Dfab house - a comprehensive demonstrator of digital fabrication in architecture. *Fabricate 2020: Making Resilient Architecture*, 04 2020, pp. 130–139.
- Kelly, M. (2017). An introduction to trajectory optimization: How to do your own direct collocation. *SIAM Review*, 59(4), 849-904.

- Lin, Y.-H., Liu, Y.-S., Gao, G., Han, X.-G, Lai, C.-Y, Gu, M. (2013) The IFC-based path planning for 3D indoor spaces. *Advanced Engineering Informatics*, 27(2), 2013, 189-205, <https://doi.org/10.1016/j.aei.2012.10.001>.
- Nägeli, T., Meier, L., Domahidi, A., Alonso-Mora, J.& Hilliges, O. (2017). Real-time planning for automated multi-view drone cinematography. *ACM Transactions on Graphics*, 36, 4, Article 132 (July 2017), 10 pages. <https://doi.org/10.1145/3072959.3073712>
- Nocedal J & Wright S. J. (2006). *Numerical Optimization*. Springer.
- Sauro, J., & Lewis, J. R (2012). *Quantifying the user experience: Practical statistics for user research*. Morgan Kaufmann, Waltham MA, USA.
- Yang, L. & Qi, J. & Song, D. & Xiao, J. & Han, J. & Xia, Y. (2016). Survey of Robot 3D Path Planning Algorithms. *Journal of Control Science and Engineering*, 2016(1),1-22. <https://doi.org/10.1155/2016/7426913>.

# UNDERSTANDING DESIGN EXPERIENCE IN VIRTUAL REALITY FOR INTERIOR DESIGN PROCESS

HWAN KIM<sup>1</sup> and KYUNG HOON HYUN<sup>2</sup>

<sup>1,2</sup>*Department of Interior Architecture Design, Hanyang University*

<sup>1</sup>*iamhwankim@hanyang.ac.kr; 0000-0001-9816-035X*

<sup>2</sup>*hoonhello@hanyang.ac.kr; 0000-0001-6379-9700*

**Abstract.** Virtual reality (VR) can enhance users' spatial perception by enabling spatial design activities. Conversely, the VR environment provides more visual information for the user to process than the desktop environment, resulting in a low efficiency of the design process. This study aims to verify whether VR can have a distinctive influence on the spatial design experience compared to the desktop environment. We conducted user studies on design experience in VR and desktop environments to accomplish this goal. The results revealed that participants' satisfaction with the design experience was higher in VR; however, the task completion was more time consuming than in the desktop environment.

**Keywords.** Spatial Design Experience; User Study; Design Process; Virtual Reality; SDG 9.

## 1. Introduction

Owing to COVID-19, space utilisation behaviour such as telecommuting has changed and has accelerated the widespread use of virtual reality (VR) technology in our daily lives. In addition to traditional video call services (e.g. Zoom, Skype), the meta-verse concept was applied to socialisation, work, or recreation in a virtual world. Specifically, meta-verse services (e.g., Roblox, Cryptovoxel, Zepeto, and Gather Town) induce general users to design spaces or objects directly in virtual reality.

Innovating the design process using virtual reality technology is not a new concept. In the early design stages, virtual reality studies were conducted that allowed detailed review and modification of architectural spaces or intuitively showed the height of surrounding buildings (Yabuki et al., 2011). In addition to research using VR's immersive sense of space, new design methods have been proposed such as supporting the architectural design process by visualising airflow (invisible) in VR (Hosokawa et al., 2016), remotely experiencing a place previously inaccessible to designers (Pletinckx et al., 2000).

VR has been widely used to improve communication between designers and clients or decision-makers (e.g. Cave Automatic Virtual Experience system for vehicle companies). In addition, VR has focused on reviewing the completed design alternatives (de Casenave & Lugo, 2017; Freeman et al., 2018; Hou et al., 2009) or

proposing a method to reveal the side effects according to the design (Yabuki et al., 2011). Design reviews in VR are sometimes more efficient and can identify more issues (Freeman et al., 2018); however, other researchers report that design reviews in VR are inefficient and detect fewer issues (Hou et al., 2009). The existing literature has shown conflicting results. Therefore, the utility of VR may vary depending on experimental settings and conditions. In literature detailing the contribution of VR in design review activities, VR was found to engage the user's spatial perception more effectively than in the desktop environment when reviewing high-complexity objects (e.g. buildings) rather than low-complexity objects (e.g. doors) classified in the theory of technical systems (de Casenave & Lugo, 2017; Freeman et al., 2018). These studies claim that VR enables an immersive and intuitive experience of the dimension, size, and shape of the target object. However, to the best of our knowledge, no studies have confirmed the specific difference between the design process considering an immersive experience in a VR environment and the design process in a general desktop environment.

Spatial design activities can be assisted by VR, which can enhance users' spatial perception. Conversely, the VR enabled immersive environment has more visual information for the user to process than the desktop environment, which may slow down the design process. Therefore, this study aims to verify whether VR can provide a distinctive influence and contribution to the spatial design experience compared to the desktop environment.

The overall design direction is determined during the early design stages and continues to be revised and developed, which makes it difficult to modify the design in the later design stages. Hence, it is essential to understand whether the VR environment can make a meaningful contribution not only in later design reviews, but also in the early stages of design. Considering these factors, we investigate the difference between designing a space in VR and in a desktop in an uncertain situation where design conditions change frequently.

By comparing the design experience of the VR environment to the desktop environment, the design process in a metaverse environment can be performed more effectively in the future. To accomplish this goal, we performed the following tasks:

- 1) We conducted user studies on design experience in both VR and desktop environments.
- 2) We conducted quantitative and qualitative analysis on the experiment results.
- 3) We discussed the advantages and disadvantages of both environments.

## **2. Related Works**

### **2.1. SPACE PERCEPTION IN VR VS. DESKTOP**

In the desktop environment, users see 3D space through the monitor screen and navigate the space by freely moving the viewpoint (zoom, pan, orbit) using the mouse and keyboard interfaces. As it is difficult for users to perceive space through their somatic sense in the desktop environment, a separate reference object (e.g. human model) or measuring tool (e.g. Tape Measure Tool in SketchUp) is needed to determine the exact distance and dimensions of the virtual object. Most 3D CAD tools provide



these functions because accurate size is essential in the design process. The desktop's omniscient point of view helps users quickly and easily explore the shapes of objects in space and their relationships. However, when experiencing virtual reality (VR) by wearing a head-mounted display (HMD), the user sees the three-dimensional virtual space surrounding him through an egocentric view. The user perceives the virtual space through the whole somatic sense while changing the viewpoint by moving his body as in the real world. Unlike viewing 3D space on a monitor screen (i.e., the desktop environment), VR enables an immersive experience of 3D space.

However, when reproducing VR through an HMD, human space perception tends to vary from that in the real world. It has been experimentally confirmed that humans perceive the distance between themselves and objects (i.e. viewer-centred depth) in VR to be closer than the actual length (Willemssen et al., 2004). This phenomenon has not yet been clarified precisely, but the leading cause is the mismatch between accommodation - the focus of each eye on the HMD's virtual image - and vergence - the location created by overlapping the virtual images perceived by both eyes - is suspected (Drascic & Milgram, 1996). According to the size-distance invariance hypothesis, the perceived underestimated depth causes the user to perceive an object's size to be usually smaller (Kelly et al., 2013). The user can recognise the space more accurately by directly moving their body in a virtual space (e.g., walking or reaching out to an object) (Kelly et al., 2013). In other words, when designing a virtual space using VR, the user needs to experience the space by directly moving their body in a 1:1 scale space for an accurate sense of space.

## 2.2. DESIGN IN VR ENVIRONMENTS

In architecture and automobile design, resources and time are consumed to produce a physical mock-up on a 1:1 scale. VR technology offers the advantage of evaluating the design work and revisions efficiently and quickly. VR is often used for simulating designs that are difficult to experiment with within spatial design. For example, Yeom et al. (2021) confirmed how green wall design affects the user's cognitive state through a VR experiment. Hong et al. (2019) investigated how the window-wall ratio of a building affects occupant satisfaction. Both aforementioned papers used VR technology to confirm the responses the building user received as it is impossible to experiment with actual buildings. The 1:1 scale model in the VR immersive space provides a much better alternative to reviewing the final design than in a desktop environment.

Researchers have been studying whether the design review process using VR provides better usability than the desktop environment for two decades. Studies have explored whether VR users can complete design reviews faster or discover more design issues than in the desktop environment. However, existing studies have reported contradictory results. According to the experimental results of Freeman et al. (2018), the VR environment allowed participants (i.e., reviewers) to complete a given task in significantly less time. In addition, users found more design issues in the VR environment than in the desktop environment. However, the experiment of de Casenave & Lugo (2017) reported that the 3D model review was completed faster in the desktop environment than in the VR environment, and there was no significant difference in the design errors found. Many studies still differ on whether the design

review process in VR shows a better time and a greater number of identified design issues. A clear advantage of the design review process in VR is that it improves human spatial perception (Horvat et al., 2019). Compared to the desktop environment, VR can naturally and realistically interact with the 3D model and allow full use of human spatial understanding (Hou et al., 2009).

In addition to design reviews conducted for the late design phase, new attempts to utilise VR technology for design concept development have also been proposed. Many design tools that operate in a VR environment have been released to general users (e.g., Tilt Brush by Google, Gravity Sketch, VR Sketch). Lee et al. (2018) proposed a method of customising the dimensions and configuration of desks and shelves to fit the body conditions of the end-user in a VR environment, free from physical constraints. Thus, design reviews and design activities in the virtual world are gradually becoming common. However, the usefulness and characteristics of the design action itself in the VR environment remains relatively unexplored. Therefore, this study will comparatively analyse the characteristics of design behaviour in a desktop environment and in a VR environment, revealing the differences between the two.

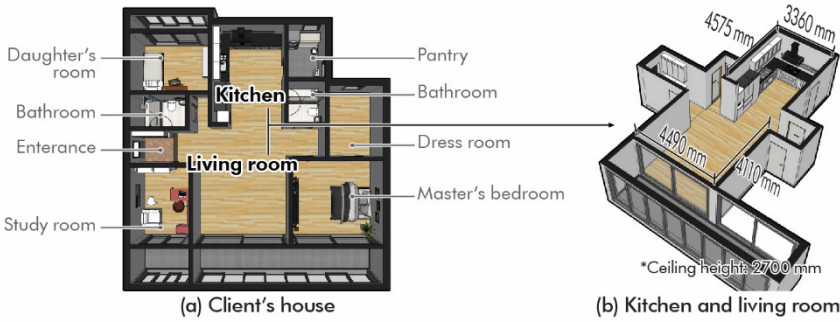
### 3. Methods

We conducted a comparative experiment to determine the difference between the VR and desktop environments in the interior design process. Interior design is one of the design domains in which designers first utilised the 3D CAD system. Owing to high complexity in interior design process, it becomes necessary to judge the shape and size of furniture, such as sofas and tables, while also determining the positional relationship of each in the space. To review the difference between the two environments, we conducted an experiment on arranging furniture in virtual and desktop environments.

#### 3.1. EXPERIMENTAL SETTINGS

The experiment required the participants to create an interior design based on the design brief. The participants were asked to decorate an apartment's living room and kitchen space in both the desktop environment and the VR environment (figure 1-a). The living room and kitchen area are approximately  $43.3 m^2$ , and the details of the space are shown in figure 1-b. The clients for the experiment were a family of three, and design requirements (figure 1-c) were included in the design brief.

The participants were instructed to design using SketchUp Pro 2021 (SketchUp) in the desktop interface (DI) and the 3D model of the living room and kitchen scaled to real-life dimensions were provided in advance (figure 2-a). In addition, 3D models of various furniture items were also supplied as a component library (figure 2-c; 20 chairs, 8 sofas, 8 tables, 6 bookcases, and 7 lights). Users were able to easily position in 3D space by drag-and-drop in SketchUp's component window (red box in figure 2). The participants selected furniture in accordance the design brief from the furniture library and then designed the space using select, move, scale, rotate, and erase functions. Participants were able to change the viewpoint (orbit, zoom, and pan) using the mouse's wheel button and the dimensions could be measured directly using the measure tool. A desktop interface with a 27-inch monitor was used, and the participants sat at a desk and performed tasks with a mouse and keyboard.



Client's family members	Age	Sex	Job	Hobby
<b>Member 1. Father</b>	44	Male	Novelist	Reading books, Having coffee time
	[Needs]	He has a personal study room, but wants to read and write novels on a laptop in the living room.		
<b>Member 2. Mother</b>	44	Female	Public official	Cooking, Having coffee time
	[Needs]	She wants to have dinner with the family not only on the kitchen table, but also in the living room.		
<b>Member 3. Daughter</b>	16	Female	Student	Reading books
	[Needs]	She wants to have a personal bookshelf in the living room because of her small room.		

(c) Client's family members and their personal information and needs

Figure 1. (a) Floor plan of client's house; (b) dimensions of kitchen and living room; (c) client's family members and their personal needs

The participants were asked to design using VR Sketch in the VR environment. VR Sketch is a program that allows 3D modelling to be performed in a VR environment through a consistent interface with SketchUp. The same design brief, 3D model, and furniture library as in the desktop environment were provided. In addition, participants could select, move, scale, rotate, erase, and measure tools even in the VR environment. Users can orbit, zoom, and pan using a two-handed controller and explore the 3D space by moving their bodies to different locations in space and changing their viewpoints. VR Sketch provides a unique teleport tool that moves the viewpoint to the location pointed to by the user as a teleport point, and at the same time, it allows the user to view the 3D model on a 1:1 scale. Oculus Quest 2 was used in this experiment, and VR design was carried out in a space of 3.5 m x 4.5 m.

To explore the design experience differences between the two environments, each participant was asked to perform the same design task in both the VR and desktop environments. To minimise the learning effect bias, half of the participants first performed in the VR environment, and the other half in the desktop environment. Next, the environments were reversed and the participants were given the same design task, but they were instructed to work toward achieving different results than they had achieved in the previous environment. Unlike the previous environment, we intended to analyse their judgments on design changes in different environments. There was no time constraint in either environment. Before designing in the VR environment, a tutorial session explaining the essential functions of VR Sketch (e.g., insert components, move, copy) was held for approximately 10 min.

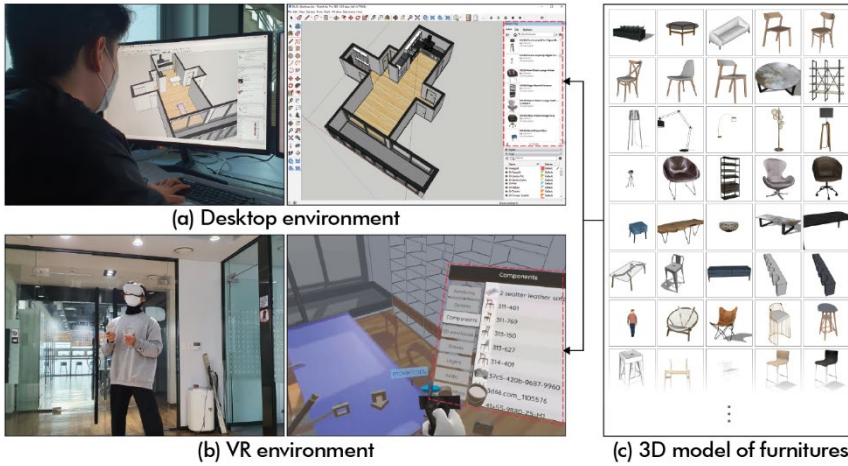


Figure 2. Experimental settings

After completing the design tasks, the participants completed a questionnaire survey on their design experience in both environments. The questionnaire consisted of the System Usability Scale (SUS) Brooke, 1996) and satisfaction questions about design results (Table 1). Responses were received on a 5-point Likert scale for each item, with one being 'strongly disagree' and five 'strongly agree'. Semi-structured interviews were then conducted. The experiment took approximately 1 hour in total.

Table 1. Items of Satisfaction Questionnaire

Satisfaction Questionnaire	
S1	I am satisfied with the quality of the design output.
S2	I am satisfied with the time it took to complete this task.
S3	I am confident that the outcome of this assignment was creative.
S4	I am confident that the results of this assignment are convincing.
S5	I am satisfied with the ease of working on this assignment.

Participants were recruited from interior design and industrial design majors who can use SketchUp with visual acuity (or corrected visual acuity) of 20/20 or higher, and four participants participated in the experiment (mean age = 27.3 y; SD = 3.2; one female). They all had an average of 3.1 years of experience using SketchUp (SD = 2.3), and all had prior experience with VR devices. One of them owned a VR device and used it more than once a week, while the remaining were infrequent users of VR devices (two people less than once a year, one more than once a year).

### 3.2. RESULTS AND DISCUSSIONS

Regardless of the experimental sequence, the task completion times were longer in VR than in DI. The average task completion time for the VR environment was 27.8 min

(SD = 3.2) while that for DI environment was 18.75 min (SD = 1.0). With a difference of approximately 10 min, the VR work took approximately 48% longer than the DI.

The results of the SUS and satisfaction questionnaire are shown in figure 3. All SUS and satisfaction results did not have normality (Shapiro-Wilk test;  $p > .05$ ). Therefore, we did not perform a statistical analysis. The mean SUS score of VR was 85.0 (SD = 13.1) and that of DI was 68.8 (SD = 13.6). The mean SUS score for both environments was calculated using the method described by Brooke (1996). An SUS score of 68 indicates the top 50% or average system usability. The experiment participants evaluated the usability of the DI environment as average and that of VR environment as above average.

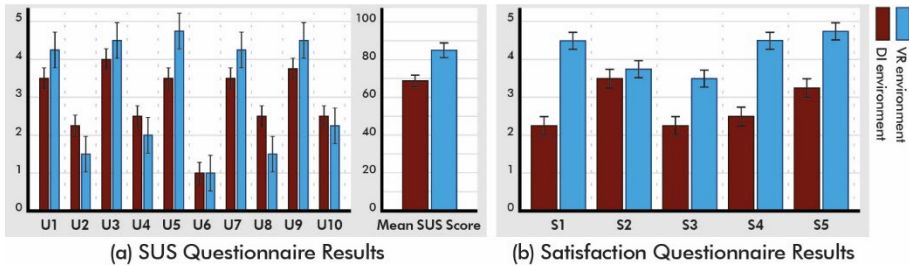


Figure 3. Questionnaire results

In addition, in the satisfaction results, the participants rated their work experience higher in VR. In particular, it showed the most significant design satisfaction compared to DI in the ‘persuasiveness of the results’ item. Participants were confident that their results in the VR environment were designed based on more rigorous evidence. In VR environment, all participants were able to visually check the size and shape of the space and furniture in a 1:1 scale space while selecting furniture and adjusting its size and location. Although it was possible to directly check the scale of the space and furniture with the Tape Measure Tool in the desktop environment, it was difficult to judge whether the size of the table or the width of the space was suitable for the human scale only with the numerical value. Following are the parts of the participant responses in the post-interview.

P1: "During the design process, it was helpful to 'feel' the size of the table directly."

P2: "Feeling a sense of space on a 1:1 scale gave me the confidence that I was doing well on my own."

P3: "VR allowed me to evaluate and modify my designs in real time. ... Checking the 1:1 scale directly helps to make a better decision."

P4: "Even if I look at the same furniture, the feeling is different (between VR and DI environments) ... It influenced my design decisions."

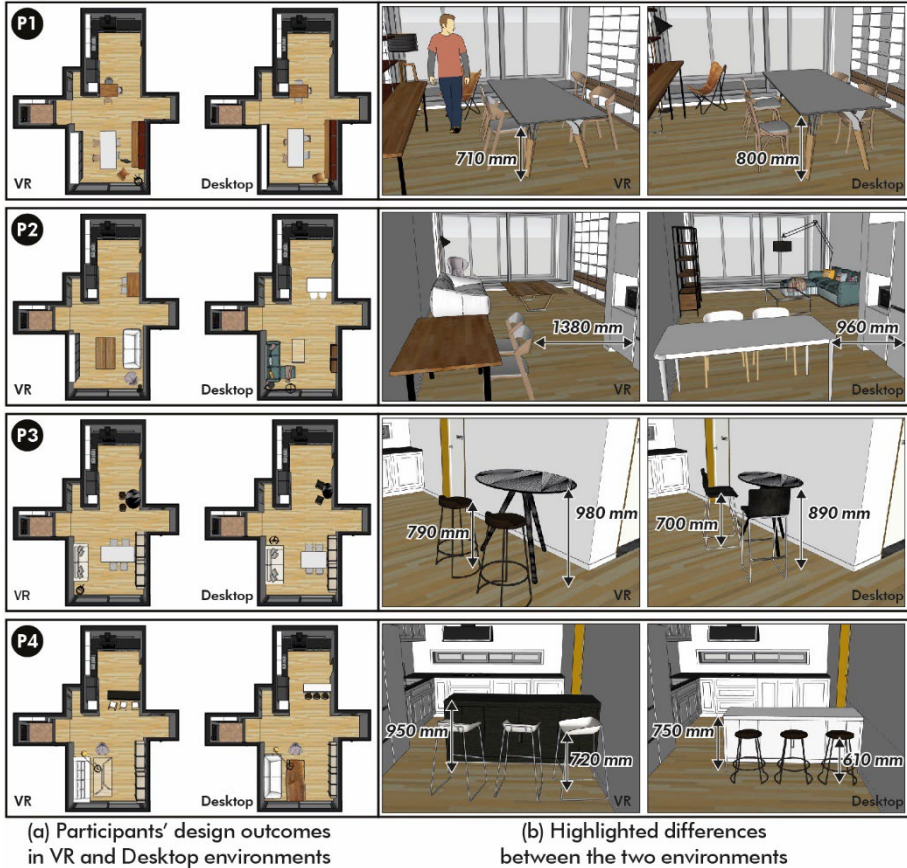


Figure 4. Design outcome from the experiment participants

The participants' design also reflected the characteristics of the VR and DI environments (figure 4). For example, in the VR design, P1 resized the height of the living room table (the original height of the table component was 800 mm) to 710 mm (figure 4-b). After arranging the table and chairs, P1 resized them to 710 mm by bending the knees and checking the height of the table from the perspective of sitting on the chair. However, P1 did not resize the height of the table in the DI work to 800 mm. Considering that the recommended table height for adults is 710–760 mm, recognising the actual size of furniture in the VR environment affected the design of P1. In another case, P3 adjusted the height of the bar table in the kitchen to be higher than that in the DI by matching the height of the bar stool in VR (figure 4-b). The natural sense of space in VR influenced the size of the furniture and the confirmation of the movement. In the DI result of P2, the aisle width next to the kitchen table was 960 mm, but in the VR result, the desk was rotated 90° to secure a wider circulation (aisle width 1380 mm). In addition, as in DI, when P2 arranged the bookshelf and P2 viewed the living room from the kitchen, P2 checked the side of the bookshelf to see that the window was not visible. Then P2 moved the bookshelf to the opposite side

(figure 4-b). In the post-interview, P2 answered the following:

“I tried to make it the same as what I did on the desktop in VR, but in reality, the space was too small... When looking from the kitchen, the bookcases covered the windows, so the space seemed small. So, I moved the bookshelf to the other side.”

In the two design results of P4, there are differences in the colour and style of the selected furniture.

“When I checked it with VR, (the bar table in the kitchen) did not match the colour of the wall. So, I switched to a black bar table and a white stool.”

Comparing the results of the two work environments, it can be confirmed that the participants used a 1:1 scale sense of space for furniture arrangement and size determination in the VR environment. Although they could check the exact dimension in DI with the Tape Measure Tool, it was more useful to check the actual size than the exact figure. This determines that the VR's accurate spatial sense made participants more confident in their designs.

#### 4. Conclusion and Future Works

This study analysed the differences in spatial design experience and design outcome depending on the VR and desktop environments. The user study results revealed that the design task completion time was longer in the VR environment than in the desktop environment. However, design satisfaction with design experience in VR was significantly higher than that in the desktop environment. Specifically, the reliability and persuasiveness of the design outcome created in the VR environment were higher than those in the desktop environment. This shows that immersive settings lead to a more rigorous design experience by allowing designers to review their work-in-progress design from various perspectives. In this study, the participants conducted short design tasks (approximately 40 min per participant); however, some participants reported nausea owing to prolonged use of the head mounted display after the study. Therefore, the results of the VR design experience may vary depending on the duration of the user study.

#### Acknowledgements

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.2021-0-00968, Developing State-of-Art 2D sketch to 3D model Conversion and Refinement Methods Inspired by NLP Translation Model)

#### References

- Brooke, J. (1996). SUS: A “Quick and Dirty” Usability Scale. *Usability Evaluation In Industry*, 207–212. <https://doi.org/10.1201/9781498710411-35>
- de Casenave, L., & Lugo, J. E. (2017). Design Review Using Virtual Reality Enabled CAD. *Proceedings of the ASME Design Engineering Technical Conference*, 1. <https://doi.org/10.1115/DETC2017-67878>
- Drascic, D., & Milgram, P. (1996). Perceptual issues in augmented reality. *Proceedings Volume 2653, Stereoscopic Displays and Virtual Reality System*. <https://doi.org/10.1117/12.237425>

- Freeman, I., Salmon, J., & Coburn, J. (2018). A bi-directional interface for improved interaction with engineering models in virtual reality design reviews. *International Journal on Interactive Design and Manufacturing*, 12(2), 549–560. <https://doi.org/10.1007/S12008-017-0413-0>
- Hong, T., Lee, M., Yeom, S., & Jeong, K. (2019). Occupant responses on satisfaction with window size in physical and virtual built environments. *Building and Environment*, 166, 106409. <https://doi.org/10.1016/J.BUILDENV.2019.106409>
- Horvat, N., Škec, S., Martinec, T., Lukacevic, F., & Perišić, M. M. (2019). Comparing Virtual Reality and Desktop Interface for Reviewing 3D CAD Models. *Proceedings of the Design Society: International Conference on Engineering Design*, 1(1), 1923–1932. <https://doi.org/10.1017/DSI.2019.198>
- Hosokawa, M., Tomohiro Fukuda, T., Yabuki, N., Michikawa, T., & Motamedi, A. (2016). Integrating CFD and VR for indoor thermal environment design feedback. *Proceedings of the 21st International Conference on Computer-Aided Architectural Design Research in Asia (CAADRIA 2016)* (pp. 663–672).
- Hou, M., Hollands, J. G., Scipione, A., Magee, L., & Greenley, M. (2009). Comparative evaluation of display technologies for collaborative design review. *Teleoperators and Virtual Environments*, 18(2), 125–138. <https://doi.org/10.1162/PRES.18.2.125>
- Kelly, J. W., Donaldson, L. S., Sjolund, L. A., & Freiberg, J. B. (2013). More than just perception-action recalibration: Walking through a virtual environment causes rescaling of perceived space. *Attention, Perception, and Psychophysics*, 75(7), 1473–1485. <https://doi.org/10.3758/S13414-013-0503-4/FIGURES/6>
- Lee, B., Shin, J., Bae, H., & Saakes, D. (2018). Interactive and Situated Guidelines to Help Users Design a Personal Desk that Fits Their Bodies. *Proceedings of the 2018 Designing Interactive Systems Conference*, 637–650. <https://doi.org/10.1145/3196709>
- Pletinckx, D., Callebaut, D., Killebrew, A. E., & Silberman, N. A. (2000). Virtual-reality heritage presentation at Ename. *IEEE Multimedia*, 7(2), 45–48. <https://doi.org/10.1109/93.848427>
- Willemsen, P., Colton, M. B., Creem-Regehr, S. H., & Thompson, W. B. (2004). The effects of head-mounted display mechanics on distance judgments in virtual environments. *Proceedings - 1st Symposium on Applied Perception in Graphics and Visualization, APGV 2004*, 35–38. <https://doi.org/10.1145/1012551.1012558>
- Yabuki, N., Miyashita, K., & Fukuda, T. (2011). An invisible height evaluation system for building height regulation to preserve good landscapes using augmented reality. *Automation in Construction*, 20(3), 228–235. <https://doi.org/10.1016/J.AUTCON.2010.08.003>
- Yeom, S., Kim, H., & Hong, T. (2021). Psychological and physiological effects of a green wall on occupants: A cross-over study in virtual reality. *Building and Environment*, 204, 108134. <https://doi.org/10.1016/J.BUILDENV.2021.108134>



# GIS-BASED EDUCATIONAL GAME THROUGH LOW-COST VIRTUAL TOUR EXPERIENCE

*Khan Game*

GUZDEN VARINLIOGLU<sup>1</sup>, SEPEHR VAEZ AFSHAR<sup>2</sup>, SARVIN ESHAGHI<sup>3</sup>, ÖZGÜN BALABAN<sup>4</sup>, TAKEHIKO NAGAKURA<sup>5</sup>

<sup>1</sup>*Department of Architecture, Izmir University of Economics; Massachusetts Institute of Technology*

<sup>2,3</sup>*M.Sc. Landscape Architecture, Istanbul Technical University*

<sup>4</sup>*Delft University of Technology, The Netherlands*

<sup>5</sup>*Department of Architecture, Massachusetts Institute of Technology*

<sup>1</sup>*guzden@mit.edu, 0000-0002-4417-6097*

<sup>2</sup>*afshar19@itu.edu.tr, 0000-0001-8380-2348*

<sup>3</sup>*eshaghi18@itu.edu.tr, 0000-0003-1754-7355*

<sup>4</sup>*ozgunbalaban@gmail.com, 0000-0002-7270-2058*

<sup>5</sup>*takehiko@mit.edu.com, 0000-0002-3219-3930*

**Abstract.** The pandemic brought new norms and techniques of pedagogical strategies in formal education. The synchronous/asynchronous video streaming brought an emphasis on virtual and augmented realities, which are rapidly replacing textbooks as the main medium for learning and teaching. This transformation requires more extensive online and interactive content with simpler user interfaces. The aim of this study is to report on the design, implementation, and testing of a game based on low-cost and user-friendly content for digital cultural heritage. In this project, a game aimed at inclusive and equitable education was developed using 360° images of the targeted architectural heritage geographically distributed in a pilot site. We promote lifelong learning opportunities for all, following the SDG4, aiming for quality education with the easy-to-use online platform and easy access to immersive education through mobile platforms. Towards a post-carbon future without the need for travel, computational design methods such as using 360° videos and images in combination with virtual reality (VR) headsets allow a low-cost approach to remotely experiencing cultural heritage. We propose developing and testing a GIS-based educational game using a low-cost 360° virtual tour of architectural heritage, more specifically, caravanserais of Anatolia.

**Keywords.** Digital Heritage; 360° Images; Educational Games; Caravanserais; SDG 4.

## 1. Introduction

The academic year of 2020 inevitably brought changes and challenges in formal education due to the pandemic. Consequently, the studies on online education in design, still experiential to date, explore both teaching and learning methods. The process of learning through experience, or "learning through reflection on doing," aka experiential learning, is an important part of architectural education. It has become a great opportunity to rethink conventional pedagogical methods and current pandemic experiments, using emerging technologies, and transporting instruction outside of the digital classroom (Estrina et al., 2021). The availability of synchronous/asynchronous video streaming brought an emphasis on virtual and augmented realities, which are rapidly replacing textbooks as the main medium for learning and teaching. However, the current multiplicity and complexity of online and interactive content created a need for simpler user interfaces.

This study aims to describe the design, implementation, and testing of a game based on low-cost and user-friendly content for digital cultural heritage. The game described in this study was developed using 360° images of the targeted architectural heritage, which are geographically distributed in a pilot site. This educational game aims to ensure inclusive and equitable education and promote lifelong learning opportunities for all. As stated by the United Nations' sustainable development goals, among the visions for 2030 are increasing gender equality, and providing equal access to all levels of education to vulnerable persons with disabilities, indigenous people, and children of both genders (UN, 2015). Thus, we assume that the aim of digital learning and teaching tools should be to ensure affordable and quality university education (SDG4). We promote lifelong learning opportunities for all through the easy-to-use online platform and easy access to immersive education through mobile platforms. The original contribution of this study is at two levels: visualisation and immersion. Immersion in interactive technologies is defined as the perception of a presence in the non-physical world; presence refers to the user's reaction to immersion (Slater, 2003). Immersion may be achieved by visual aspects, auditory images, and also narrative.

Digital Heritage, as defined by UNESCO (2009), "is made up of computer-based materials of enduring value that should be kept for future generations." The cultural heritage has a concrete value derived from its authenticity, as well as its digital interpretation, which has its own value (Aydin and Schnabel, 2015). Serious games (SG) -videogames designed for educational objectives- on cultural heritage, as one of many interpretations, are a new tool aiming to increase engagement with cultural content. Mortara et al. (2014, 318) argue that "the design process of an SG differs from the one of a common e-learning application as an intrinsic balance between learning and gaming should be found." Thus, the learning content has a key role, as the game interactions and mechanics cannot be regarded as a separate layer, but rather, are highly dependent on the heritage content.

Games have much in common with the architectural design process, including the emphasis of graphical representations, the dominance of a narrative, and possibilities of collaborative teamwork (Di Mascio, 2017). Similarly, Zarzycki (2016) highlights the similarity between building and video game design because, besides the visual quality, both are based on narrative. Current video games are complex and costly by nature, and this is leading scholars to explore simpler, cost-effective tools and methods.

360° videos together with images and virtual reality (VR) headsets are a low-cost way of remotely experiencing cultural heritage. In contrast to previous attempts at using 360° videos, we propose developing and testing a GIS-based educational game using a low-cost 360° virtual tour of architectural heritage, more specifically, the caravanserais of Anatolia.

The main aim of the Digital Caravanserai Project is to create a geolocated virtual experience of the caravanserais, allowing the exploration of Anatolian networks of the past. Using this heritage data, we aim to address the following three questions: How can we extend the notion of a relational database into a more holistic geolocated system? How can 360° photographic images be used as an immersive experience using simple and low-cost systems? How can the photogrammetric reality and game elements create an authentic self? The project consists of three stages. In the first, we collected data on caravanserais in architectural literature and also studied the heritage objects in the field, both to experience their geographical context and to collect images with 360° cameras. Second, through these collected panoramic images, we designed and implemented both virtual tours and an educational game using 3dvista Virtual Tour PRO software. Finally, we designed the game to encourage users to visit all the virtual sites and interact with the content, providing exhaustive coverage on the material. The game involves puzzle-solving using the learned information, which further guarantees the users attend to the content. The user interaction is monitored and stored in online platforms, allowing assessment of the interaction rate of all elements, which can be used to improve later versions of the game.

## **2. Immersion and Visualisation in Heritage Game**

Game design is about the design and aesthetics of creating a game for entertainment, while educational or serious game design focuses specifically on supporting learning on certain subjects, skills, etc. Whether educational or not, game mechanics are constraints of rules and feedback loops intended to produce enjoyable gameplay. Gameplay is the way in which players interact with a game. Educational games and game-based learning have become prominent tools in delivering an innovative learning experience. The tradition of learning by doing was conceived of by Bauhaus, and an extension of this conception is learning by playing. This is the focus of an increasing number of studies on emerging technologies, which move instruction beyond the digital classroom (Estrina et al., 2021). Before the pandemic, a few futuristic studies had already emerged on the integration of architectural education and immersive technologies, such as Augmented Reality (AR), Virtual Reality (VR), Mixed Reality (MR), but the online teaching era brought many more explorations of the interaction potential of live tools. Also, at the same time, there was an expansion of interaction from simple chatting to more attention-demanding tools, such as games.

Within the focus on heritage educational games, two definitions need to be clarified: immersion and visualisation. Immersion in interactive technologies is defined as the perception of the presence in the non-physical world. Presence is the user's reaction to immersion (Slater, 2003). Immersion may be achieved by visual aspects, auditory images, and also narrative; visualisation, on the other hand, is any technique for creating images to communicate a message. The use of immersive visualisation and more natural interaction increases the sense of presence, creating an enhanced game

experience.

This paper is a combined effort incorporating two text-based games designed and implemented by the authors: one on Iranian caravanserais (Eshaghi et al., 2021) and another on Anatolian caravanserais (Vaez Afshar et al., 2021). Also, further research has been published covering other heritage sites: a mobile platform game (Varinlioglu et al., 2017), a virtual reality application (Varinlioglu and Kasali, 2018), and an augmented reality application (Varinlioglu and Halici, 2019). Drawing on the authors' expertise both in the subject of caravanserais and in various immersion and visualization methods, this paper focuses on two components of the Khan Game: GIS-based and 360° immersive experiences.

Maps have always been an integral component of games, as the contextual layer for a gaming experience, (Tomaszewski et al., 2018: 369). An early use of geographic information was the strategy board game Risk and its video game version in 1988, based on a political map of the world. Similarly, Microsoft Flight Simulator 2.0 in 1984 portrays planet Earth with varying degrees of detail, both major landmarks and populous cities, and with a sparse landscape. The Stanford Geospatial Network Model of the Roman World (ORBIS, 2013) portrays the Roman world, not as it would have physically appeared from space, but captures environmental constraints that govern the flows of people, goods, and information. By simulating movement along the principal routes of the Roman network, the main navigable rivers, and sea routes, this interactive model reconstructs the duration and financial cost of travel in antiquity. Although produced as a commercial game, this interactive model is a unique educational source for understanding pre-modern society.

The VR game design combines the audience's live experience, using 360° panorama shooting, 3D modelling, virtual reality, and intelligent question-answering technology. Some studies suggest simpler methods and tools such as 360° videos, excluding even the headsets to prevent distraction by virtual reality (Alamäki et al., 2021). 360° videos and panoramic photos have only recently been introduced in the heritage sector, for example, Argyriou et al. 2020 produced a set of design tasks and techniques to exploit immersion in cultural heritage tour applications. Thus, our case study, as a solution to the high-cost content creation for VR, is designed to apply low-cost immersion and visualisation techniques to the phenomenon of caravanserais, the waypoints along the Silk Roads.

The Silk Roads were a network of routes connecting the ancient societies of East, Central, and Western Asia with the Mediterranean (Frankopan, 2015). This network represents one of the world's preeminent long-distance communication networks, covering the Anatolian landscape through caravan routes (UNESCO 2013). ICOMOS's thematic study on Silk Roads covers the east-west extent as far as Antioch - modern-day Antakya. Despite the importance of Anatolia in this respect, there is still doubt over Turkey's full involvement in this project, and thus over the inclusion of the Silk Road network in Anatolia (Williams 2014).

A caravanseraï, caravan palace, or han in Turkish, is a roadside inn for travellers to rest during long journeys. These are deliberately positioned within a day's journey apart, on average of every 30-40 kilometres. According to the Project Old World Trade Routes (OWTRAD), there are 154 Seljuk caravanserais on the route from Denizli to Doğubeyazıt, Turkey (Ciolek, 1999). Independent studies on Anatolian urban

networks include an in-depth but accessible research project by Branning (2019) on the Turkish khans, and studies on Anatolian caravanserais (Erdmann, 1961; İltter, 1969; Özergin, 1965; Tuncer, 2007). In contrast to these more general studies, the specific focus of the current study will be the caravanserais of Southern Anatolia during the Anatolian Seljuk State. Our pilot location is Antalya, the intersection of eight caravan roads, and thus, a prominent focus in the discovery of ancient Anatolian networks (Bakkal, 2019).

### 3. Game Design, Implementation, Playtesting, and Web Analytics

The design and development of Khan Game involved an interdisciplinary team of designers, architectural historians, and engineers collaborating to explore the possibilities offered by an easy-to-use virtual tour platform of the software, e.g., 3dvista Virtual Tour PRO. Through the modification of 360° images of existing caravanserais, the game encourages the user to solve puzzles based on historical facts in a game scenario. In the game, the avatar, a digital archaeologist, explores the geographically distributed caravanserais, each of which has a unique puzzle game that involves finding missing pieces of an inscription or other missing objects. These caravanserais are presented in an online 3D map of Anatolia, and by clicking a particular one, users can access its puzzle.

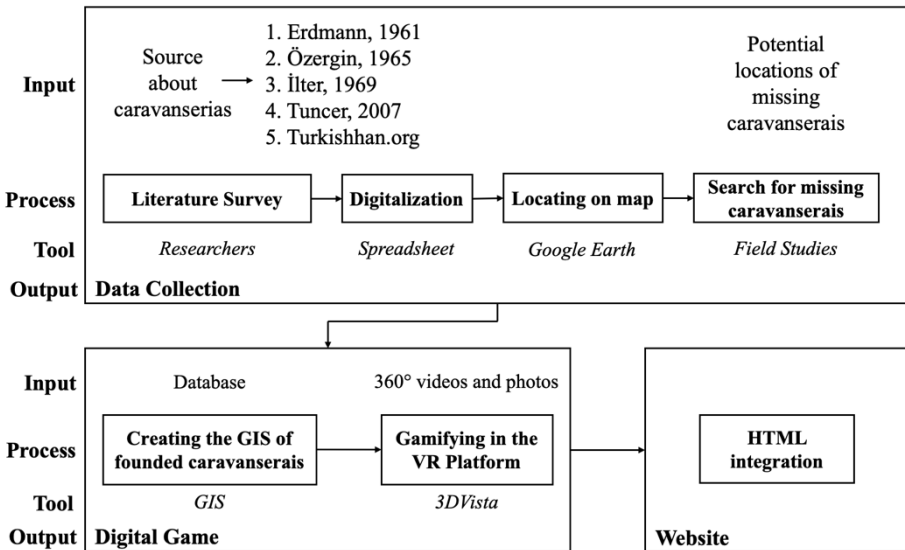


Figure 1. Methodology of the Khan Game design

There were three steps to developing the game. First, information about the selected architectural heritage, i.e., Caravanserais of Antalya, was accessed from the existing academic literature, consisting of publications, book chapters, and unpublished research, as well as notes and blogs of travellers, tourist guides, and local people (Fig. 1). We compiled this unstructured data into a spreadsheet, allowing us to test the validity of information on the location of the caravanserais, the interpretations of its

characteristics, date of construction, and other relevant aspects. The location data was compiled on the Google Earth platform, both to geolocate the data and allow comparisons within this collaborative geodatabase. The next step was the field studies, where the geo-data was reassessed to resolve any conflicting information. Finally, for each caravanserai, 360° images were carefully collected using low-budget devices, e.g., Go Pro max. This visual data was further used in virtual tours of the caravanserais.

Second, to achieve a distinctive visual language, the 360° images were compiled, and used in the design and implementation of a web-based virtual tour. For this stage, we used a commercial 3D virtual tour tool, i.e., 3dvista Virtual Tour PRO. This tool, which enables the creation of interactive virtual tours using 360° images, and is compatible with virtual reality devices and smart devices running Web browsers. Easy to publish web-based tours require no reviewing time, unlike in AppStore or Google Play. 3dvista also allows adding information boxes, media, URLs, etc. These capabilities allow the user to increase the content's immersiveness and interactivity for the target audience (Nemtinov et al. 2020). E-learning functions, which have been recently incorporated in the software, enabled generating educational virtual tours. However, while gaming is not the main aim of the software, the presence of question or quiz cards, integrated scoring, and the reporting system allows the creation of an entertaining treasure hunt game. Additionally, while studies in the literature are available using the virtual tour-making capability of the 3dvista (Perdana et al., 2019), there seems to be little research exploiting this opportunity for game development. Hence, this novel technology is worthy of deeper consideration.



Figure 2. Visual language and layers of the game

For the game's visual language, first, we edited the 360° photos to create an atmosphere of fantasy, while sharpening the focus of the architectural heritage objects and sites. Artistic experimentation resulted in 360° images composed of three layers - a layer of "heritage", displaying the currently existing architectural values of caravanserais, a layer of "landscape", displaying the actual context of caravanserais, and a layer of "fiction", narrating the scenario (Fig. 2). The team created sprite sheets

of the characters and artefacts, as well as typefaces for the visual language components of the game. The content of these tours was provided through drone and 360° videos to enhance the immersion, with the addition of maps, photos, and other information collected through the literature review.

The final step was incorporating a narrative in order to develop the gameplay. A story involving a female archaeologist was verbally and visually transcribed, set in one of the distinctive locations of the ancient site. Through intuitive movements, the player directly navigates the realm of immersive space in these 360° settings, allowing interaction with the architectural settings. Further, through the short quizzes and puzzle games, the user is able to absorb the historical information in an attractive and easy-to-digest form. The inspiration for the game's story came from real information about the Alarahan, one of the caravanserais on the list. According to Branning (2019), the upper part of the inscription at the top of the entry door of the Alarahan is missing. This information inspired us to create a story in which, during her journey, an archaeologist, the game's main character, encounters a thief in a caravanserai stealing a part of an inscription. The main goal of the game's scenario is to find the thief, but in the search, the archaeologist experiences a mysterious journey, seeking hidden objects, answering questions, and finding clues to the correct route through the caravanserai. The player first encounters the clues to find the answer to questions during the gameplay, then needs to retain that content to progress to the next part of the game, and achieve a certain score within a time limit. Hence, the game's structure encourages players to retain the provided information (Fig. 3). After completing the first level of the game in the first caravanserai, the archaeologist follows the thief's footprints through other caravanserais. The game map includes caravanserais from the Seljuk period located on the Silk Roads, and was created in the GIS using geographical locations identified during the data gathering stage.

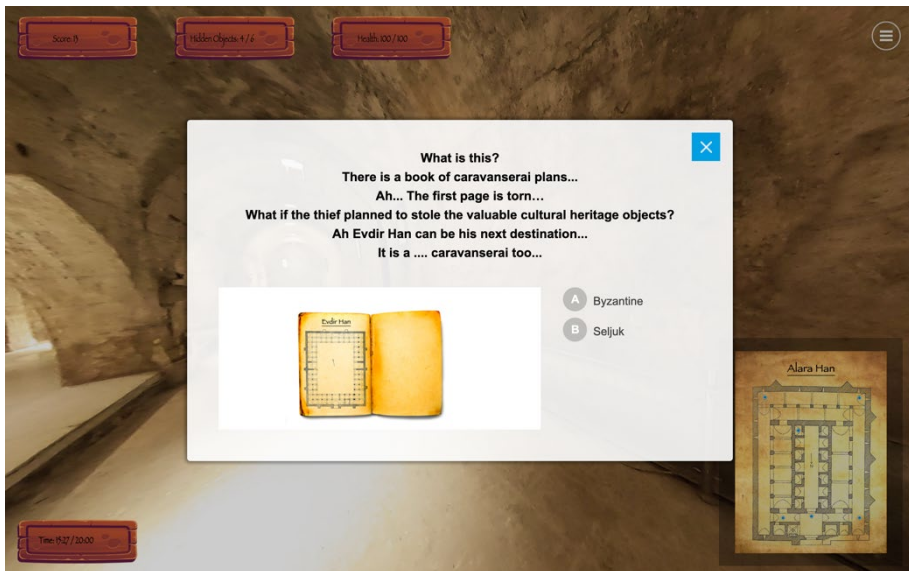


Figure 3. Screenshot of the game

Khan game is a web-based game hosted on a Linux server. When the users enter the game, a 3D map opens up displaying the terrain and the caravanserais (<https://makingindigital.com/khangame/>). This main map is implemented using Mapbox GL JS map service, a JavaScript-based map library. With the use of Mapbox, it is possible to create a web-based map, and it is also possible to link data from GIS to show up on the map. We imported the caravanserais' database location data in the game map in the form of the GeoJSON file. Moreover, it is possible to adapt the maps with the use of Mapbox Studio. Links to the other caravanserai games developed with 3dvista are visible on the info windows for each caravanserai. Within each game, users interact with puzzles and gain points when they find the objects but lose health points if they exceed the time limit. This data on the users' interaction with the puzzles, scores, health points, and user information is stored in a firebase server.

We conducted a playtest session with 30 students in November 2021 using different mediums, such as computers and mobiles. Their interactions were recorded with Google Analytics and a script. Google Analytics records all user clicks along with other anonymous user information, such as location and the browser used. The script stores the scores of each user originating from the different caravanserai games in a database. These features make it possible to track users to determine which parts of the caravanserais are less visited and which puzzles are more difficult.

#### **4. Conclusion**

In this era, there is a growing need for quality content creation in the heritage field, and to meet this need, 360° images and videos provide a cost-efficient immersive experience. To maximise its learning potential, this immersive content requires game elements to be incorporated. This research showcases a methodology for creating serious games in the heritage field. Developing a game requires expertise in programming along with many different types of software. This method allows the creation of games in a time- and cost-efficient way. However, the game is limited to the functions of 3dvista, and extra work is needed, for example, to connect the game to a database to store scores, which involves editing the game's script files. The currently available software functions in the e-learning part were not originally designed specifically for game development, but can, if used imaginatively, enable the user to create immersive and educational gameplay. In conclusion, the contribution of this study is its exploration of the intersection of game studies and heritage supported by a multidisciplinary group of professionals, without the need for popular gaming engines. The immersiveness of the 360° images is the key to maintaining a high level of engagement. As future work, we aim to shift the focus from the architectural buildings to the intangible heritage. While tangible heritage is still visible to people and is touchable, to achieve a more holistic approach to the exploration of the era of Anatolian Seljuks, we aim to gather data on their lifestyles, including their clothing, tools, mechanical devices, and other material details, as well as data on the landscape. These data are more abstract, and thus in more danger being lost. This valuable information holds within it the spirit of ancient times, and deserves to be in the spotlight of current studies.



## Acknowledgements

The field studies on Antalya caravanserais are supported by Koç University Suna & İnan Kıraç Research Center for Mediterranean Civilizations, project number AKMED 2021/P.1050. The 3dvista Virtual Tour PRO software and GoPro Max hardware has been partially supported by Izmir University of Economics, Scientific Research Commission, "Outdoor Augmented Reality for Alternative Art Spaces", contract no: BAP 2019\_08. We would like to thank Dr. Pınar Yoldaş, from UCSD for joining the field studies, and Hilal Kaleli for illustrations of game characters, Emre Kılıçkaya, Efe Şür, and Ezgi Yılmaz for testing gameplay.

## References

- 3D Vista Pro (2021). *3DVISTA*, Retrieved December 20, 2021, from <https://www.3dvista.com>.
- Alamäki, A., Dirin, A., Suomala, J., & Rhee, C. (2021). Students' Experiences of 2D and 360° Videos With or Without a Low-Cost VR Headset: An Experimental Study in Higher Education. *Journal of Information Technology Education: Research*, 20, 309-329. <https://doi.org/10.28945/4816>.
- Argyriou, L., Economou, D., & Bouki, V. (2020). Design methodology for 360 immersive video applications: the case study of a cultural heritage virtual tour. *Personal and Ubiquitous Computing*, 24(6), 843-859.
- Aydin, S., & Schnabel, M. A. (2015). Fusing Conflicts Within Digital Heritage Through the Ambivalence of Gaming. In *20th International Conference of the Association for Computer-Aided Architectural Design Research in Asia: Emerging Experience in Past, Present and Future of Digital Architecture, CAADRIA 2015* (pp. 839-848). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA).
- Bakkal, A. (2019). Antalya Selçuklu Kervansarayları. *Turkish Academic Research Review - Türk Akademik Araştırmalar Dergisi*.
- Branning, K. (2019). *The Seljuk Han of Anatolia Retrieved*. December 20, 2021, from <http://www.turkishhan.org>.
- Ciolek, T. M. (1999). *Georeferenced historical transport/travel/communication routes and nodes - Dromographic Digital Data Archives (ODDDA): Old World Trade Routes (OWTRAD) Project*. Retrieved December 20, 2021, from [www.ciolek.com](http://www.ciolek.com).
- Di Mascio, D. (2017). 3D Representations of Cities in Video Games as Designed Outcomes. In *22nd International Conference of the Association for Computer-Aided Architectural Design Research in Asia: Protocols, Flows and Glitches, CAADRIA 2017* (pp. 33-43). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA).
- Erdmann, K. (1961). *Das anatolische Karavansaray des 13. Jahrhunderts*, I, Verlag Gebr Mann, Berlin
- Eshaghi, S., & Sepehr Afshar, Varinlioglu, G. (2021). The Sericum Via: A Serious Game for Preserving Tangible and Intangible Heritage of Iran. In *9th International Conference of the Arab Society for Computer Aided Architectural Design, ASCAAD 2021: Architecture in the Age of Disruptive Technologies: Transformations and Challenges* (pp. 306-316).
- Estrina, T., Hui, V., & Ma, L. (2021). The Digital Design Build-Modes of Experiential Learning in the Pandemic Era. In *26th International Conference on Computer-Aided Architectural Design Research in Asia: Projections, CAADRIA 2021* (pp. 41-50). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA).
- Frankopan, P. (2017). *The Silk Roads: A New History of the World*. Bloomsberry.
- İlter, İ. (1969). *Tarihi Türk hanları*. Karayolları Genel Müdürlüğü Yayınları.

- Mortara, M., Catalano, C. E., Bellotti, F., Fiucci, G., Houry-Panchetti, M., & Petridis, P. (2014). Learning cultural heritage by serious games. *Journal of Cultural Heritage*, 15(3), 318-325.
- Nemtinov, V., Borisenko, A., Nemtinova, Y., Tryufilkin, S., & Morozov, V. (2020). Development of Virtual Tours of Memorable Places Associated with Residency and Activities of Famous Personalities. *International Multidisciplinary Scientific GeoConference: SGEM*, 20(2.1), (pp. 127-134).
- ORBIS (2013). *The Stanford Geospatial Network Model of the Roman World*. Retrieved December 20, 2021, from <https://orbis.stanford.edu>.
- Özergin, M. (1965). Anadolu'da Selçuklu kervansarayları. *Tarih Dergisi*, 15(20), 141-170.
- Perdana, D., Irawan, A. I., & Munadi, R. (2019). Implementation of a web based campus virtual tour for introducing Telkom university building. *International Journal if Simulation—Systems, Science & Technology*, 20(1), 1-6.
- Slater, M. (2003). A note on presence terminology. *Presence Connect*, 3(3).
- Tomaszewski, B.M., Konovitz-Davern, A., Schwartz, D.I., Szarzynski, J., Siedentopp, L., Miller, A., & Hartz, J. (2018). GIS and Serious Games. In B. Huang (Ed.), *Comprehensive Geographic Information Systems*, (pp. 369-383). Elsevier.
- Tuncer, O. C. (2007). Anadolu Kervan Yolları. *Vakıflar Genel Müdürlüğü Yayınları*.
- UNESCO (2009). *Charter on the Preservation of Digital Heritage*. Retrieved November 24, 2021, from <https://unesdoc.unesco.org/ark:/48223/pf0000179529.page=2>.
- UNESCO (2013-2022). *Silk Roads Programme*. Retrieved December 20, 2021, from <https://en.unesco.org/silkroad/content/caravanserais-cross-roads-commerce-and-culture-along-silk-roads>.
- United Nations Department of Economic and Social Affairs (2015, September). *#Envision2030 Goal4: Quality Education*. Retrieved November 24, 2021, from <https://www.un.org/development/desa/disabilities/envision2030.html>.
- Vaez Afshar, S., Eshaghi, S., Varinlioglu, G., & Balaban, Ö. (2021). Evaluation of Learning Rate in a Serious Game-Based on Anatolian cultural heritage. In *39th International Conference of Education and Research in Computer Aided Architectural Design in Europe: Towards a new, configurable architecture, eCAADe 2021* (pp. 273-280).
- Varinlioglu, G., & Halici, S. M. (2019). Gamification of Heritage through Augmented Reality. In *37th International Conference of Education and Research in Computer Aided Architectural Design in Europe, eCAADe 2019: Architecture in the Age of the 4th Industrial Revolution* (pp. 513-518).
- Varinlioglu, G., & Kasali, A. (2018). Virtual Reality for a Better Past. In *the 36th International Conference of Education and Research in Computer Aided Architectural Design in Europe, eCAADe 2018: Computing for a Better Tomorrow* (pp. 243-250).
- Varinlioglu, G., Aslankan, A., Alankus, G., & Mura, G. (2017). Raising Awareness for Digital Heritage through Serious Game-The Teos of Dionysos. In *35th International Conference of Education and Research in Computer Aided Architectural Design in Europe, eCAADe 2017: ShoCK!* (pp. 647-654).
- Williams, T. D. (2014). The Silk Roads: an ICOMOS Thematic Study. *International Council of Monuments and Sites* (ICOMOS).
- Zarzycki, A. (2016). Epic video games: Narrative spaces and engaged lives. *International Journal of Architectural Computing*, 14(3), 201-211.

# A METHOD OF VR ENHANCED POE FOR WAYFINDING EFFICIENCY IN MEGA TERMINAL OF AIRPORT

SHUYANG LI<sup>1</sup>, CHENGYU SUN<sup>2</sup> and YINSHAN LIN<sup>3</sup>  
<sup>1,2,3</sup>*Tongji University.*

<sup>1</sup>*lishuyang1995mail@gmail.com, 0000-0002-8780-7469*

<sup>2</sup>*ibund@126.com, 0000-0002-5686-5957*

<sup>3</sup>*lilianlin003@126.com, 0000-0002-6687-6457*

**Abstract.** The airport is one of the most essential infrastructures of cities. An important issue of the airport design is that passengers must be able to find their way efficiently. Although the designers adopt the post-evaluation after the operation, it takes a long time to conduct the on-site wayfinding experiment, and the number of participants of the experiment is also limited. Moreover, conventional post-occupancy evaluation suffers from security control and quarantine inspection that can not be carried out in the field. We proposed a VR enhanced POE approach that carries out an online wayfinding experiment to obtain numerous and detailed data, which significantly improves the efficiency of the post-occupancy evaluation project, and is validated by an affordable small-scale on-site experiment. Meanwhile, the cause for low wayfinding efficiencies, such as the symmetric space, the ambiguous direction and the redundant information on signboards are found and corresponding optimization suggestions are presented. The following signage system optimization project conducted in the terminal is welcomed by the passengers according to monthly questionnaires.

**Keywords.** Transportation Building; Post-Occupancy Evaluation; Digital Twins; Signage System Design; Wayfinding; Virtual Reality; Eye-Tracking; SDG 9.

## 1. Introduction

With the increase of global economic cooperation and cultural exchanges, the airport has become one of the most important infrastructures of cities. Tourism will be booming if there is an airport in the developing area, and the produces export will increase at the same time, so that the income and living standard of residents will be improved. The scale of terminal buildings has been expanding gradually because the globalization and the travel on business, accompanied by the accelerating complexity of passenger paths in the terminal. As the first stop for visitors, terminal design, especially the signage system design, will greatly affect their feelings, thus the efficiency of wayfinding becomes a crucial indicator in the post-occupancy evaluation for the airport.

### 1.1. OBSTACLES IN WAYFINDING EFFICIENCY EVALUATION

Through the post-occupancy evaluation of wayfinding, we can accumulate knowledge about airport design and signage system design. Anthony adopted the personal survey and behaviour tracing to evaluate the signage system in O'Hare International Airport (Anthony, 1991). He found that the corridor directional signs play a more important part in wayfinding rather than you-are-here maps, then he designed an experiment to compare the signs before and after optimization, the conclusion was used to suggest specific guidelines for the effective redesign. Barbara and others conducted a post-occupancy evaluation of wayfinding in a hospital through the interviews, behaviour observation, and tracking of visitors (Barbara et al, 1997). Results reveal the problems of radial floor layouts, signs, colors, and other wayfinding cues, these conclusions can provide references for hospital design in the future.

Most wayfinding research take on-site experiment, this is a huge challenge for researchers, and it is impossible to comprehensively evaluate the whole indoor space of the large-scale buildings. Considering that the strict user control of the terminal, the number of participants is limited, especially in the post-pandemic era, organizing large numbers of people to take part in on-site experiments suffer more risks. Then, the researchers are forced to choose only some section of the building to design wayfinding experiments, which makes the post-occupancy evaluation is incomplete and less credible. Moreover, researchers spend a lot of time recording and collating data of participants' trajectories and behaviours in on-site experiments, which results in a long period for the post-occupancy evaluation.

Compared with the on-site wayfinding experiments, the virtual wayfinding experiments are more feasible, but it also suffers from credibility. Virtual environments can be reproduced verisimilitudinous with mapping the material textures to 3D model (Kuliga et al, 2020), but the surrounding crowds, lighting, and other factors also affect the participants' wayfinding behaviour. Therefore, some researchers design the same wayfinding experiment conducted both on-site and in the virtual environments to verify the reliability of virtual wayfinding experiment (Schwarzkopf et al, 2013). Besides, the number of participants is limited because the virtual experiments need a set of VR equipment with the head-mounted display or the CAVE VR, similar to on-site wayfinding experiments.

### 1.2. EXPLORATION OF VIRTUAL WAYFINDING EXPERIMENT

From the perspective of influencing factors of the wayfinding process, Carpman, Grant, and other scholars summed up three elements: behaviour elements, design elements, and operational elements (Carpman and Grant, 2002). Both the conventional on-site wayfinding experiment and the virtual wayfinding experiment in recent years are based on the above theory. Then, a VR wayfinding experiment is designed to minimize the differences between the experimental scene and the real scene in the three elements. The performance and trajectory of the participants are recorded. Finally, the researchers analyze the data and make conclusions.

Helmut and others built a virtual interior environment of the main railway station of Vienna (150m × 300m, 3 levels). With the help of mobile eye-tracking, they saved lots of time in the wayfinding research because the eye-tracking data can be projected

onto the 3D model (Helmut et al, 2016). Xu and others constructed a virtual space of three subway stations of Shanghai and conducted experiments to explore how the layout of the signboards affects passengers' wayfinding behaviour (Xu et al, 2010). Benefiting from the virtual environment can be adjusted according to the experimental requirements, this experiment was completed in 4 days with 120 participants at an affordable cost. Sun and Yang proposed a research approach to study the wayfinding process in the virtual environment based on eye-tracking technologies (Sun and Yang, 2019). Through the analysis of the trajectories (recorded by virtual wayfinding experiment platform) and cognitive map, they investigated the mechanism of how architectural characteristics impose an impact on individuals' wayfinding behaviour.

Previous studies have shown that passengers mainly focus on the signboards and architectural environment in the wayfinding process. The above two objects can be completely reproduced in the virtual environment, so the virtual wayfinding experiment is adopted by more and more researchers.

## 2. Methodology

As a powerful supplement to the conventional wayfinding experiment and post-occupancy evaluation methods, this study proposes a VR enhanced post-occupancy evaluation method based on the self-developed online virtual wayfinding experiment platform (Desktop VR). This method can facilitate systematic virtual experiments with large numbers of participants in that it can avoid obstacles in the on-site experiment. It also betters the credibility of the virtual experiment that used to be only tested in a small-scale on-site wayfinding experiment. Eye-tracking technology is used in the on-site experiment so that abundant data can support the conclusion of the virtual experiment.

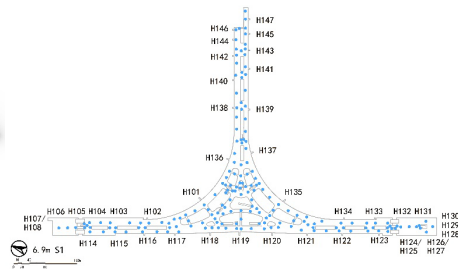
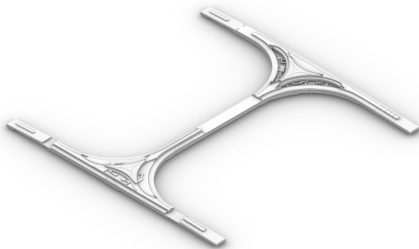


Figure 1. The low-level detailed 3D model

Figure 2. The layout of wayfinding decision points



*Figure 3. The high-resolution panorama was taken in the Satellite Terminal of PVG Airport*

## 2.1. ONLINE WAYFINDING EXPERIMENT PLATFORM

Based on Unity 3D, a virtual experiment platform is developed. First, after the field research, a low-level detailed 3D model (Fig. 1) within the scope of this evaluation is constructed according to the CAD drawings and BIM model of the PVG airport. Then, as stated by the theory of wayfinding, the walkable area and wayfinding decision points (passengers will decide to turn left, right, or go ahead at this point) are determined (Fig. 2). To make the participants have a fluent experience of wayfinding, a point is added between two wayfinding decision points if they are far away from each other, and a total of 1235 panorama images (Fig. 3) with 8K resolution are taken at each point. Based on the method of “the high-resolution panoramas + the low-level detailed 3D model”, the interior space of the Satellite Terminal of PVG airport is reproduced as a virtual reality scene.

Desktop VR wayfinding experiment has some credibility and is a suitable option for VR experiments with larger numbers of participants. Compared with HDM VR, Desktop VR is overall similar when it comes to the user experience and wayfinding decision. Some experiments provide empirical evidence supporting researchers to choose non-immersive VR when studying passengers' wayfinding behaviour (Yan, 2021).

Publishing the virtual wayfinding experiment platform as a Web-based application can attract more people to take part in this experiment. Participants can access the website ([pvg.plans.run](http://pvg.plans.run)) through various browsers (Chrome, Firefox, Edge, and Safari), and the online experience requires affordable computer hardware which is beneficial for the experiment. The experiment was started by filling in personal information (gender and age) on the login interface, and a random assignment mechanism was set up for the objectivity of the experiment.

## 2.2. SETTING AND PARTICIPANTS

The VR enhanced post-occupancy evaluation consists of two experiments. The first is the virtual wayfinding experiment with a large number of participants, and the second is the on-site wayfinding experiment with a small number of passengers as subjects.

In the online wayfinding experiment, 175 participants performed as passengers to

complete the given wayfinding task (Fig. 4). These participants were college students, teachers, and volunteers between 18 and 46 years, with an approximately equal number of men and women. All the participants had the experience of taking the plane and knew the boarding process, but they had never been to the Satellite Terminal of PVG airport.

In the on-site wayfinding experiment, participants were equipped with a mobile eye tracking device (Dikablis Glass 3). We recruited 8 passengers (in the Satellite Terminal of PVG airport) as subjects who took part in this experiment (Fig. 5).

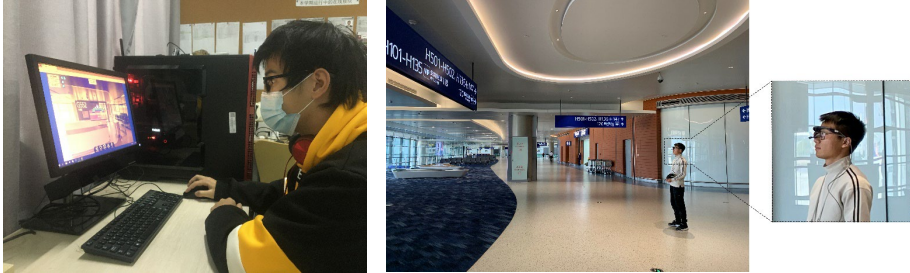


Figure 4. Online wayfinding experiment Figure 5. On-site wayfinding experiment with eye-tracking

### 2.3. VIRTUAL WAYFINDING EXPERIMENT

There are 252 kinds of specific wayfinding tasks including different entrance match with different boarding gate. Participants received random tasks when they enter the virtual environment, the essential information of the task (boarding deadline, gate number, and flight number) was recorded on the virtual ticket (Fig. 6) so that participants can check it anytime.

Participants spent about 45 minutes in the virtual wayfinding experiment. In the first 5 minutes, participants had to get familiar with the operation on the online experiment platform. Then, in the next 40 minutes, participants were required to complete randomly assigned wayfinding tasks, each participant completed about 10 tasks on average in this period, a total of 1861 wayfinding experiments were completed. The same wayfinding task was assigned to at least 5 participants to avoid occasional decision errors, but no participant will receive the same task again.



Figure 6. Virtual ticket in online experiment platform Figure 7. Trajectories of virtual experiment

The movement of participants would be recorded as points in Cartesian coordinates, and the trajectory is obtained by connecting these points in chronological order. By overlapping the paths of different participants who complete the same task on the building plan, we can quickly find the wayfinding decision points where participants make wrong decisions (Fig. 7).

#### 2.4. ON-SITE WAYFINDING EXPERIMENT

The on-site experiment consists of 10 wayfinding decision points. We selected 8 wayfinding decision points that participants make wrong decisions frequently in the virtual experiment as a trial group. 2 wayfinding decision points without any decision error were added as a control group.

A total of 80 on-site wayfinding tasks were completed in this experiment. 8 wayfinding tasks were set up at each decision point, with 8 passengers wearing the mobile eye tracking device who need to complete a randomly selected wayfinding task.

After hearing the experimenter dictate the wayfinding task, the participants began to observe the interior space and search for the signboards. The experiment ended when the participants reached the right boarding gate. Finally, participants described the experience feeling and reviewed the most confusing moment during the wayfinding process, which are recorded by the experimenter.

Based on the data recorded in the mobile eye-tracking device, the reason for the participants to make wrong wayfinding decisions (for example: ignoring a specific signboard or being guided by incorrect information on the signboard) can be found. Moreover, the participants' visual behaviour during the wayfinding process can be explored.

### 3. Results and Discussion

This is an example of the first paragraph of body text after a heading. These paragraphs do not have an indentation of the first line. Use the style 'CAADRIA text first' for the first paragraph after any type of heading.

All other paragraphs should be formatted using the 'CAADRIA text' style. These paragraphs have a 0.5 cm indentation of the first line. Please do NOT adjust any of the styles, fonts, line spacing etc., only the CAADRIA styles should be used to format your document.

#### 3.1. VALIDATION OF VIRTUAL WAYFINDING EXPERIMENT

By comparing the trajectories of the virtual experiment and the on-site experiment in the same wayfinding tasks, we can test the credibility of the virtual wayfinding experiment.

Assessing the 8 wayfinding decision points of the on-site experiments, we find that the trajectories of the on-site experiment are highly similar to the trajectories in the virtual experiment, that is, the decision errors made by the participants being the same as those in the virtual experiment.

For instance, in the on-site experiment, passengers look for the boarding gate H138 and choose the path on the right (Fig.8 and Fig.9), which is the same choice of the



participants in the virtual experiment (Fig. 7), but the shortest path is on the left.

However, in terms of the 2 wayfinding decision points of the control group, all the participants made the right wayfinding decision which are same as the virtual experiment. Therefore, based on the results of the virtual wayfinding experiment being consistent with the results of the on-site experiment, we assume that the VR enhanced POE method is a valid technical path and we can obtain reliable evaluation results from it.

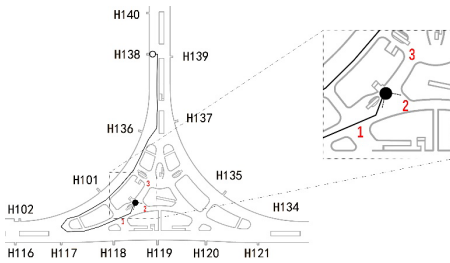


Figure 8. Trajectory of on-site experiment



Figure 9. Data visualization of eye-tracking

### 3.2. PROBLEMS OF WAYFINDING AND OPTIMIZATION

#### 3.2.1. The Symmetric Space

From the layout of the trajectories, around 78% of passengers (137 participants) took a detour when they pass through the atrium. However, only 13% of passengers took a detour in the airside concourse. The starting point and some critical wayfinding decision points of the international departure are located at the intersection of the T-shaped atrium, and the spatial arc interfaces to the three directions seem extremely similar from the standing locale (Fig. 10). Reviewing the eye-tracking data, more than three quarters of passengers looked around and observed repeatedly at these wayfinding decision points, and looked back frequently at the direction they came from although there are no signboards. Oral statements also prove the above conjecture, passengers are confused when they observe, sometimes they mistakenly think they have been to this place when they first arrive, which is the main cause of the wayfinding problem.

Considering that the form of architecture space cannot be changed, it is suggested that interior designs should be used to imply the difference between different directions, such as setting unique sculptures and installations or using different colors in three directions (Fig. 11).



Figure 10. T-shaped atrium



Figure 11. Installations and colors in three directions

### 3.2.2. The Ambiguous Direction and The Redundant Information on Signboards

The oral statement showed that around 65% of participants were confused and feel anxious when they see excessive signboards in the observation process, especially some of which had the same wayfinding information about the boarding gates but pointing to different directions (Fig. 12). From the layout of trajectories (Fig.13), we can see the consistent paths become to diverge in some specific wayfinding decision points, and some paths are obvious detours. Thus, designers can accurately locate the signboards with ambiguous directions and then correct them.

Besides, the eye-tracking data reflects that passengers spend more time before making a decision when they are located at a multi-directional decision point with a large number of signboards. A large number of signboards can inevitably be seen in the multi-directional wayfinding decision points, the same wayfinding information are repeated on signboards which located in different directions, and all text on signboards are the same size and color (Fig. 14). About half of the participants ignored some important information when observing the signboards because they have lost patience in the long-period wayfinding process. It can be said that information overload triggers typical brain fatigue, passengers therefore may make wrong wayfinding decisions.

In fact, certain rules of the sequence of observation have been found through analyzing eye-tracking data: the participants' sight will be first attracted by some signboards with different background colors from a large number of signboards, and the luminous part of the signboard will also be given priority. So, it is suggested to define the priority level of the wayfinding information, and distinguish different levels of attraction via multiple visual expressions (Fig.15). Besides, it is desirable to reduce the redundant information on signboards to avoid misleading and to improve the wayfinding experience.

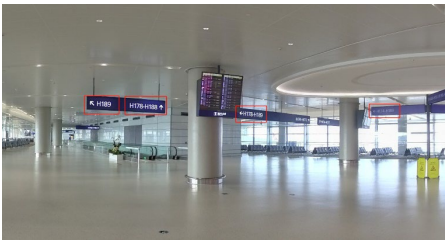


Figure 12. The ambiguous direction on signboards

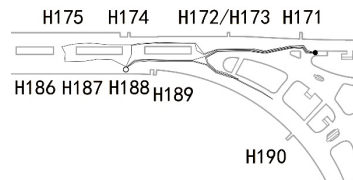


Figure 13. Discrepancy trajectories

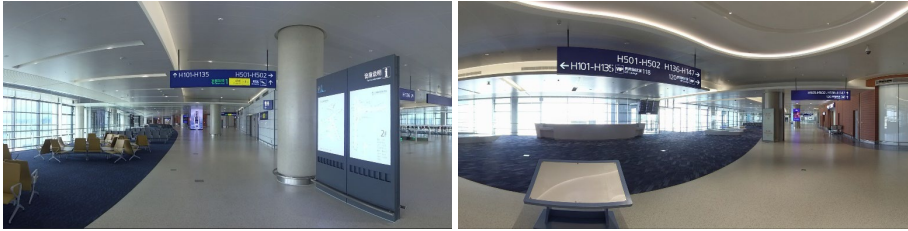


Figure 14. Signboards in Satellite Terminal of PVG airport before optimization



Figure 15. Optimization of signboards design

#### 4. Summary and Outlook

In architectural practices and research, post-occupancy evaluation (POE) is an effective way to understand to what extent the built environment runs as it was designed. Only with proper evidence can further optimization and improved knowledge be expected.

We proposed a VR enhanced POE method that through online VR wayfinding experiments to obtain sufficient data, which significantly improves the efficiency of the POE projects and is validated by the affordable on-site data. VR enhanced POE took only 2 weeks to complete the wayfinding evaluation, which saved nearly 3 months compared with the conventional on-site wayfinding evaluation. Meanwhile, using the VR enhanced POE method can reach the aim of the conventional POE project under the condition of limited budget and epidemic spread. So, the VR-enhanced POE approach is feasible and replicable.

Furthermore, we can quickly locate the wayfinding decision points that may have potential problems that lead to detours through the VR enhanced POE project. Then, three major reasons are summarized, such as the symmetric interior space, the ambiguous direction and the redundant information on signboards, which claims subsequent design updates in interior space and signage system. As a result, the designers accepted our suggestions and made improvements to the interior design and signage system. The following signage system optimization project conducted in the terminal is welcomed by the passengers according to monthly questionnaires, around 60% reduction of complaints compared to the previous.

It is also found that in the on-site experiment, the participants are subconsciously

influenced by passengers around, especially when they are confused, which coincides with the results of some previous studies. Therefore, we are trying to introduce avatars as background crowds into the virtual environment to improve the credibility of the virtual experiment.

### Acknowledgements

The research was supported by a project of Natural Science Foundation of China titled "An internet-plus-based approach of crowd simulation for public buildings" (no. 51778417).

### References

- Anthony, A. (1991). Human Orientation and wayfinding in airport passenger terminals. *Transportation Research Record*. (1298). 25-32.
- Arthur, R., Passini, R. (1992). *Wayfinding: People, signs, and architecture*. McGraw-Hill Book Company.
- Barbara, B., Holly, W., Craig, B. (1997). A post-occupancy evaluation of wayfinding in a pediatric hospital: research findings and implications for instruction. *Journal of Architectural and Planning Research*, 14(1), 35-51.
- Carpman, J., Grant, M. (2002). Wayfinding: A broad view. In Bechtel, R. B., & Churchman, A. (Eds.), *Handbook of Environmental Psychology* (pp. 427-443). John Wiley & Sons.
- Fewings, R. (2001). Wayfinding and airport terminal design. *Journal of Navigation*, 54(2), 177-184.
- Kuliga, S., Charlton, J, Rohaidi, H. F., Isaac, L. Q. Q., Joos, M., Christopher, H., & Michael, J. (2020). Developing a replication of a wayfinding study: From a large-scale real building to a virtual reality simulation. In *Spatial Cognition XII - 12th International Conference, Spatial Cognition 2020*, (pp. 126-142).
- Schrom-Feiertag, H., Settgast, V., Seer, S. M., (2016). Evaluation of indoor guidance systems using eye tracking in an immersive virtual environment. *Spatial Cognition & Computation*, 17(1-2), 163-183.
- Schwarzkopf, S., von Stülpnagel, R., Büchner, S. J., Konieczny, L., Kallert, G., & Hölscher, C. (2013, September). What lab eye tracking tells us about wayfinding a comparison of stationary and mobile eye tracking in a large building scenario. In *Eye Tracking for Spatial Research, Proceedings of the 1st International Workshop (in conjunction with COSIT 2013)* (pp. 31-36).
- Skorupka, A. (2009). Comparing human wayfinding behavior in real and virtual environment. In *Proceedings of the 7th International Space Syntax Symposium*, (pp. 1-7).
- Sun, C., Yang, Y. (2019). A study on visual saliency of way-finding landmarks based on eye-tracking experiments as exemplified in harbin kaide shopping center. *Architectural Journal*, 02, 18-23.
- Tang, M., Auffrey, C. (2018). Advanced digital tools for updating overcrowded rail station: Using eye tracking, virtual reality, and crowd simulation to support design decision-making. *Urban Rail Transit*, 4(4), 1-8.
- Weisman, J. (1981). Evaluating architectural legibility: Way-finding in the built environment. *Environment and behavior*, 13(2), 189-204.
- Xu, L., Zhang, W., Tang, Z. (2010). A virtue reality study of wayfinding and the sign layout in subway station. *Architectural Journal*, 1, 1-4.
- Yan, F., Dorine, C. D., Serge, P. H. (2021). Wayfinding behaviour in a multi-level building: A comparative study of HMD VR and Desktop VR. *Advanced Engineering Informatics*, 51, 101475.

# DEVELOPING AN AUGMENTED REALITY SYSTEM WITH REAL-TIME REFLECTION FOR LANDSCAPE DESIGN VISUALIZATION

*Using real-time ray tracing technique*

HAO CHEN<sup>1</sup>, TOMOHIRO FUKUDA<sup>2</sup> and NOBUYOSHI YABUKI<sup>3</sup>

<sup>1,2,3</sup>*Division of Sustainable Energy and Environmental Engineering,  
Graduate School of Engineering, Osaka University*

<sup>1</sup>*chen.hao@it.see.eng.osaka-u.ac.jp, 0000-0002-5296-9775*

<sup>2</sup>*fukuda.tomohiro.see.eng@osaka-u.ac.jp, 0000-0002-4271-4445*

<sup>3</sup>*yabuki@see.eng.osaka-u.ac.jp, 0000-0002-2944-4540*

**Abstract.** In landscape design, visualization of a new design on the site with clients can greatly improve communication efficiency and reduce communication costs. The use of augmented reality (AR) allows the projection of design models into the real environment, but the relationship between the models and the physical environment, such as reflections, which are often thoughtfully considered in waterfront landscape design, is difficult to express in existing AR systems. The aim of this study is to accurately render and express the reflections of virtual models in the physical environment in an AR system. Different from traditional rasterized rendering, this study used physically correct ray-tracing algorithms for reflection rendering calculations. Using a smartphone and a computer, we first constructed a basic AR system using a game engine and then performed ray-tracing computations using a shader kernel in the game engine. Finally, we combined the rendering results of reflections with the video stream from a smartphone camera to achieve the reflection effect of a virtual model in a physical environment. Both designers and clients could review the design with a realistic reflection on an actual water surface and discuss design decisions through this system.

**Keywords.** Augmented Reality (AR); Reflection; Landscape Design; Interactive Visualization; Real-time Rendering; Planar Reflection; Real-time Ray Tracing; SDG 11.

## 1. Introduction

For visualization of landscape designs, it is crucial to review the context between design targets and their surrounding environment. If they are inadequate, time and social resources will be wasted in the subsequent construction process. Augmented reality (AR) technology can visualize the relationship between a design object and its

surrounding environment (Giunta et al., 2018). Visual AR is a composition wherein a virtual object is rendered and then is placed upon the image of the physical environment by synchronizing their cameras (viewing direction, field of view and maybe more like lighting adjustments, colour corrections).

One of the current problems in applying AR to the design process is the lack of inherent interaction between the design model and its intended actual environment. For example, although the phenomenon of reflection and refraction on a water surface or glass is one of the essential design elements in landscape design (Booth, 1989), it cannot be achieved with existing AR systems. In other words, physical phenomena that a design model exerts on its surrounding environment cannot be completely rendered in physical space.

Adding reflection as a physical phenomenon to an AR system also requires attention to the accuracy of the reflection calculation, as poor accuracy can negatively affect the practical application of the system. The goal of this research is to develop an AR system that expresses real-time reflections between a virtual model and its physical environment. Thus, we used a more physically correct rendering method, real-time ray tracing, to improve the accuracy of reflections.

## **2. Literature Review**

### **2.1. AUGMENTED REALITY**

In one of the most accepted definitions, AR means the technology that possesses three characteristics (Azuma, 1997): 1) It combines real and virtual content; 2) It is interactive in real time; 3) It is registered in 3D.

An AR system has a display that can combine real and virtual images, a computer system that can generate interactive graphics that responds to user input in real-time, and a tracking system that can recognize the motion trail which enables the virtual contents to appear fixed in the real world (Billinghurst et al., 2015).

AR implementation relies heavily on the hardware. In this era of rapidly changing technology, at the point of this writing, there are already many organizations that have introduced iteratively improved AR devices. Most current AR devices can layer virtual 2D images onto the real world through the display of smart devices (such as tablets or smartphones) or AR glass. The latter can provide higher quality rendering and more complex interactions according to hand gestures, even immersive sound effects, etc., but its development costs are higher while the development process is more complex, requiring multi-disciplinary collaborative development. The former, using smart devices, can achieve fundamental AR effects with much smaller development costs (Phillip, 2018).

### **2.2. AR IN THE LANDSCAPE DESIGN FIELD**

AR is a technology that projects virtual information onto the physical world. In other words, the virtual information can be accessed in 3D physical environments (Kipper and Rampolla, 2013).

The complexity and site-specific nature of design projects makes it crucial for the landscape architectural designer to take field trips to the project locations (Kerr and

Lawson, 2020). The real-time rendering feature of AR and the symbiotic display of reality make the application of AR in design review sessions very promising (Azuma, 1997).

Using AR can project the virtual model of a design directly onto a real site, and even non-professional people can perceive the relationship between the design and the environment in an intuitionistic way (Broschart and Zeile, 2015), which makes the traditional landscape design move more towards participatory design.

### 2.3. THE REFLECTION BETWEEN PHYSICAL ENVIRONMENT AND THE VIRTUAL MODEL

The relationship between virtual objects and physical environments is always a crucial issue. Reflection, as one of these relationships, includes reflections of physical environments on virtual objects and reflections of virtual objects on physical surfaces.

One solution for realizing reflections of physical environments on virtual objects is capturing the physical environment in a skybox texture, so that reflections on virtual objects can be rendered by sampling that skybox texture, and this method has been widely adopted in commercial AR engines. Dos Santos et al. (2012) used Kinect depth information and real-time ray tracing rendering method to achieve reflections of a real object on a virtual one. However, the reflection of virtual objects on physical surfaces remains unavailable.

Chen et al. (2021) proposed a method for landscape design review based on rasterized rendering that can render the reflection of virtual objects on a physical surface. In this method, the virtual model's reflection is rendered by calculating its symmetric coordinates about the reflective plane. Because of the limitations of the rasterized rendering technique, the accuracy of the reflection is insufficient, and a more physical-based rendering method is needed to improve its performance.

### 2.4. REFLECTION RENDERING USING REAL-TIME RAY TRACING METHOD

Since one of the key characteristics of AR is the interactivity between users and virtual content, the virtual content is produced in the way of real-time rendering. Real-time rendering contains rasterized rendering and ray tracing. Rasterization is the process of pixelating graphics composed of vector vertices, while ray tracing is the process of tracking ray traces backward from the viewpoint to each pixel of the screen and finding all object surface points and light source information that intersects with the viewpoint to calculate an accurate rendering result (Glassner, 1989). In ray tracing, the colour of each pixel is calculated by emitting a ray from the camera to that pixel and tracing the reflection path of that ray to calculate the pixel colour based on the material, colour, and light source information of the points along the path (Tomas et al., 2021). If an intersection point on that path has a reflective attribute, then the colour information of the previous intersection point is passed on, and the final display on the screen includes the effect of the reflection. Ray tracing is a more physical-based real-time rendering technique compared with rasterized rendering.

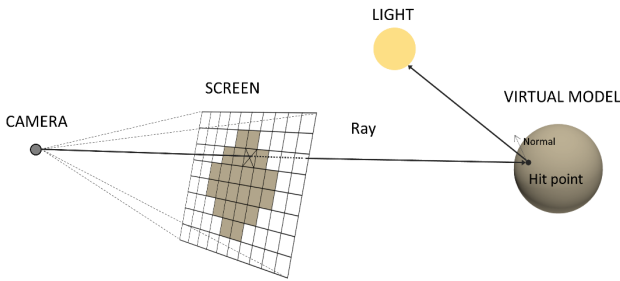
### 3. Methodology

#### 3.1. PROPOSED SYSTEM

The proposed system is composed of a smartphone and a computer. The smartphone captures a video stream of the physical environment and transmits it to the computer, and the computer renders the reflection by ray tracing and overlays the rendering results onto the video stream.

We use a game engine to build the virtual scene and perform the rendering. By passing the position information of the mobile device to the virtual camera in the virtual scene in real time, as the device moves, the camera in the virtual scene also moves according to the same trajectory. Thus, the rendering angle matches the motion perspective of the physical device.

#### 3.2. REFLECTION CALCULATION USING RAY TRACING



*Figure 1. Ray tracing rendering principle*

Since Microsoft introduced the DirectX Ray Tracing (DXR) API in 2018, different game engines have subsequently added support for DXR. This allows us to render directly in the game engine by using the encapsulated ray-tracing functionality. However, because ray tracing is highly integrated into the game engine, it becomes very difficult to make local adjustments to the algorithm. Therefore, in this study, we built a simplified version of the reflection function in the game engine based on the ray-tracing principle, so that some algorithmic adjustments can be made to make the reflection rendering layer overlay well onto the video stream.

The ray tracing is rendered according to the principle of how the human eye sees objects in realistic situations. Figure 1 shows how a scene is normally rendered using ray tracing: rays are emitted from a light source, projected onto a screen through reflections between objects, and finally converged onto the camera. Conversely, a ray is sent from the camera to the screen, the reflection of this ray in the scene is determined, and the mesh information of the intersection point is used as data for calculating the rendering result.



However, this method for directly rendering results has a problem. Given that the whole scene is in a virtual world, in addition to the virtual model, the reflection plane also reflects the whole skybox background, which can seriously degrade the immersive experience of the AR effect. To solve this problem, we developed a method to eliminate reflection from pixels in the background as follows: check the intersection of each ray, and if there is a second reflection and the intersection point of the second reflection is on the virtual model, then this ray can be recognized as reflected from the model to the reflection plane. This whole ray will be retained, while all other rays will be discarded (Figure 2). In this way, we get the rendering result that represents only the reflected part of the object on the plane.

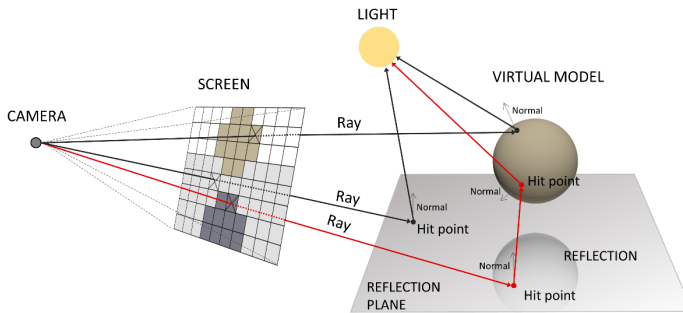


Figure 2. Ray tracing reflection rendering in our research (red ray: kept; black ray: discarded)

## 4. Implementation

### 4.1. SYSTEM CONFIGURATION

At present, the AR display methods are divided mainly into "combining real and virtual view images" and the "eye-to-world spectrum" (Billinghurst et al., 2015). To consider the effectiveness of practical application of these methods, we chose the former as the display method for practical application in our system. Moreover, we adopted the system configuration of a smartphone combined with a computer, where the smartphone is connected to the computer using a USB cable, the operator holds the smartphone to survey the scene, the AR image is output on the smartphone screen and computer at the same time, and the placement of the design model and related system parameters can be adjusted on the computer. This configuration can help us to balance the computing performance and operation in the outdoor environment.

### 4.2. DEVICE TRACKING

Tracking the motion of the viewpoint is one of the fundamental elements of AR (Azuma et al., 1998). We used the commercial AR development tool ARcore for our system configuration. ARcore calculates the exact location and movement of the viewpoint, which is based on the changes of the feature points in the image combined with the inertial measurement unit system in the device. This information is passed to the virtual camera of the game engine in real time so that the virtual camera keeps the same motion trajectory as the physical camera. Hence, when the point of view in the

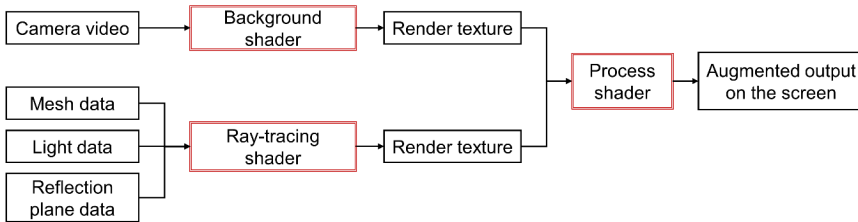
display is in motion, the virtual image rendered in the game engine will always render the corresponding angle correctly.

#### 4.3. RAY TRACING REFLECTION IMPLEMENTATION

In this study, we used compute shader, a program running on the GPU that is suitable for handling parallel computing (e.g., computing image pixels). It runs independently from the rendering processing stream and is more flexible than a normal shader.

First, the mesh of the virtual models in the scene are read by the script and transferred to the compute shader, where the position of the reflection plane is already set. A ray vector is created based on the camera coordinates and screen pixel coordinates, and the collision point coordinates (where the ray intersects the scene mesh and the reflection plane) are calculated. We set the maximum number of ray bounces to eight, which means that we ignore the ray after eight reflections. Finally, in the shading function, the colour of each pixel in the rendering result is calculated based on the reflectivity, albedo, and other material information of each collision point along the entire ray trace.

#### 4.4. MASKING AND OVERLAYING RENDERED RESULTS



*Figure 3. The workflow for masking and overlaying the rendered results*

How the reflections are properly overlaid on the camera-captured video stream greatly affects the immersive effect of AR. We developed an overlaying method for our system.

The video stream captured by the camera is processed by the AR background shader. At the same time, our ray tracing rendering result will be passed to the process shader (Figure 3). At this point, if without any processing, the ray tracing rendering result will completely overwrite the video stream. To avoid this, we masked the raytraced rendering result based on the methodology described in Section 3.2.

The specific procedure is, firstly according to our proposed ray tracing method, we obtain ray tracing extracted result (c) from the ray tracing result (a), in which all pixels except reflections are pure black. We multiply the ray tracing processed image's RGB value by 1000 (in most cases gives us a pure black and pure white image), the result is a mask (d) and is applied to the video stream (e) to get video stream masked (f). The ray tracing result is added together with the model rendered in the scene (a), masked video stream (f) and video stream(e) to get an augmented reality effect with reflections (g). We can also change the transparency of ray tracing result to increase realism (Figure 4).

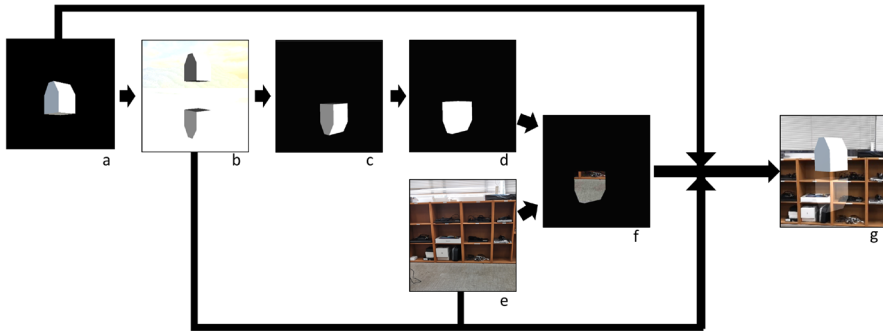


Figure 4. The composition of the video stream and rendering result. 'a' is the model rendered in the scene, 'b' is the ray tracing result, 'c' is the reflection part extracted from 'b', 'd' is the mask image generated by 'c', 'e' is the video stream from phone, 'f' is the result of applying the mask onto the video stream, 'g' is the final augmented result

## 5. Results and Discussion

### 5.1. EXPERIMENT AND RESULT

Our experiments were conducted near a water pond on Suita campus, Osaka university (Figure 5). The hardware and software used are listed in Table 1.

To make the result more realistic in the common weather situation, we made a transparent reflection in the process shader described in Section 4.4. We chose a laptop computer as the computing device and a smartphone connected to the laptop computer as the video streaming input device. The data connection is by wireless network using the AR foundation remote plug-in. The virtual model we used is a shade pavilion model with standard material in Unity. The dimensions of the base are 0.5 m\* 0.5 m\*0.5 m, the length of the umbrella pole is 3 m, and the shape of the umbrella surface is a quadrilateral cone surface with a diagonal length of 3.8 m. The mesh number of this model is 40.

In the experiment, we first positioned the virtual model within the game engine on the computer side to make it fit the physical environment. We then ran the AR system and verified that the reflection image was rendered accurately. By moving the camera of the smartphone, we could see the model and its reflection effect from different viewpoints, as shown in Figure 6. An average display rate of 15–20 frames per second (FPS) was the result.

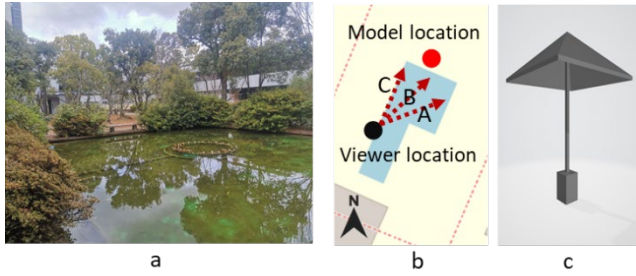


Figure 5. The left image is the photo of the experiment site; The middle image is the plan of the experiment site, in which the three arrows corresponds to the three different viewpoints directions in this experiment; The right image is the sunshade model used in this experiment

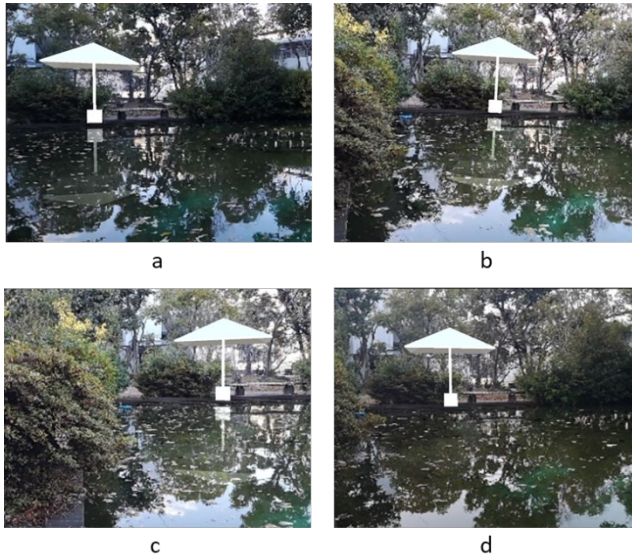


Figure 6. Experiment results, 'a' 'b' 'c' corresponds to the AR result in three different viewpoints. "d" is the AR result without reflection as a control experiment

Table 1. Experiment configuration

Hardware	Hardware Product Name	Software	Software Plugin
Smartphone	Huawei p20	Unity	Unity 2020.3.13f1 Personal
PC-CPU	11th Gen Intel(R)Core (TM)i5-1135G7@2.4GHz	Visual Studio Code	DX11 (embedded in Unity)
PC-GPU	GeForce MX450	ARCore SDK 4.1.7	AR foundation 4.1.7
PC-RAM	16GB		AR foundation remote 1.4.15

## 5.2. DISCUSSION

### 5.2.1. Achievements

- Current AR systems can only realize the reflection of real objects on virtual surfaces. Using the predefined plane combined with ray tracing rendering method proposed in this paper, we can effectively realize the reflection of virtual objects on real surfaces in AR systems.
- Compared to the previous research (Chen et al., 2021), the reflection could be generated accurately by ray tracing, which is a physical correct rendering method.
- Our proposed AR system can be used in landscape review scenarios, which is usually with relatively large scales models and outdoors environments.

### 5.2.2. Limitations

- The positioning of the model and the reflection plane still need to be configured manually, which is time-consuming.
- The hardware requirements are demanding and the model mesh number must be controlled; otherwise, the display rate (in FPS) will be seriously affected.

## 6. Conclusion and Future Work

The system we developed can effectively project a design model into the physical environment and can generate the accurate reflections of the model in real time by ray tracing, a physically accurate rendering method.

However, the system cannot automatically locate the reflective plane and the position of the model according to the physical environment. Thus, it requires manual operation, which is time-consuming and laborious. Future work to automate the reflective plane positioning requires the system to understand the real environment. Moreover, we need to make the rendering results run more efficiently on mobile devices and improve the display rate (in FPS), which requires further optimization of the computation algorithms.

## Acknowledgement

This research was partly supported by JSPS KAKENHI, Grant Number JP19K12681.

## References

- Akenine-Mo, T., Haines, E., & Hoffman, N. (2018). *Real-Time Rendering, Fourth Edition*. A. K. Peters, Ltd.
- Azuma, R. T. (1997). A survey of augmented reality. *Presence: teleoperators & virtual environments*, 6(4), 355-385.
- Azuma, R., Hoff, B. R., et al (1999). Making augmented reality work outdoors requires hybrid tracking. *In Augmented Reality: placing artificial objects in real scenes: proceedings of IWAR '98* (pp. 219-224).
- Ballo, P. (2018). Hardware and software for AR/VR development. *Augmented and virtual reality in libraries* (pp.45-55). Rowman & Littlefield.

- Billinghurst, M., Clark, A., & Lee, G. (2015). A survey of augmented reality. *Foundations and Trends in Human-Computer Interaction*. 8(2-3). 73-272.  
<https://doi.org/10.1561/11000000049>
- Booth, N. (1990). *Basic elements of landscape architectural design*. Waveland press.
- Broschart, D., & Zeile, P. (2015). Architecture: augmented reality in architecture and urban planning. In *digital landscape architecture, DLC 2015*. The Association for Digital Landscape Architecture Conference (DLC).
- Chen, H., Fukuda, T., & Yabuki, N. (2021). Development of an Augmented Reality System with Reflection Implementation for Landscape Design Visualization using a Planar Reflection Method in Real-Time Rendering. In *39th International Conference on Education and Research in Computer Aided Architectural Design in Europe: Towards a new, configurable architecture, eCAADe2021* (pp. 547-554). The Association for Education and Research in Computer Aided Architectural Design in Europe (eCAADe).
- Dos Santos, A. L., Lemos, D., Lindoso, J. E. F., & Teichrieb, V. (2012). Real time ray tracing for augmented reality. In *14th Symposium on Virtual and Augmented Reality* (pp. 131-140). Institute of Electrical and Electronics Engineers (IEEE).
- Giunta, L., O'Hare, J., Gopsill, J., & Dekoninck, E. (2018). A review of augmented reality research for design practice: looking to the future. In *13th biennial Norddesign Conference: Design in the era of digitalization*.
- Glassner, A. S. (1989). *An introduction to ray tracing*. Morgan Kaufmann.
- Kerr, J., & Lawson, G. (2020). Augmented reality in design education: landscape architecture studies as AR experience. *International Journal of Art & Design Education*, 39(1), 6-21.  
<https://doi.org/10.1111/jade.12227>
- Kipper, G., & Rampolla, J. (2012). *Augmented Reality: an emerging technologies guide to AR*. Elsevier.

# A NATURAL HUMAN-DRONE INTERFACE FOR BETTER SPATIAL PRESENCE EXPERIENCES

CHUN-YU LO<sup>1</sup> and JUNE-HAO HOU<sup>2</sup>

<sup>1,2</sup>*Graduate Institute of Architecture, National Yang Ming Chiao Tung University.*

<sup>1</sup>*chunyulo@arch.nycu.edu.tw, 0000-0002-7910-849X*

<sup>2</sup>*jhou@arch.nycu.edu.tw, 0000-0002-8362-7719*

**Abstract.** As many remote construction projects increase in size and complexity, being able to manage personnel schedules, delegate tasks, and check work progress can improve work efficiency and productivity. Hence, video conferencing and remote monitoring software have been attempting to pursue an immersive and intuitive experience, but with limited developments. To better achieve that, we propose a system with a natural user interface (NUI) that can offer a vivid experience, facilitating AEC personnel who is novice drone operator to interact with the Unmanned Aerial Vehicle (UAV) by voice instructions and body posture to conduct remote site surveying, monitoring, and inspections instead of physical visiting. In addition, the proposed system is capable of on-demand path planning and camera movements for various tasks and enhances the spatial experience. We integrate these techniques to develop a human-drone interface, including a VR simulator and a haptic vest system, which offer a perceivable experience of spatial presence for different purposes. Compared with other relative works, the proposed system allows users to actively control the viewing angle and movements in the remote space more intuitively. Moreover, drones can augment human vision and let users gain mobile autonomy.

**Keywords.** Spatial Presence; Natural User Interface; Human-drone Interaction; Virtual Reality; Remote Working; Body Posture Recognition; Speech Recognition; SDG 9.

## 1. Introduction

This paper proposes a system with a natural user interface (NUI) for AEC personnel who is novice drone operator to interact with the Unmanned Aerial Vehicle (UAV) by voice instructions and body posture to conduct remote site surveying, monitoring, and inspections instead of physical visiting. In addition to NUI, the proposed system is capable of on-demand path planning and camera movements for various tasks and enhances the spatial experience. Besides, we develop a system including a VR

simulator for a plausible experience of spatial presence and training purposes.

The AEC industry has been impacted by the use of UAV (Albeaino et al., 2019). Yet it is still one of the domains with low remote working capabilities (Dey et al., 2020), which implies it has a high potential for developing remote communication technologies. Researches in telepresence suggested that human beings have been eager to participate in everything in-situ (Barfield et al., 1995). And drones are capable of augmenting human vision to enormous innovative applications (Erat et al., 2018). However, non-intuitive flight operation and camera control of UAVs not only becomes cognitively demanding during long-term operations but also discourages the use of drones as an experiential platform or physical avatar by novice users (Peschel and Murphy, 2012).

Fernandez et al. (2016) pointed out that the NUI can help the operator control the drone in an intuitive way. And the use of VR and first-person perspective can increase the user's sense of presence in avatars (Macchini et al., 2021). Thus, this research aims to (1) allow UAV to respond to user's intention by spoken instructions and body posture via proposed NUI; (2) simplify viewpoint control of UAV by semi-autonomous path planning; (3) improve the experience of drone operations in VR by enhancing spatial presence.

## **2. Research Background**

### **2.1. REMOTE WORKING IN AEC INDUSTRY**

Remote construction projects exist in many regions worldwide, such as deserts, polar regions, mountains, and sparsely populated areas; construction site monitoring of those areas can be a time-consuming task requiring users' presence for observations, decisions, and actions. In addition, since the outbreak of the epidemic, most industries have been forced to work remotely. However, many jobs cannot be performed remotely, and workers are required to be present in person. Thus, solutions for remote construction will save AEC personnel from visiting the site by letting them do the monitoring remotely. Furthermore, relevant research pointed out that the AEC industry is still one of the areas with low remote working capabilities (Dey et al., 2020), which means that it has great potential for the development of remote communication technology.

To provide this domain with the ability to work remotely, Cote et al. (2011) show that installing cameras on-site and assembling them into a 360 panorama can remotely view construction sites in an immersive way, which has been considered a very positive and effective method for remote monitoring and inspection.

### **2.2. FROM VIRTUAL REALITY TO SPATIAL PRESENCE**

Many technology companies have been developing many immersive video conferencing software these years, such as Facebook Horizon Workroom, Spatial., which implies that human beings have been eager to experience space immersively. However, compared with the technique mentioned above, the AEC industry needs a remote communication technology that allows users to shuttle in the remote space freely.



Spatial presence is usually defined as the feeling of “being there”. The term “there” refers to a virtual location or a real remote location. Therefore, spatial presence includes the user's ability to experience a sense of presence in any environment where they are transported. According to the results of user testing, the remote environment unexpectedly has a higher sense of spatial presence or “hyper-presence” at certain times compared with the real environment (Khenak et al., 2020).

Moreover, drones can augment human vision and let operators gain mobile autonomy, and the use of VR and first-person perspective can increase the user's sense of presence in avatars (Macchini et al., 2021). Therefore, combining drones with VR can provide users with a safe and fascinating flight environment for training to achieve spatial presence.

### 2.3. DRONE USER INTERFACE

Natural User Interfaces (NUIs) intend to make use of innate human features, such as speech, gesture, posture, and vision to interact with technology. Similarly, it is crucial for humans to interact and command drones in natural and efficient ways in Human-Drone Interaction (HDI) frameworks (Fernandez et al., 2016). Yam-Viramontes and Mercado-Ravell (2020) integrated body gestures into the control system of the drone. In the study, the proposed strategy is validated with different human users through a standard User Experience Questionnaire (UEQ), showing good results in usability and user experience. It proved that using NUI to control drones has good results and positive feedback.

## 3. System Design and Implementation

### 3.1. SYSTEM STRUCTURE

This research developed a drone simulator that has the configuration of an actual UAV to test the feasibility in the future. Furthermore, we proposed a NUI with body posture and voice control for the drone according to the survey of the AEC industry's communication dialect.

Based on these two control principles, this research proposed three systems to control UAVs: (1) Responsive Control System (RCS), (2) Body Posture Control System (BPCS), and (3) Voice Control System (VCS). With these systems, the operator can receive the current flight status information from the simulator through vision and haptic feedback during the flight. Simultaneously, the recognition system will capture the body's response, transmit it to the workstation for data processing, and send the results to the simulator. Moreover, the operator can get the vision of the simulation from the screen or a head-mounted display (HMD), switch various scenes, perspectives, and give instructions through the NUI.

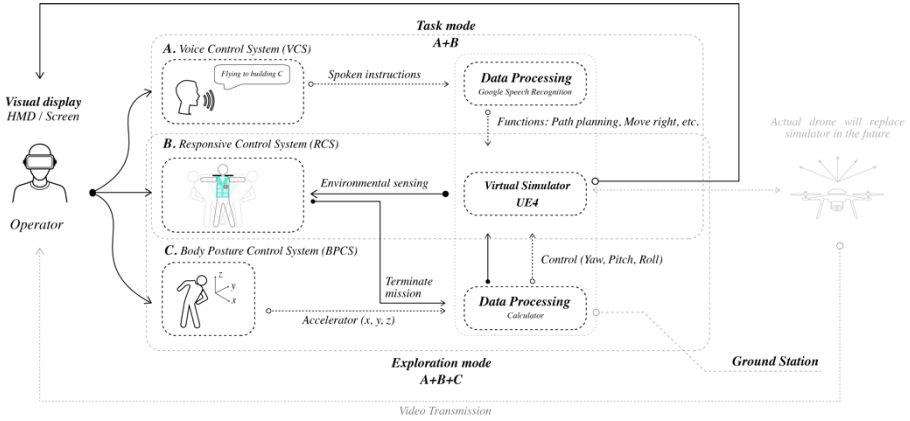


Figure 1. System structure

### 3.2. SYSTEM ANALYSIS AND DESIGN

At present, the applications of UAVs in the AEC industry are primarily used in monitoring, photography, scanning, and urban observation (Albeaino et al., 2019). According to the purposes of the above tasks, this research organizes these control systems into two UAV control modes: (1) Task mode and (2) Exploration mode. In addition, it is expected that the proposed NUI can automatically recognize the operator's intention to switch the control mode to achieve a natural use.

- (1) The Task mode uses VCS as the primary input of the UAV, and the RCS serves as the emergency control. When performing simple tasks and flights such as photography and scanning, the drone only needs a simple flight path and functions. Hence, the operators can reduce their burden with the Voice control system and free their hands for other works or tasks. Moreover, if the body receives an environmental warning during the flight, the RCS will be activated and decide whether to terminate the ongoing mission based on the operator's behavior.
- (2) The Exploration mode uses BPCS as the primary input of the UAV, VCS and RCS serve as the function inputs and the emergency control. When performing remote city exploration or surveying, the drone needs to be fully controlled by the operators to follow their intention. Besides, the operators usually imagine themselves as the avatar of a drone during the flight. Thus, the operators can fully control the drone's flight through the BPCS in this control mode. In addition, VCS in this mode is used for taking off, landing, perspective switching, and other functions. Furthermore, if the body receives an environmental warning during the flight, the RCS will be activated and decide whether to terminate the ongoing mission based on the operator's behavior.

To sum up, the VCS is mainly used to control UAV's take-off, landing, perspective-switching, path planning, image recording, etc. The BPCS is mainly used for controlling the flight (Roll, Yaw, Pitch). In addition, the RCS serves as emergency control; when the virtual drone perceives that the environment is facing an emergency,

the body reaction of the operator will be captured by the RCS to achieve obstacle avoidance.

For safety reasons, the RCS is designed as the highest-level system, covering any commands under any circumstances. Similarly, the operator can give instructions through voice, and if the BPCS is activated simultaneously, the previously spoken instructions will be overwritten.

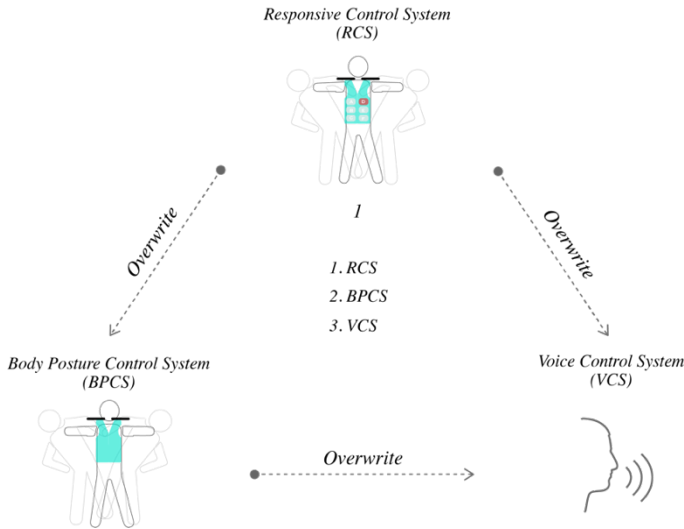


Figure 2. Hierarchy of control systems

### 3.3. THE SIMULATOR

This research uses Unreal Engine to develop a drone simulator in which the configurations of the actual drone are included, such as the flight status, distance, speed, and altitude. The virtual drone can detect the distances between itself and the surroundings in real-time and export the results to the external program for data processing. The ambient light and climate in the simulator can be changed accordingly.

In the simulator, users can conduct perspective-switching and path planning, as well as change scenes. It also simulates various environmental conditions. In order to show the typical scenes of remote construction, this article chooses “City Park” as the simulation field in this phase.



Figure 3. (a) Distance detection system (b) Third-person perspective (c) Path planning

### 3.4. NUI - BODY CONTROL AND SENSING

#### *3.4.1. Responsive Control System (RCS) - Vest As a Passive Controller*

When facing an emergency, humans will respond physically based on experience to avoid disasters (Dunsmoor and Murphy, 2015). Similarly, the operator will also perform body dodge actions when experiencing a dangerous state. Based on these phenomena, this research develops a Responsive Control System (RCS) that will be activated when the situation is severe to prevent accidents. Besides, it is difficult for the operator to clearly understand the actual distance between the aircraft and the surrounding obstacles through the visual display in dynamic flight.

Research has shown that haptic feedback can enhance environmental perception and effectively trigger the operator's emotions. In addition, jackets and vests can give users a multi-part perception of the body; the pressure's position, intensity, frequency, and temperature can correspond to various feelings defined in the emotional list of the social psychology model (Arafsha et al., 2012). Therefore, this paper applies a developed haptic vest, which can control the size of the airbag according to external information, and integrates it into the drone simulator to enhance the operator's understanding of the flight environment.

The haptic vest we used can activate specific airbags through inflating for pressuring on the pilot's body to achieve the function of direction warning when the environment is severe. The obstacle avoidance system in the simulator can instantly detect the environment around the aircraft; the detection range is set as five meters as a default, which can also be adjusted according to different aircraft types, velocities, and detection principles.

During the process, the RCS sends the detected distance value to the external program file for data processing. The processed results are sent to the specific motors on the vest through the Message Queuing Telemetry Transport (MQTT) protocol to inflate the specific airbags. Finally, the current airbag inflation rate has been optimized to 0.2 seconds to minimize the latency.

In order to quantify the swing of current body posture, this research sets a micro:bit on the user's shoulder, which is embedded in the vest. In addition, the micro:bit serves as the accelerometer for capturing body posture and the device for wireless communication. Thus, if the RCS detects a drastic change of the operator's body posture when the airbag is activated, it will immediately pause the drone's flight.

#### *3.4.2. Body Posture Control System (BPCS) - Vest As an Active Controller*

The vest acts as an active controller in the BPCS. It obtains the current data of body posture in real-time through the micro:bit set on the shoulder; sends the data to the external program file by setting a specific radio channel for data processing. The processed results are sent to the simulator for controlling the drone. After validation, the values of each axis captured by the accelerometer can be effectively mapped to the drone's control parameters (Roll, Yaw, Pitch) to change its flight path to achieve the BPCS.

It has also been tested that the captured value can effectively correspond to the body posture. Moreover, we have successfully optimized the update rate of the communication system to 30 data per second to minimize the latency.



*Figure 4. The virtual drone follows the operator's body posture in BPCS*

### 3.5. NUI – VOICE CONTROL SYSTEM (VCS)

#### 3.5.1. *Speech Recognition System*

In order to allow AEC personnel to carry out simple tasks and free up their hands easily, this research developed a Voice Control System (VCS) for drones. This system currently uses Google Speech Recognition as a standalone subsystem for Speech To Text. Besides, the operator can control the drone by simply speaking the requirements, such as flight parameters (Roll, Yaw, Pitch) adjustments, path planning, perspective mode-switching, screenshots; the system will automatically capture the keywords from the spoken instructions and execute them.

#### 3.5.2. *Speech Command Dictionary*

The Speech Command Dictionary refers to the professional dialect in the AEC domain and on-site communication terms. Besides, this research also studies how the designers communicate with remote workers in simple, clear, and understandable terms (Christenson et al., 2013). At this phase, three types of functions have been developed, including basic controls, functions, and various types of path planning. The VCS can recognize sentences containing keywords; when the keywords are detected, the VCS will execute the corresponding commands.

During the whole process, the VCS can continuously recognize the spoken instructions during the flight and decide whether to execute them based on the weights of the instructions. For example, the system will overwrite the previous command if the drone receives a stop command while flying to the target; if the drone receives a screenshot command during the flight, the system will continue to fly and take a screenshot. The following examples illustrate several commands and actual functions.

- Take off = Move up 10 meters above the ground to stand by.
- Move right = Adjust the degree of the Yaw axis of the drone.
- Circular movement on building B and screenshot = Use the current distance between the drone and the target building as the radius to circle around and take a

screenshot.

- Switch to FPV = Switch the screen of the simulator to first-person perspective.

Basic control	Function	Path planning
Take off	Screenshot	Linear movement to building A
Land	Start recording	Circular movement on building B
Turn right	Stop recording	SemiCircular movement on building C
Turn left	Switch perspective	Grid movement on building D
Move forward		
Move backward		
Stop		

Figure 5. Speech command dictionary

### 3.5.3. Path Planning and Photography

By designing a Speech command dictionary, the operator can use voice to give the drone flight commands. This paper refers to the current UAV software and proposes the following five types of movements in the preliminary stage.

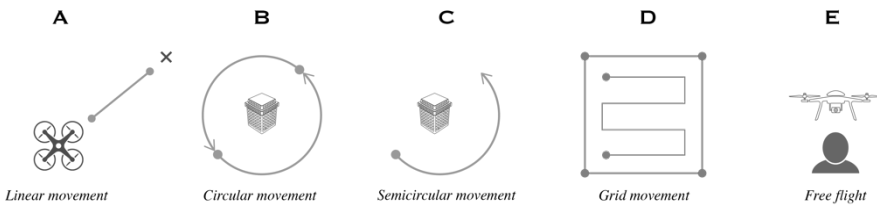


Figure 6. Types of path planning



Figure 7. (a) Linear movement (b) Circular movement

## 4. Tests and Preliminary Feedback

This research plans a set of application scenarios for the AEC domain, which provides architects, planners, and practitioners abilities to perform remote operations with

intuitive body posture and voice to obtain a better spatial presence and experience. First of all, this research conducts a user test involving five professionals who work in the AEC field to verify the feasibility of using the drone NUI to assist remote tasks. Secondly, we described the goal of the test and show them how to use the NUI. During the process, the words and body postures they used are recorded to verify the system's availability deficiency, and the user's response and feedback will also become the focuses of the next stage.

The test requires users to fly around the building to take pictures. During the process, they need to set the origin, take off, approach the building, switch perspective, stop, screenshot, fly back to the origin, and land.

Finally, the users provided a lot of feedback and expectations, including that (1) it is expected operators in different regions can control the same UAV at the same time in the future; (2) it is expected the system can automatically perform tasks by observing the dialogue between users rather than by instructions; (3) it is expected the system can record the last flight path and perspective to execute tasks repeatedly; (4) it is expected the practitioners at the remote site can know who the drone operator is to increase interactivity; (5) it is expected this system can bring users to visit any indoor space, making the virtual tour experience more immersive.

## **5. Conclusion and Future Work**

In summary, this research has developed a drone simulator and proposed a NUI strategy that combines RPS, BPCS, and VCS to allow operators to interact with the drone intuitively. Moreover, compared with the previous video conferencing and remote monitoring software, the proposed system allows users to actively control the viewing angle, which can help spatial designers improve the perception of remote space. Furthermore, users can perform multi-modal interaction with UAVs by switching the proposed NUIs to meet their application requirements. The use of NUI for HDI (Human-drone interaction), especially for tasks in the AEC industry, enhances spatial presence and expands its scope of application; We expect the system can recognize requirements automatically in the next step to achieve a more natural experience.

However, the current simulator still has room for improvements, including the efficiency of speech recognition and the limitation of the Speech command dictionary. Therefore, this research will apply natural language processing (NLP) to improve the VCS and apply the proposed system to the physical drone for validation in the next stage.

Lastly, the application of this research can be used in AEC industries, building and bridge inspections, progress monitoring, urban planning, architectural photography, and surveying, as well as taking guests and customers to any spaces for virtual tours and become a new type of telepresence robots in the future. In the future, if our simulated environment can precisely match the physical drone flight, the proposed NUI will be of great benefit to enhance the interactivity of the drone control system.

## Acknowledgements

This research is supported by the Ministry of Science and Technology, and the grant number is MOST 110-2634-F-009-018; and thank the Architectural Informatics Lab (GIAAIL), National Yang Ming Chiao Tung University for technical support and assistance.

## References

- Albeaino, G., Gheisari, M., & Franz, B. W. (2019). A systematic review of unmanned aerial vehicle application areas and technologies in the AEC domain. *J. Inf. Technol. Constr.*, 24(Jul), 381-405.
- Arafsha, F., Alam, K. M., & El Saddik, A. (2012). EmoJacket: Consumer centric wearable affective jacket to enhance emotional immersion. Paper presented at *the 2012 international conference on innovations in information technology (IIT)*.
- Barfield, W., Zeltzer, D., Sheridan, T., & Slater, M. (1995). Presence and performance within virtual environments. *Virtual environments and advanced interface design*, 473-513.
- Christenson, M., Ahmed, B. A., & Nassar, K. (2013). A verbal communication game for architecture, engineering and construction students. Paper presented at *the American Society for Engineering Education North Midwest Section Conference*.
- Côté, S., Latulippe, M., bilodeau, v., Poirier, S., gervais, r., & Barette, Y. (2011). Immersive remote construction site monitoring using live augmented panoramas.
- Dey, M., Frazis, H., Loewenstein, M. A., & Sun, H. (2020). Ability to work from home. *Monthly Labor Review*, 1-19.
- Dunsmoor, J. E., & Murphy, G. L. (2015). Categories, concepts, and conditioning: how humans generalize fear. *Trends in cognitive sciences*, 19(2), 73-77.
- Erat, O., Isop, W. A., Kalkofen, D., & Schmalstieg, D. (2018). Drone-augmented human vision: Exocentric control for drones exploring hidden areas. *IEEE transactions on visualization and computer graphics*, 24(4), 1437-1446.
- Fernandez, R. A. S., Sanchez-Lopez, J. L., Sampedro, C., Bavle, H., Molina, M., & Campoy, P. (2016). Natural user interfaces for human-drone multi-modal interaction. Paper presented at the *2016 International Conference on Unmanned Aircraft Systems (ICUAS)*.
- Go, Y.-G., Kang, H.-S., Lee, J.-W., Yu, M.-S., & Choi, S.-M. (2021). Multi-User Drone Flight Training in Mixed Reality. *Electronics*, 10(20), 2521.
- Khenak, N., Vézien, J., Théry, D., & Bourdot, P. (2020). Spatial Presence in Real and Remote Immersive Environments and the Effect of Multisensory Stimulation. *Presence*, 27(3), 287-308.
- Macchini, M., Lortkipanidze, M., Schiano, F., & Floreano, D. (2021). The Impact of Virtual Reality and Viewpoints in Body Motion Based Drone Teleoperation. Paper presented at the *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*.
- Peschel, J. M., & Murphy, R. R. (2012). On the human-machine interaction of unmanned aerial system mission specialists. *IEEE Transactions on Human-Machine Systems*, 43(1), 53-62.
- Yam-Viramontes, B. A., & Mercado-Ravell, D. (2020). Implementation of a Natural User Interface to Command a Drone. Paper presented at the *2020 International Conference on Unmanned Aircraft Systems (ICUAS)*.



# **Generative, Algorithmic & Evolutionary Design**



# CALIBRATING A FORMFINDING ALGORITHM FOR SIMULATION OF TENSIONED KNITTED TEXTILE ARCHITECTURAL MODELS

FARZANEH OGHAZIAN<sup>1</sup>, NATHAN BROWN<sup>2</sup> and FELECIA DAVIS<sup>3</sup>

<sup>1,2,3</sup>*The Pennsylvania State University.*

<sup>1</sup>*fxo45@psu.edu, 0000-0002-1728-5191*

<sup>2</sup>*ncb5048@psu.edu, 0000-0003-1538-9787*

<sup>3</sup>*fud12@psu.edu, 0000-0001-9417-0518*

**Abstract.** This paper presents an optimization-based calibration process for tuning a digital formfinding algorithm used with knitted textile materials in architectural tension structures. 3D scanning and computational optimization are employed to accurately approximate a physical model in a digital workflow that can be used to establish model settings for future exploration within a knit geometric typology. Several aspects of the process are investigated, including different optimization algorithms and various approaches to data extraction. The goal is to determine the appropriate optimization method and data extraction, as well as automate the process of adjusting formfinding settings related to the length of the meshes associated with the knitted textile behavior. The calibration process comprises three steps: extract data from a 3D scanned model; determine the bounds of formfinding settings; and define optimization variables, constraints, and objectives to run the optimization process. Knitted textiles made of natural yarns are organic materials and when used at the industrial level can satisfy DSG 9 factors to promote sustainable industrialization and foster innovation in building construction through developing sustainable architectural systems. The main contributions of this paper are calibrated digital models of knitted materials and a comparison of the most effective algorithms and model settings, which are a starting point to apply this process to a wider range of knit geometries. These models enhance the implementation and further development of novel architectural knitted systems.

**Keywords.** Tensioned Knitted Textiles; Computational Design; Formfinding; Calibrating; Optimization, SDG 9.

While knitted fabrics are not common in architecture, recent improvements in CNC knitting and computational design have increased their viability, leading to increasing interest (Tamke et al., 2020; Thomsen et al., 2016; Ahlquist & Menges, 2013; J. E. Sabin, 2013). Knitted textiles offer many potentials, which makes them a perfect material for developing more sustainable architectural systems and building

construction. This research positions itself in the Sustainable Development Goal 9 [SDG 9] sector which aims to build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation ("HLFP Thematic Review of SDG-9," 2017). Knitted textiles made of organic yarns enable architects to apply these materials for making sustainable structures. Additionally, knitted textiles because of their specific structures allow for integration of different materials such as conductive yarns into their structures. When used as architectural structures these materials can collect energy and be used as a source of electricity. Much more can be added to the list that support the need for more in-depth investigation of these materials for architectural application. Successful design and fabrication of knitted textiles in architectures requires developing a digital model parallel to physical experiments that represents the behavior of the material accurately.

However, developing digital models for architectural knitted textiles is often a challenge, especially in early design. While tools such as Kangaroo 1 and 2, which use Dynamic Relaxation and the Projective Constraint method respectively, are immensely powerful for explorative formfinding of flexible materials, they require the user to specify lengths and stiffnesses or related model properties to get an accurate representation of the tensioned shape. Especially for knitted structures, a specimen might have substantially different properties in different directions or regions of the textile, making these settings tedious to model correctly. Making digital models thus requires an iterative process of experimentation with real materials, learning from the physical modeling, and translating material logic to simulation settings in the digital environment. Yet converting material behavior into numerical data has challenges that stem from the range of simulation inputs required to accurately represent a tensioned shape. Such manipulations include translating non-linearity in the behavior of the knitted textile in different directions depending on yarn types, patterns, and knitting setting used for making knitted structures, as well as boundary conditions, and stabilizing applied forces.

In connection to ongoing research for developing formfinding methods of knitted tensioned structures (Oghazian, Farrokhsiar, & Davis, 2021), the authors of this paper implement and test optimization methods as tools for tuning a digital formfinding algorithm to aid in the design of architectural knitted tension structures.

## **1. State-of-the-Art**

Formfinding and simulation of architectural tension structures made of real materials with complex characteristics are challenging. Although some researchers emphasize that the formfinding process is a geometrical problem and material independent (Dutta & Ghosh, 2019), others argue that in working with specialized materials, the characteristics of the material along with the geometry inform particular behaviors (Oxman & Rosenberg, 2007). This research positions itself in the second category, where the geometry and material characteristics are both informative in formfinding process. However, for materials such as knitted textiles used in architectural tension structures, working with the material properties requires an iterative learning process. This learning process is a feedback loop that synthesizes information from physical models and manipulates digital formfinding algorithms.

In design and fabrication with knitted textiles, three scales can be recognized in

their structure: micro-scale (one stitch or loop), mesoscale (combination of stitches in patterns), and macro-scale (overall form) (Oghazian & Vazquez, 2021). Knitted textiles possess some uncertainty in their behavior, derived from how they were formed at these varying scales. No matter the yarn type, different knit structures affect the elasticity of the textile. Therefore, when combined with the complex irregular architectural forms not common in textile design, predicting and simulating the exact shape and behavior of the knitted material is challenging. Many research studies in the literature tried to remove the extra elasticity added because of the structure of the textile by using yarns with the least elasticity, such as Dyneema, and patterns that add minimum elasticity to the textile (Tamke et al., 2020). In order to limit stretchability, techniques such as in-laying yarns are incorporated during the manufacturing process (Pal, Chan, Tan, Chia, & Tracy, 2020).

Calibrating digital models based on the results of a physical model is not new in architecture. However, in many studies, the calibration process is not elaborated. The simulation result is not compared with the actual physical models, especially with materials that possess some uncertainty in their behaviors, such as knitted textiles. In a study by Cuvilliers, Yang, Coar, & Mueller (2018), two common algorithms of Kangaroo1 and 2 are compared regarding the reliability, speed, and accuracy of the formfinding process by defining numerical calculation from physical model measurements and comparing them with the form found models. Their main argument is that it is not clear how the available algorithms can be calibrated to get meaningful results in physical units. Specifically for knitted textiles, in calibrating the simulation process for a piece of knitted textile by Schmeck & Gengnagel (2016), the authors explored the cyclic manipulation process to make their algorithm correspond to certain characteristics of the material. However, complete accordance has yet to be achieved.

While simulating and calibrating a digital model of a square shape of knitted textile that is deformed to a conical shape by applying force in the middle of the textile, Baranovskaya, Tamke, & Thomsen (2020) used a genetic algorithm through Galapagos to calibrate the digital model. The meshes represent graded knitted textiles with different stiffness characteristics manipulated to obtain a digital model close to the physical 3D scanned shape of the knitted textile. In connection to this last study, in our research, we propose a procedure and examine different optimization tools as a method of calibrating digital formfinding algorithms for conical 3D knitted tension structures. The goal is to speed up the calibration process by systematically limiting the number of data points and selecting the most reliable algorithms.

## **2. Introducing Workflow, Case Study, Challenges, and Research Goals**

In this research, 3D scanning captures the overall shape of a physical knitted textile structure to compare it with the digital models. The selected shape is a conical form with 48 stitches in the course and wale direction in the conical part. The cone is selected because it is common in tensile architectural structures. Seamless knitting allows to knit tube and cones easily. While simple, there are challenges in simulating the exact behavior of the conical shapes to avoid a bottle-neck effect around the tip of the cone and wrinkles over the stretched surface. Figure 1 shows some of projects that used conical knitted textile shapes for architecture. The formfinding procedure implemented by Oghazian, Farrokhsiar, & Davis (2021) is used as a starting point. The formfinding

uses Kangaroo 2 and a simplified quad-based mesh representing the knitted textile structure.



Figure 1: Case studies of projects made of conical knitted textiles. Left: Hybrid Tower (Thomsen et al., 2015). Right: Fabricating Networks, Flower Antenna Sculpture (Davis, 2021)

Usually in the formfinding process the mesh settings should be modified to obtain a shape that corresponds to the physical model. A more accurate model will have smaller geometric differences between the digital and physical artifacts as measured by the method in Section 3.1. Accurate models are critical because designers use them for feedback during rapid iteration, and a final product that looks substantially different is undesirable. The tedious task of calibration requires looking at the model repeatedly and determining the length and stiffness of the mesh edges in formfinding components. In this paper, we use optimization methods to automate this calibration process. Pre-requirements for the calibration process include first overlapping the formfound and 3D scanned mesh, considering the upper and lower boundaries as the fixed points (Figure 2). The distance between two meshes is then minimized through calibration process. The cumulative distance is considered as an objective function for optimization, while the length for seven mesh edge categories are the variables. These mesh categories will be explained in more detail in section 4.2. The same process can be used for tuning formfinding simulation of other soft and knitted textile materials and forms. However, the mesh categories might differ based on the overall form used.

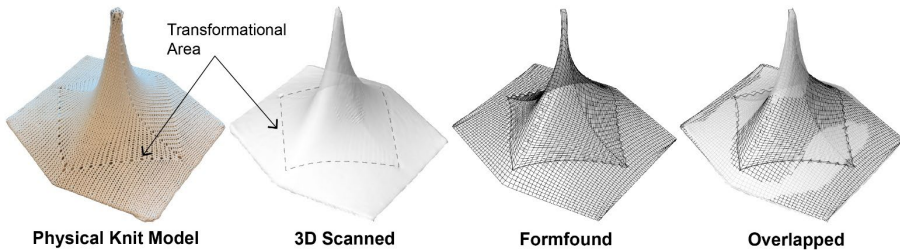


Figure 2: Physical Knit Model, 3D scanned, Formfound Models, and 3D scanned/Formfound Overlapped meshes

In the conical shape selected in this study, two critical parts can be recognized: the cone and the planar surface. After post tensioning the connection between these two parts is more consistent. Since the stitch effect is different in Transformational Area, as it is shown in Figure 2 over the Physical Knit Model, we can use it as a qualitative criterion to determine the performance of the optimization algorithms.

### 3. Process

The main questions we address in this research are: what are the best data extraction methods from the 3D scanned model that represent the overall form and behavior of the tensioned knitted textiles? What are the most appropriate optimization algorithms that can assist architect designer for tuning formfinding algorithms? The calibration process includes a series of steps. The first step is to extract data from the 3D scanned model that is being used as a test case, and the second step is to determine the range for springs length in the formfinding algorithms based on the results of the physical modeling. Once the simulation problem has been established, the third step is to set the optimization constraints and repeatedly run different combinations of data extraction and optimization settings to determine the most effective method.

#### 3.1. DATA EXTRACTION

A 3D Systems SENSE 2 3D Scanner was first used to scan the knitted samples. The output of the 3D scanning process is a 3D mesh. The formfound model also includes many points associated to the corners of the mesh faces/stitches. To determine the distance between two meshes, the points from formfound model can be projected vertically or perpendicularly to the 3D scanned mesh (Figure 3).

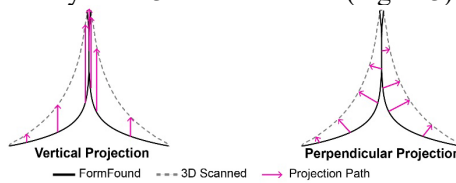


Figure 3: Projection methods

While both were attempted, the perpendicular projection method is used because it produces better results regarding the curvature fit between the models. Vertical projection was also problematic for the points near the tip of the cone, where the projection lines were almost tangent to the mesh, or not touching the other mesh at all.

Next, the set of points for distance measurement had to be selected. Three individual approaches are explored: Random Points, Section Points, Transformation Area Points (Figure 4A). Except for the Random Points the other methods are selected because of the specific geometry of the overall model. Combinatorial approaches are also explored by combining the individual strategies to investigate the combined effect on enhancing the performance of the optimization process (Figure 4B).

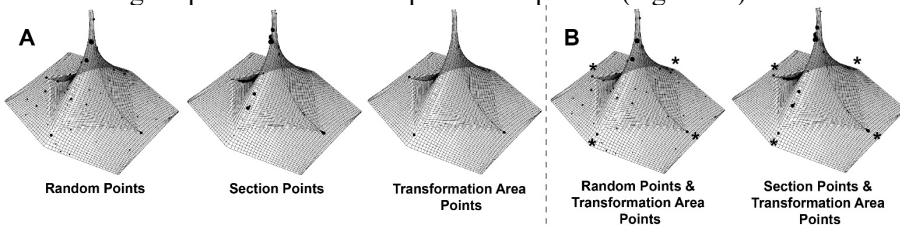


Figure 4: Data extraction approaches A: Individual Approaches. B: Combinatorial Approaches (Weighted /not Weighted. The weighted points are labelled with \*)

### 3.2. MESH EDGES AND MESH LENGTH FACTOR BOUNDS

Once the points are selected for comparison, the formfinding process was implemented using Kangaroo 2. The main steps for any formfinding process are 1) defining the initial mesh, 2) determining the boundary conditions and external forces, 3) converting the mesh edges to the springs considering the mesh length and strength, and 4) running the formfinding solver. The central activity of this paper is Step 3, where we convert mesh edges to springs and input characteristics of the knitted textiles. Studying physical conical models used in this study show, while characteristics and size of the stitches are similar before applying the tension, their behaviors are varied at different sections of the conical tensioned structures. As illustrated in Figure 5A, the mesh lines are divided into seven categories to have better control over the shape change of different parts of the model.

Another critical element is the length criteria. This element determines the length change during formfinding. Here, initial mesh lengths are multiplied by a length factor to determine the length change. These length factors are used as variables in the optimization process. To adequately reflect the size change of the mesh edges, the bounds of the length change associated with each of these categories are limited to reliable bounds as presented in Figure 5A.

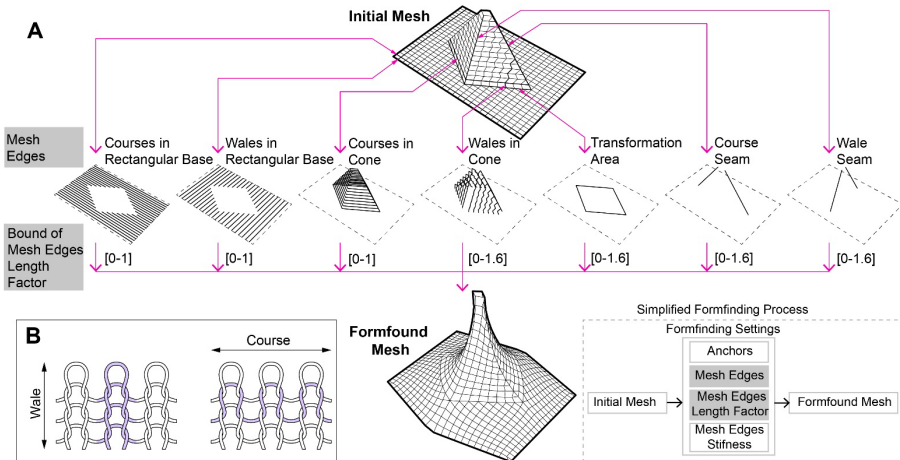


Figure 5: A: Seven categories of mesh edges and associated length factors, B: Wale and Course

In general, two main directions can be recognized in knitted textiles structures, and these are course and wale (Figure 5B). In the physical model used in this study, the initial knitted model is attached to the boundary frame, keeping the overall length and width around the rectangular head as it was knitted and without stretching. Therefore, the elongation only happens in the wale direction, which is the direction of the exterior force. Consequently, the stitches in course direction mainly shrink.

### 3.3. OPTIMIZATION

For optimization, the plugins Radical and Opossum are selected. Radical is a tool in Design Space Exploration (DSE) that incorporates algorithms for constrained,



gradient-based optimization for numerical and geometric design variables (Brown, Jusiega, & Mueller, 2020). Opossum is a black-box optimization tool applicable for time-intensive architectural problems (Wortmann, 2017).

After an initial test period, we selected nine promising algorithms that performed better on minimizing the differences between the simulated and physical structure. There are six algorithms from Radical DSE including: SBPLX, GN\_ORIG\_DIRECT, GN\_ORIG\_DIRECT\_L, ISRES, LN\_BOBYQA, LN\_COBYLA. There are three from Opossum including: RBF, CMAES, and CMAES\_Random. All algorithms in Radical are from the NLOpt library (G. Johnson, 2020), which can be consulted for more information about each one. The algorithms from Opossum are from RBFOpt, an open-source library for black-box optimization (Costa & Nannicini, 2018). The starting point for all the algorithms is setting the length factor to zero, which can be the start of a generalized process. Figure 6 graphically illustrates the starting point.

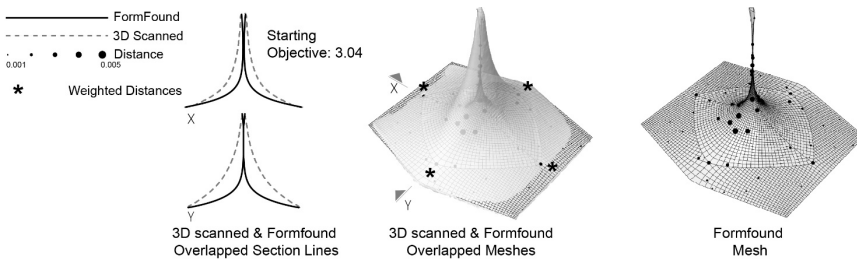


Figure 6: Starting point in optimization process where all the variables are set to zero

### 3.4. RESULTS OF OPTIMIZATION PROCESS

The results of the optimization processes for different data extraction methods and algorithms are provided in this section. Figure 7-left illustrates the performance of different algorithms for the three individual data extraction approaches. The Random Points and Section Points techniques both give better results considering the shape of section lines. But the Transformation Area Points technique did not yield curves in the formfound model that completely followed the 3D scanned model shape. GN-ORIG-DIRECT-L Radical is the best solver for both approaches. In the Transformational Area Points method, because the distance between the four points in this approach was less than the other approaches, the distance is multiplied by 25. Transformational area lines in the formfound model completely overlap the 3D scanned model. However, none of the algorithms yield section lines that are close to the 3D scanned model. Based on these observations we decided to combine the four-point Transformational Area Points strategy with the Random Points and Section Points approaches as a combinatorial approach (Figure 7 Right). In this approach we added the four Transformational Area Points with and without weight.

In general, adding weight to the four critical points of the Transformational Area in combinatorial approaches does not improve the performance of the algorithms. Almost all the algorithms will give better section lines in unweighted strategies. Performance of the algorithms regarding the Transformational Area shape is a bit better in weighted strategies, but it is not significant.

LN\_BOBYQA and LN\_COBYLA usually stopped at the earlier stages of the optimization process for either individual or combinatorial approaches and not giving good results regarding the objective of the project. GN\_ORIG\_DIRECT\_L, which is a global optimization algorithm that works through systematic division of the search domain into smaller hyperrectangles, is the most reliable algorithm for different types of the data extraction such as Random, Section Points, and Random Points & unweighted Transformational Area Points. 1077 seconds is the minimum, and 3750 s is the maximum time for optimization of the models that ran for all 400 iterations. However, some algorithms were able to obtain their best results faster and within less than 100 iterations such as SBPLX.

#### **4. Conclusion and Contribution**

This study explores the potentials of using optimization strategies for calibrating the digital model out of the formfinding process and automating the process of adjusting formfinding settings in simulating architectural knitted textiles. The results of our study, a contribution of this paper, show that among different individual and combinatorial data points, we found the Random Points and Random Points & Not Weighted Transformational Area Points technique to provide good data points for a calibration process. The number of data points obtained from the 3D scanned models are many and if all of them are used to minimize the distance between the 3D scanned mesh and formfound mesh, it increases the computational process during the optimization process. Therefore, we limited the data to the minimum critical points to obtain the desired results.

The performance of different optimization algorithms in Grasshopper plug-ins was then investigated for tuning a simulation process of conical knitted tensioned structures. Another contribution is that GN-ORIG-DIRECT-L from the Radical plugin was one of the best algorithms for this research problem. To speed up the calibration process during the design process, it was important to understand which kinds of data, variables, and optimization tools should be selected at the first step.

While this research uses small conical models as a case study, the whole process of calibrating the formfound model can be applied to other projects and other tension structure shapes that use similar materials. Researchers can benefit from the detailed calibration process introduced here that include 3D scanning the model and extracting the essential data, Determining the bounds of formfinding settings, and setting the optimization process and minimizing the distance between the target surface and initial surface. The procedure can be adopted as a method to automatically determine the length factor settings for the mesh edges during the formfinding and design process.

#### **Acknowledgement**

The authors are grateful for the support of ICDS Seed Grant at Penn State University 20212022.

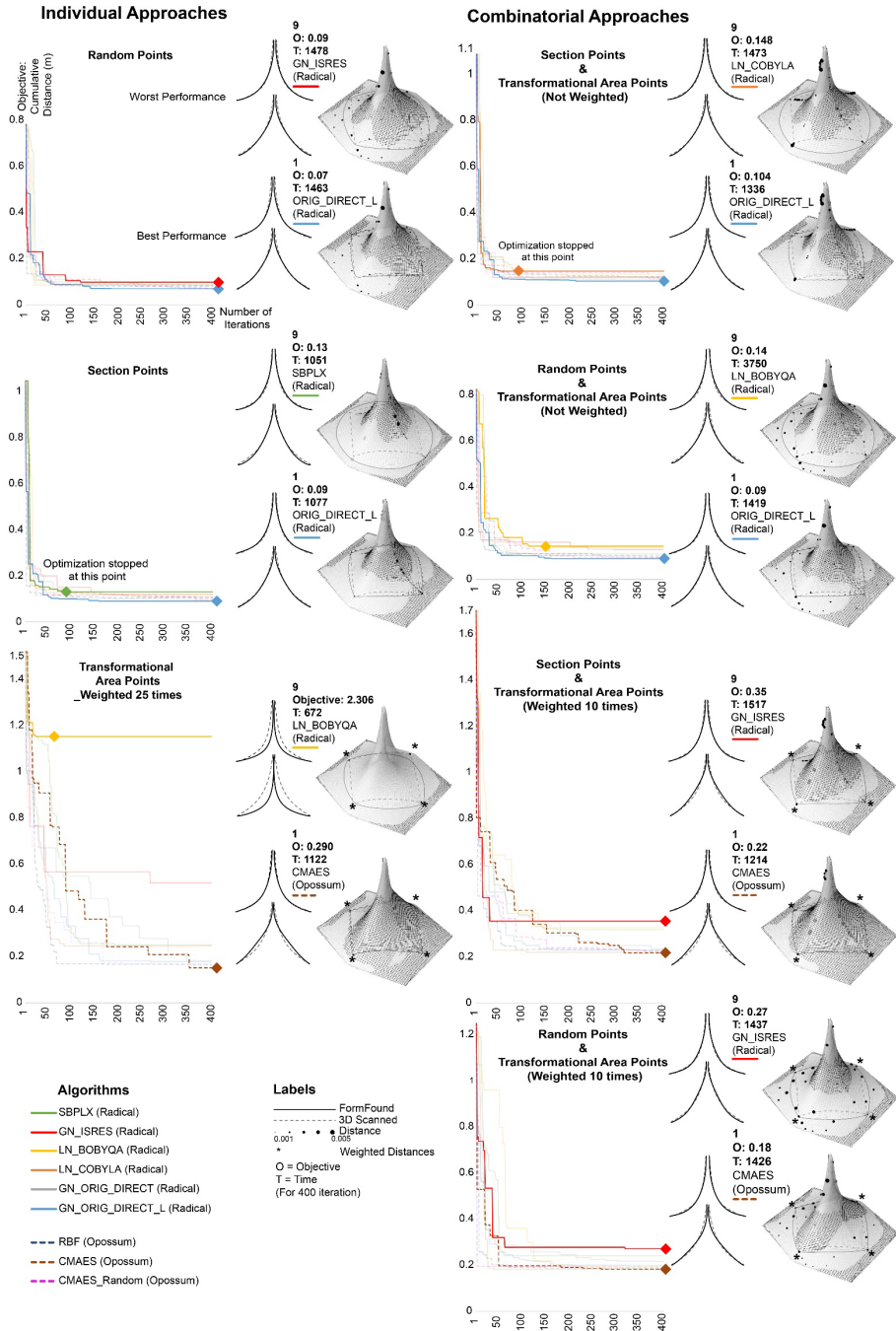


Figure 7: Comparison of individual and combinatorial data extraction strategies for different Radical and Opossum algorithms (Left: Individual approaches. Right: Combinatorial approaches)

## References

- Ahlquist, S., & Menges, A. (2013). Frameworks for Computational Design of Textile Micro-Architectures and Material Behavior in Forming Complex Force-Active Structures. *ACADIA Conference*.
- Baranovskaya, Y. S., Tamke, M., & Thomsen, M. R. (2020). Simulation and Calibration of Graded Knitted Membranes. *ACADIA Conference*.
- Brown, N. C., Jusiega, V., & Mueller, C. T. (2020). Implementing data-driven parametric building design with a flexible toolbox. *Automation in Construction*, 118, 16. doi:10.1016/j.autcon.2020.103252.
- Costa, A., & Nannicini, G. (2018). RBFOpt: an open-source library for black-box optimization with costly function evaluations. *Mathematical Programming Computation*, 10(4), 597-629.
- Cuvilliers, P., Yang, J. R., Coar, L., & Mueller, C. (2018). A comparison of two algorithms for the simulation of bending-active structures. *International Journal of Space Structures*, 33(2), 73-85.
- Davis, F. (2021). *Fabricating Networks: Transmissions and Receptions from Pittsburgh's Hill District, Flower Antenna*. Available from: <https://www.moma.org/audio/playlist/312/4039>. Retrieved [04/05/2021].
- Dutta, S., & Ghosh, S. (2019). Analysis and Design of Tensile Membrane Structures: Challenges and Recommendations. *Practice Periodical on Structural Design and Construction*, 24(3).
- Johnson, S. (2020). *The NLOpt nonlinear-optimization package*. Retrieved from <http://github.com/stevengi/nlopt>.
- HLFP Thematic Review of SDG-9. (2017). Retrieved from <https://sustainabledevelopment.un.org>.
- Oghazian, F., Farrokhsiar, P., & Davis, F. (2021). A simulation process for implementation of knitted textiles in developing architectural tension structures. *IASS 2020/21*.
- Oghazian, F., & Vazquez, E. (2021). A Multi-Scale Workflow for Designing with New Materials in Architecture: Case Studies Across Materials and Scales. *CAADRIA 26*.
- Oxman, N., & Rosenberg, J. L. (2007). Material-based Design Computation an Inquiry into Digital Simulation of Physical Material Properties as Design Generators. *IJAC journal*, 5(1), 25-44.
- Pal, A., Chan, W. L., Tan, Y. Y., Chia, P. Z., & Tracy, K. J. (2020). Knit Concrete Formwork. *CAADRIA Conference*.
- Sabin, J. E. (2013). My thread Pavilion: Generative Fabrication in Knitting Processes. *ACADIA Conference*.
- Schmeck, M., & Gengnagel, C. (2016). Calibrated Modeling of Knitted Fabric as a Means of Simulating Textile Hybrid Structures. *Procedia Engineering*, 155, 297-305. doi:<https://doi.org/10.1016/j.proeng.2016.08.032>
- Tamke, M., Sinke Baranovskaya, Y., Monteiro, F., Lienhard, J., La Magna, R., & Ramsgaard Thomsen, M. (2020). Computational knit – design and fabrication systems for textile structures with customised and graded CNC knitted fabrics. *Architectural Engineering and Design Management*, 1-21.
- Thomsen, M. R., Tamke, M., Karmon, A., Underwood, J., Gengnagel, C., Stranghöner, N., & Uhlemann, J. (2016). Knit as Bespoke Material Practice for Architecture. *ACADIA*.
- Thomsen, M. R., Tamke, M., Deleuran, A. H., Tinning, I. K. F., Evers, H. L., Gengnagel, C., & Schmeck, M. (2015). Hybrid Tower, Designing Soft Structures. In *Design Modelling Symposium 2015*.
- Wortmann, T. (2017). OPOSSUM: Introducing and Evaluating a Model-based Optimization Tool for Grasshopper. *CAADRIA Conference*.

# OPTIGAN: TOPOLOGICAL OPTIMISATION IN DESIGN FORM-FINDING WITH CONDITIONAL GANS

XUYOU YANG<sup>1</sup>, DING WEN BAO<sup>2,\*</sup>, XIN YAN<sup>3</sup> and YUCHENG ZHAO<sup>4</sup>

<sup>1</sup>*IDesignLab*

<sup>2,\*</sup>*School of Architecture and Urban Design & Centre for Innovative Structures and Materials, RMIT University*

<sup>3</sup>*Tsinghua University*

<sup>4</sup>*ByteDance*

<sup>1</sup>*xuyou.yang.92@gmail.com, 0000-0002-6294-378X*

<sup>2,\*</sup>*nic.bao@rmit.edu.au, 0000-0003-1395-8747*

<sup>3</sup>*yanxin13@mails.ucas.ac.cn, 0000-0002-5033-3597*

<sup>4</sup>*zhaoyucheng.joe@bytedance.com, 0000-0002-9291-2554*

**Abstract.** With the rapid development of computers and technology in the 20th century, the topological optimisation (TO) method has spread worldwide in various fields. This novel structural optimisation approach has been applied in many disciplines, including architectural form-finding. Especially Bi-directional Evolutionary Structural Optimisation (BESO), which was proposed in the 1990s, is widely used by thousands of engineers and architects worldwide to design innovative and iconic buildings. To integrate topological optimisation with artificial intelligence (AI) algorithms and to leverage its power to improve the diversity and efficiency of the BESO topological optimisation method, this research explores a non-iterative approach to accelerate the topology optimisation process of structures in architectural form-finding via conditional generative adversarial networks (GANs), which is named as OptiGAN. Trained with topological optimisation results generated through Ameba software, OptiGAN is able to predict a wide range of optimised architectural and structural designs under defined conditions.

**Keywords.** BESO (Bi-directional Evolutionary Structural Optimisation); Artificial Intelligence; Deep Learning; Topological Optimisation; Form-Finding; GAN (Generative Adversarial Networks); SDG 12; SDG 9.

## 1. Introduction

Structural optimisation, including topology optimisation, plays a significant role in architectural design. It can increase the performance of structures and the efficiency of material use and thus reduce material waste and carbon impact in the fabrication and construction process. By integrating topology optimisation with artificial intelligence

(AI) for more efficient use of materials in the industry, it seeks to help achieve the United Nations Sustainable Development Goal 12: Ensure sustainable consumption and production patterns (United Nations, 2015).

### 1.1. STRUCTURAL OPTIMISATION

Structural optimisations aim to achieve the best structural performance while meeting the requirement of various constraints. Over the past three decades, high-speed computers and rapid improvements in algorithms have been used to develop better structural optimisation solutions by a number of engineering researchers.

#### 1.1.1. Topology Optimisation

Topology optimisation is one of the most popular optimal structural design methods for discrete structures, such as trusses and frames. It is developed to search for the optimal spatial order and connectivity of the bars. Topology optimisation of continuum structures is to find optimal designs by determining cavities' best locations and geometries in the design domains.

Topology optimisation can be readily used to perform shape optimisation by simply restricting the structural modification to the existing boundaries (Huang & Xie, 2010). In the field of topology optimisation, there are several notable methods based on finite element analysis (FEA) developed, such as the homogenisation method (Bendsøe & Kikuchi, 1988), the solid isotropic material with penalisation (SIMP) method (Bendsøe & Sigmund, 1999), the evolutionary structural optimisation (ESO) (Xie & Steven, 1993), the bi-directional evolutionary structural optimisation (BESO) (Huang & Xie, 2010; Huang et al., 2007) and the level-set method (LSM) (Wang et al., 2003). In this paper, bi-directional evolutionary structural optimisation (BESO) proposed by Huang and Xie (2010) is adopted to develop a new integrated topology optimisation algorithm (Figure 1).

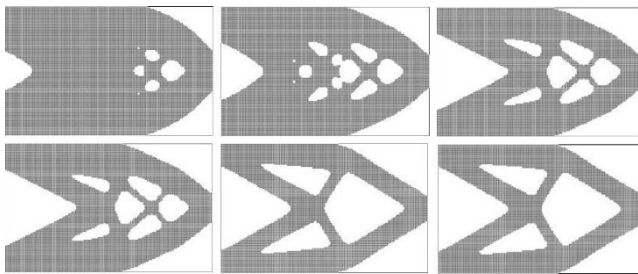


Figure 1 Bi-directional evolutionary structural optimisation (BESO) result

#### 1.1.2. Bi-directional Evolutionary Structural Optimisation (BESO)

Bi-directional evolutionary structural optimisation (BESO) is the emerging technology that is an extension of evolutionary structural optimisation (ESO) developed by Xie and Steven in 1992 (Xie & Steven, 1993). Both ESO and BESO algorithms are based on finite element analysis (FEA) for topology optimisation of continuum structures. BESO algorithm aims to find the solution with the highest structural performance

under certain material limitations by removing or adding material elements step by step (Bao et al., 2020). The ESO method also inspires the Extended ESO method, widely used in architecture design projects, such as the Akutagawa River Side Project in Japan by Ohmori and Qatar National Convention Centre by Arata Isozaki, to generate an optimised model with not only high structural performance but also some different characteristics to meet more functional requirements or aesthetic preferences. In the past few years, Mike Xie and his team have modified many detailed control strategies for topology optimisation in architectural design and development during the process (Yan et al., 2021).

### *1.1.3. Ameba Software*

Because of the benefit of form-finding through topology optimisation and Bi-directional evolutionary structural optimisation (BESO), more and more designers and architects seek to use topology optimisation methods to design buildings and furniture. However, due to the complexity and slow speed to directly use the algorithm for architectural design, a new Rhinoceros plug-in named Ameba, a topology optimisation tool based on the BESO method and FEniCS open-source computing platform (Zhou et al., 2018), has been developed. More and more architects and designers have gained opportunities to use this intelligent method to work with computers interactively to create innovative, efficient, and organic architectural forms using Ameba. In this work, the authors use it as the topology optimisation tool to form the dataset for training generative adversarial networks to assist and investigate the research.

## 1.2. GENERATIVE ADVERSARIAL NETWORK AND ITS APPLICATION IN TOPOLOGY OPTIMISATION

Allowed by the development of deep learning algorithms and fast-growing computational power, artificial neural networks, including generative adversarial networks (GANs), have been increasingly used in architectural and structural explorations such as topology optimisation in the design process.

### *1.2.1. Generative Adversarial Networks*

A generative adversarial network (GAN) is a particular artificial neural network that learns from a collection of examples and their probability distribution. It is then able to generate more examples from the estimated probability distribution (Goodfellow et al., 2020). A typical GAN often consists of a generator that defines a prior probability distribution  $P(z)$  based on a vector  $z$  as input and a discriminator which examines whether data  $x$  is real (sampled from the training examples) or fake (sampled from the output of the generator).

GANs can further be extended to conditional models (cGANs) where both the generator and discriminator are conditioned on extra information  $y$  as input (Mirza & Osindero, 2014). Besides examining whether  $x$  is real or fake, the discriminator of a cGAN also evaluates whether it matches the condition  $y$ . For example, when using cGANs to solve topology optimisation problems,  $x$  can be the expected optimisation results, given the corresponding boundary and load conditions of  $y$ .

### *1.2.2. Topology Optimisation via Deep Learning*

In recent years, there has already been some research into solving topology optimisation problems through artificial neural networks, especially GANs. For example, the TopologyGAN (Nie et al., 2021) is developed on a cGAN, whose generator combines a U-Net (Ronneberger et al., 2015) and ResNet (He et al., 2016). It takes displacement, load boundary conditions and target volume fraction augmented with dense initial fields computed over the unoptimised domain as the input to the model. By doing so, this method vastly improves the accuracy of predicting topology optimisation results compared to some baseline models. Another research by Yu et al. (2019) transforms the boundary conditions into multi-channel images as input to a convolutional-neural-network-based encoder to generate low-resolution topology optimisation results. It inputs the low-resolution results into a GAN to produce final results in high resolution. Differently, the proposed method utilises cGANs directly. It requires very brief input to keep the models easy to operate and thus broaden the potential range of users with or without professional structural knowledge.

### 1.3. PROPOSED METHOD

This research suggests an approach to accelerate the topology optimisation process of structures in architectural form-finding by replacing iterative calculation procedures with an end-to-end algorithm via conditional generative adversarial networks. This method is named OptiGAN by the authors as the ultimate goal is to generate topology optimisation results efficiently and accurately. Trained with a small number of topological optimisation results generated with Ameba software, the proposed method is able to predict a wide range of optimised two-dimensional structural forms under defined conditions.

## **2. Methods**

To achieve the research target, a coarse-to-fine network of cGANs is developed and trained with a dataset collected by the authors.

### 2.1. DATA COLLECTION AND PRE-PROCESSING

To train OptiGAN, an original dataset of topology optimisation results is collected with Ameba software. In detail, the material is kept as steel during the data generating process, and the volume fraction is set to 0.5 consistently. The parameters that can vary from case to case are design domain, fixing edge and load conditions, which are the very parameters to use as input parameters of OptiGAN. In practice, in the Ameba script, the fixing edge of a square design domain is always set to the left edge as fixing at other edges are seen as the same as rotating left-edge-fixing conditions in the later data augmentation process when training the cGAN models. In the first stage of the research, a total number of 1385 optimisation results are included in the dataset.



During the pre-processing, the input parameters are translated into a three-channel input image at  $256 * 256$  pixels in size. By doing so, the model is provided with two-dimensional spatial clues. Thus, the difficulty for training the model to predict two-dimensional results can be reduced compared to using merely numerical inputs directly without spatial suggestions. Specifically, as Figure 2 demonstrates, the design domain and fixing edge are expressed in all three channels by assigning the value of 0 to the corresponding pixels representing the geometries. Load locations and load unit vectors projected onto the X and Y axis are documented in the first and second channel respectively in values remapped into a range between 0 and 255. Values of the other pixels are assigned 255 in all three channels by default.

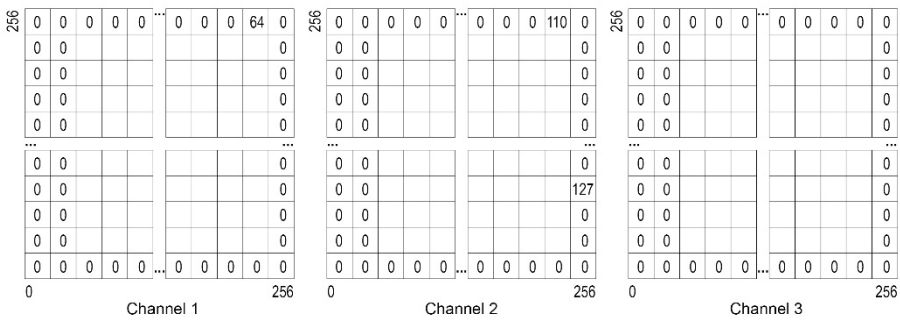


Figure 2 Translated input image of  $1000 * 1000$  mm design domain, fixing at the left edge, load one at  $(900, 1000)$  location in  $45^\circ$  angle, load two at  $(1000, 700)$  location in  $180^\circ$  angle. The values of the rest pixels are all 255.

## 2.2. OPTIGAN ARCHITECTURE

Unlike the previous researches mentioned in section 1.2.2, there are no dense initial fields, low-resolution results, or any other inputs than design domain, fixing edge and load conditions for the OptiGAN generator. Keeping the inputs simple can make it potentially as easy as adjusting a few number sliders for the OptiGAN users to operate. At the same time, it dramatically increases the difficulty for the generator to speculate the results by providing less input information. To respond to the conflict, OptiGAN adopts a coarse-to-fine network architecture.

Specifically, as Figure 3 demonstrates, there are two generators and discriminators in the network. The initial input of conditions in the form of translated input images is first fed to the coarse generator, which then predicts a rough output examined by the coarse discriminator. Then the rough output together with the initial conditions are input to the fine generator, which outputs the final results. Although compared to conventional cGANs, the input of OptiGAN consists only the conditions  $y$  alone without vector  $z$ , the networks can also learn from only the conditions. This choice of input is also being suggested in Pix2Pix (Isola et al., 2017), one of the most successful image-to-image translation models.

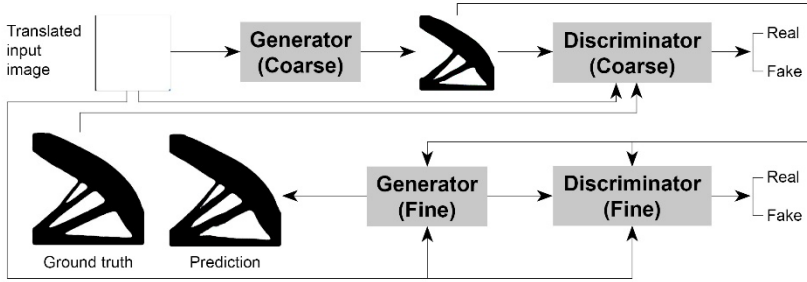


Figure 3 The architecture of OptiGAN

### 2.2.1. Generators and Discriminators

Both the coarse and fine generators used in OptiGAN are U-Net (Ronneberger et al., 2015), which has a mirrored encoder-decoder network architecture with skip connections between symmetrical layers. It can work effectively with very few training data. Meanwhile, both discriminators are PatchGAN (Isola et al., 2017), which focuses on  $156 * 156$  patches as through testing, such patch size provides the best outcome in the tasks.

### 2.2.2. Objective

The objective that OptiGAN tries to optimise can be expressed as equation (1). It consists of two parts, the cGAN loss (equation (2)) and L1 loss (equation (3)). Besides generating images that look real, as the other goal is to eventually achieve results as close to the BESO optimisation outcomes as possible, L1 loss is added to the total loss with a considerable weight of  $\lambda$  to force the output to be close to the ground truth. L1 loss is chosen out of L2 loss because it encourages less blurring effect of images compared to L2 loss. In the experiments, the weight is set to 125 ( $\lambda = 125$ ) to emphasise the importance of L1 loss in this particular task.

$$G = \arg \min_G \max_D \mathcal{L}_{cGAN}(G, D) + \lambda \mathcal{L}_{L1}(G) \quad (1)$$

$$\mathcal{L}_{cGAN}(G, D) = \mathbb{E}_{x,y}[\log D(x, y)] + \mathbb{E}_x[\log(1 - D(x, G(x)))] \quad (2)$$

$$\mathcal{L}_{L1}(G) = \mathbb{E}_{x,y}[\|y - G(x)\|_1] \quad (3)$$

## 2.3. TRAINING OPTIGAN

OptiGAN is trained in two steps: the coarse generator and discriminator are trained first, after which the fine generator and discriminator are trained. Both parts of the network are trained for 200 epochs to achieve the demonstrated results. Following the conventions, instead of minimising  $\log(1 - D(x, G(x)))$ , the trainings maximise  $\log(D(x, G(x)))$  to avoid saturating  $\log(1 - D(x, G(x)))$  in the early training stage. The initial learning rate is 0.0005 decaying to 0 in the last 100 epochs and the Adam optimiser is used. Figure 4 shows the history of the cGAN loss and L1 loss during the

two training procedures.

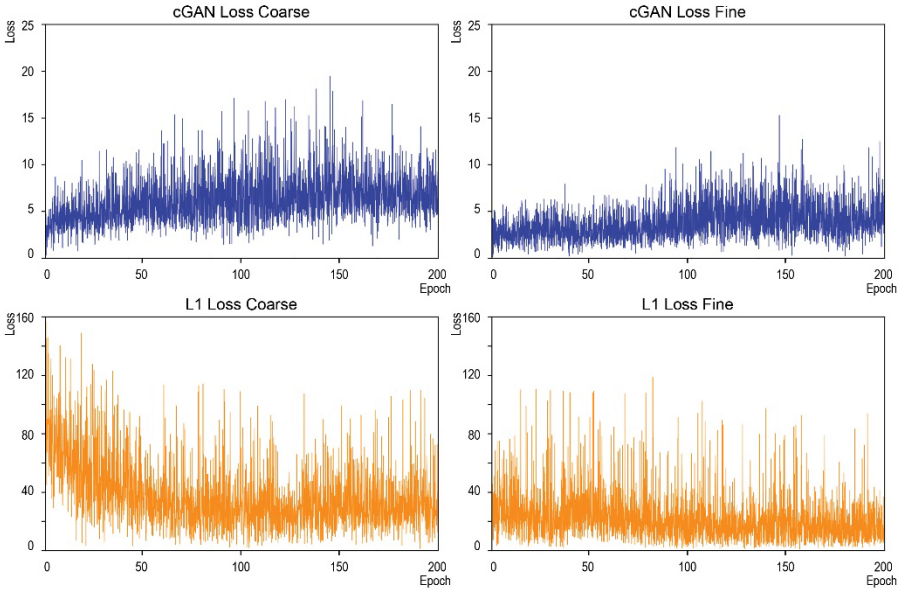
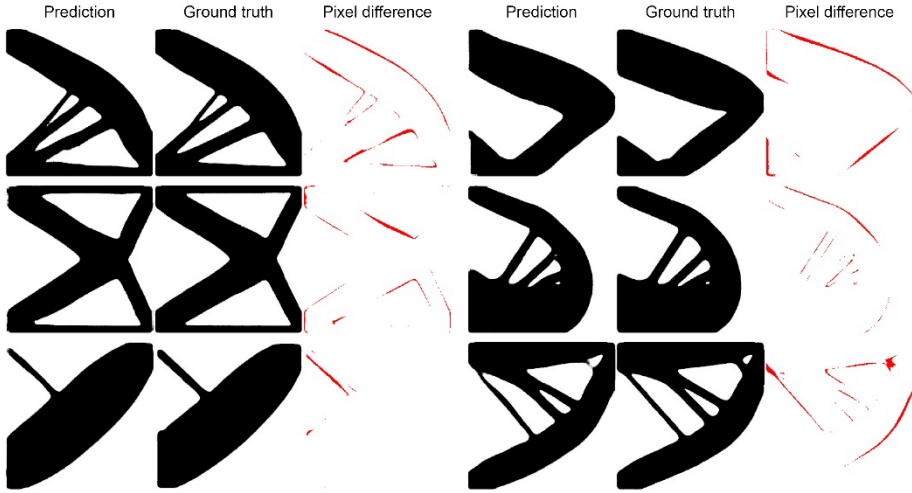


Figure 4 The cGAN loss and L1 loss during the coarse training and fine training

The coarse generator and fine generator take different types of input and are paired with different discriminators as introduced in section 2.1.1, so it is not considered practical to compare the absolute value of the corresponding cGAN losses of the coarse and fine models. Even though, the cGAN loss improves from the beginning to the last epoch in both cases, as shown in Figure 4. In contrast, for both generators, the L1 loss is calculated according to the same ground truth, and it can be discovered that the fine generator further decreases the L1 loss based on the coarse generator, which indicates that the fine generator is able to further improve the accuracy of predicted topology optimisation results by OptiGAN.

### 3. Results

During the training process, the L1 loss of OptiGAN reduced from over 80 to less than 20. More importantly and precisely, the pixel-wise accuracy is used to evaluate the performance of the models. It is equal to the percentage of accurate pixels in a prediction out of the total pixels of that image. Tested with 150 randomly selected pieces of data different from the training set, the average pixel-wise accuracy of OptiGAN is able to achieve 83.15%. Figure 5 demonstrates some of the testing results in different load conditions with pixel-wise differences between the predictions and ground truth visualised for each case.



*Figure 5 Some prediction results of OptiGAN with ground truth and pixel-wise difference*

#### 4. Discussion

OptiGAN demonstrates the ability of a novel approach and its application in architectural and structural form-finding. It is the extension of the SwarmBESO (multi-agent-based topology optimisation) method proposed by Bao & Yan in 2020 (Bao et al., 2021) to improve the diversity of the topological optimisation generative method. It has the potential to significantly help architects and engineers save material and produce more efficient structural layouts and building envelopes. It is valuable to integrate two intelligent computational design methods, deep learning and topology optimisation, for designers in the conceptual design phases.

However, the research is in a rudimentary phase and is temporarily constrained in a range of two-dimensional solutions. Although set as an input variable, the design domain in the current dataset includes only square geometries, despite that it can perform well in this geometric range, as demonstrated in the testing results. To truly diversify the spectrum of results and further increase the accuracy, it is very crucial that OptiGAN must be trained with much more data of various design domains and load conditions. Future line of research also includes further equipping the model with the ability to solve three-dimensional topology optimisation problems.

#### 5. Conclusion

This research develops OptiGAN, a non-iterative method to accelerate the topology optimisation process of structures in architectural form-finding via conditional generative adversarial networks with high accuracy. It demonstrates the process of integrating topology optimisation and generative adversarial networks to establish an artificial intelligence (AI) based structural optimisation technique. This new methodology holds great potential for practical application in architecture and

engineering fields. It increases the diversity of outcome of the topology optimisation generative design such as the application of shell (Figure 6).

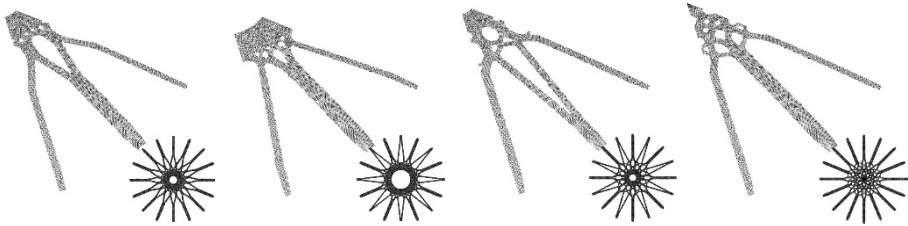


Figure 6 Diverse BESO results of shell optimisation

### Acknowledgements

We thank Nanjing Ameba Engineering Structure Optimization Research Institute for providing educational version of Ameba software to support our research. (For Ameba, see Ameba.xieym.com)

### References

- Bao, D. W., Yan, X., Snooks, R., & Xie, Y. M. (2020). Bioinspired Generative Architectural Design Form-Finding and Advanced Robotic Fabrication Based on Structural Performance. In *Architectural Intelligence* (pp. 147-170), Springer.
- Bao, D. W., Yan, X., Snooks, R., & Xie, Y. M. (2021). SwarmBESO: Multi-agent and evolutionary computational design based on the principles of structural performance. In *26th International Conference of the Association for Computer-Aided Architectural Design Research in Asia: PROJECTIONS CAADRIA 2021*, Hong Kong (pp. 241-250).
- Bendsøe, M. P., & Kikuchi, N. (1988). We are generating optimal topologies in structural design using a homogenisation method. *Computer methods in applied mechanics engineering*, 71(2), 197-224. doi:[https://doi.org/10.1016/0045-7825\(88\)90086-2](https://doi.org/10.1016/0045-7825(88)90086-2)
- Bendsøe, M. P., & Sigmund, O. (1999). Material interpolation schemes in topology optimisation. *Archive of applied mechanics*, 69(9), 635-654. doi:<https://doi.org/10.1007/s004190050248>
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, 63(11), 139-144. doi:10.1145/3422622
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, Las Vegas, (pp. 770-778).
- Huang, X., & Xie, M. (2010). *Evolutionary topology optimisation of continuum structures: methods and applications*: John Wiley & Sons.
- Huang, X., Xie, Y. M., & Burry, M. C. (2007). Advantages of bi-directional evolutionary structural optimisation (BESO) over evolutionary structural optimisation (ESO). *Advances in Structural Engineering*, 10(6), 727-737. doi:<https://doi.org/10.1260/136943307783571436>
- Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1125-1134).

- Mirza, M., & Osindero, S. (2014). Conditional generative adversarial nets. *arXiv preprint arXiv: 1411.1784*.
- Nie, Z., Lin, T., Jiang, H., & Kara, L. B. (2021). Topologygan: Topology optimisation using generative adversarial networks based on physical fields over the initial domain. *Journal of Mechanical Design*, 143(3), 031715.
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, (pp. 234-241).
- United Nations. (2015, September 25). *Transforming our world: The 2030 agenda for sustainable development*. Retrieved January 20, 2022, from: [https://https://www.un.org/ga/search/view\\_doc.asp?symbol=A/RES/70/1&Lang=E](https://https://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E).
- Wang, M. Y., Wang, X., & Guo, D. (2003). A level set method for structural topology optimisation. *Computer methods in applied mechanics engineering*, 192(1-2), 227-246. doi:[https://doi.org/10.1016/S0045-7825\(02\)00559-5](https://doi.org/10.1016/S0045-7825(02)00559-5)
- Xie, Y. M., & Steven, G. P. (1993). A simple evolutionary procedure for structural optimisation. *Computers & structures*, 49(5), 885-896. doi:[https://doi.org/10.1016/0045-7949\(93\)90035-C](https://doi.org/10.1016/0045-7949(93)90035-C)
- Yan, X., Bao, D., Zhou, Y., Xie, Y., & Cui, T. (2021). Detail control strategies for topology optimisation in architectural design and development. *Frontiers of Architectural Research*, 10, 1-17. doi:<https://doi.org/10.1016/j.foar.2021.11.001>
- Yu, Y., Hur, T., Jung, J., & Jang, I. G. (2019). Deep learning for determining a near-optimal topological design without any iteration. *Structural Multidisciplinary Optimization*, 59(3), 787-799. doi:<https://doi.org/10.1007/s00158-018-2101-5>
- Zhou, Q., Shen, W., Wang, J., Zhou, Y. Y., & Xie, Y. M. (2018). Ameba: A new topology optimisation tool for architectural design. In *International Association for Shell and Spatial Structures (IASS)*, (pp. 1-8).

# VISUAL CHARACTER ANALYSIS WITHIN ALGORITHMIC DESIGN

## *Quantifying Aesthetics Relative to Structural and Geometric Design Criteria*

ROBERT STUART-SMITH<sup>1</sup> and PATRICK DANAHY<sup>2</sup>

<sup>1,2</sup>*University of Pennsylvania, <sup>1</sup>University College London*

<sup>1</sup>*rssmith@design.upenn.edu, 0000-0003-3644-3906*

<sup>2</sup>*pdanahy@design.upenn.edu, 0000-0003-3393-8102*

**Abstract.** Buildings are responsible for 40% of world CO<sub>2</sub> emissions and 40% of the world's raw material consumption. Designing buildings with a reduced material volume is essential to securing a post-carbon built environment and supports a more affordable, accessible architecture. Architecture's material efficiency is correlated to structural efficiency however, buildings are seldom optimal structures. Architects must resolve several conflicting design criteria that can take precedence over structural concerns, while material-optimization is also impacted from limited means to quantitatively assess aesthetic decisions. Flexible design methods are required that can adapt to diverse constraints and generate filigree material arrangements, currently infeasible to explicitly model. A novel approach to generative topological design is proposed employing a custom multi-agent method that is adaptive to diverse structural conditions and incorporates quantitative analysis of visual formal character. Computer vision methods Gabor filtering, Canny Contouring and others are utilized to evaluate the visual appearance of designs and encode these within quantitative metrics. A matrix of design outcomes for a pavilion are developed to test adaptation to different spatial arrangements. Results are evaluated against visual character, structural, and geometric methods of analysis and demonstrate a limited set of aesthetic design criteria can be correlated with structural and geometric data in a quantitative metric.

**Keywords.** Generative/Algorithmic Design; Computer Vision; Environmental Performance; Multi-Agent Systems; Visual Character Analysis; SDG 10; SDG 11; SDG 9; SDG 12; SDG 13.

## 1. Introduction

Buildings are responsible for 40% of world CO<sub>2</sub> emissions and 40% of the world's raw material consumption (Bergman, 2013, pp. 24–25). Reducing material volume in building designs is an essential step towards a post-carbon built environment that can support a more equitable, affordable and sustainable future. Recent developments in

large format Additive Manufacturing (AM) potentially enable the fabrication of more materially efficient buildings. AM methods encompass a wide range of material and manufacturing constraints, yet typically allow substantial formal and topological complexity. Although these can support greater material efficiency and design expression, design methods that can generate and evaluate both criteria concurrently are relatively nascent. Architecture's material efficiency is correlated to structural efficiency; however, buildings are seldom optimal structures. Architects must resolve several conflicting design criteria that can take precedence over structural concerns. Examples include planning or site-related constraints, that might impose variations in column or wall spacings. Additionally, formal, or aesthetic criteria might undermine optimization of material volume. Flexible design methods are required that can adapt to diverse planning conditions while minimising material usage. The utility of such methods is contingent on their ability to provide materially quantitative feedback to the designer relative to their aesthetic design concerns. While optimisation methods can be integrated within generative architectural design processes (E.g., reinforcement-learning or evolutionary solvers (Mitchell, 2019)), these require fitness data to be obtained from design iterations. To integrate visual aesthetic criteria is challenging, as it requires a quantitative means of assessing qualitative formal character. Material volume offers one means to assess environmental and cost impacts of an architect's spatial, formal, or ornamental design propositions. Large-scale design decisions (E.g., floor planning, 3D massing or structural grids) are often constrained by programmatic or structural parameters that require explicit design input. Fortunately, AM architecture can support these constraints while enabling design expression and material optimization at smaller scales, traditionally considered to be the preserve of ornament. However, explicitly modelling such high-definition designs is prohibitively laborious, necessitating algorithmic or software-automated approaches.

## 2. State of the Art

Increases in a design's topological complexity can support greater material and structural efficiencies, as can be seen in the use of Topological Structural optimization (TSO) for the design of slabs, pavilion canopies, and large-span halls (Jipa et al., 2016; Sasaki et al., 2007). However, TSO operates primarily as a design-rationalization method with limited potential to directly inform aesthetic variability. Several algorithmic design approaches that encode a formal-aesthetic condition in addition to structural parameters have been developed for AM architectural elements such as columns and floor slabs (Anton et al., 2020). These have primarily focused on the design and manufacture of individual elements with limited demonstration of adaptation to variable spatial-structural conditions and no investigation into the relationship between their visual formal character and material volume. Multi-Agent design methods are particularly well suited to resolving spatial-formal problems and have already been utilized for architectural design (Snooks & Stuart-Smith, 2012). To date, multi-agent methods for AM architecture have not been developed that are adaptive to the varied structural conditions found in many buildings' structural layouts. Algorithmic methods also encode a designer's formal intentions to a large degree. Programmers/designers engage in subjective aesthetic decision-making when writing algorithms, and by assigning values to variables within code. During this process, a



designer has limited means to reflect or quantify visual character's impact on material efficiency. As algorithmic design and optimization methods operate on quantitative data, it is difficult to correlate qualitative concerns such as aesthetics. In the field of computer science, computer vision research has produced image processing methods that enable the quantification of visual features. Methods for pixel filtering, contouring, or clustering are critical to several data collection and machine learning algorithms (Davies, 2017). Recent research into the programming of aesthetics principles within deep learning approaches (Shaji, n.d.) suggests aesthetics can be engaged within algorithmic design. While architectural researchers recognise algorithmic processes may generate new aesthetics (Rehm, 2020), little research has been conducted into how evaluation of such aesthetics might inform optimisation processes, or impact a design's carbon footprint. Although clustering techniques have been used to classify architectural precedent buildings (Alymani et al., 2020), such methods have not been employed to evaluate visual formal character in algorithmically generated designs.

### **3. Algorithmic Design with Visual Formal Character Analysis**

This research proposes a novel approach to topological design through a custom Multi-Agent Design (MAD) method that is adaptive to spatial and structural inputs, and incorporates Visual Character Analysis (VCA) together with Structural Analysis (SA) and Geometric Analysis (GA). GA also supports estimation of Embodied Carbon (ECO2). Computer vision methods such as Gabor filtering, Hough Line, and Canny Contouring (OpenCV, 2021) are utilised to evaluate and encode the visual appearance of designs within quantitative metrics in addition to SA deflection and principal stress, and GA metrics for volume and surface area. The term 'visual character' can be used to describe a wide range of visible attributes. Within this research, a narrow set of visual characteristics were selected that were deemed relatively easy to quantify and aesthetically related to the design method's ability to produce increased topological complexity, including heterogeneity, intricacy, continuity, and recesses. Proposed visual characteristics and corresponding analytical criteria offer an extremely limited form of aesthetic evaluation. As a first foray into the quantification of visual character, these categories are utilised to establish a proof of concept, sufficiently general to be of relevance to a broad number of manufacturing methods and design approaches. No optimisation routines are undertaken within the research, instead, novel research into design and analysis methods is presented that could be utilised in conjunction with optimisation methods. A matrix of design outcomes for a pavilion was developed to test the method's adaptability to different spatial and structural arrangements. Outcomes are tested against visual character, structural, and geometric methods of analysis, to evaluate whether a limited set of aesthetic design criteria can be correlated with structural and material volumetric efficiency as a quantitative metric.

### **4. Methods**

A Multi-Agent Design method (MAD) (Figure 1) was developed that included quantitative analytical methods to provide SA, GA, and VCA metrics (Figure 2). Using the method, a matrix of design outcomes (Figure 3) was produced to evaluate the effectiveness of the approach, and its suitability to support optimisation processes.

#### 4.1. MULTI-AGENT DESIGN METHOD (MAD)

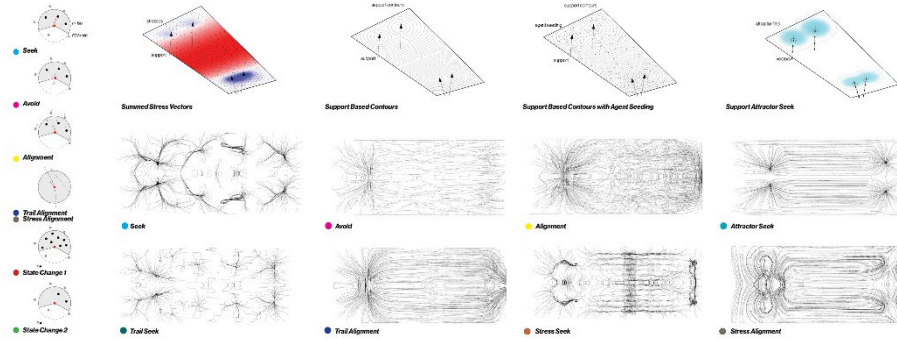


Figure 1. MAD Method: a) Agent rules b) Set-Out data and agent seeding. c) rule trajectories

A custom multi-agent design algorithm was developed to enable geometrically integrated ceilings, slabs and column/wall designs to be developed from a flexible, user-specified spatial set out of vertical supports, adapting designs to different planning and structural constraints within the Rhino3D software environment (Figure 3). A ceiling/slab mesh model with user-specified column and wall support locations is first structurally analysed in Karamba3D™, with resulting principal stress and deflection data stored in Python lists. A custom Python agent class inspired by Craig Reynolds Boids algorithm (Reynolds, 1987) was written that governs the motion of particles over simulated time. The ceiling mesh surface is populated with instances of the agent class, that seek and align to high stress areas, and attempt to travel down support locations. Agent motion trajectory curves are interpolated as a single mesh topology.

To ensure each simulation's initial set-out is consistent across diverse permutations of user-specified structural support conditions, a method was developed that provides a variable density distribution of agents relative to structural stress and surface area. A contour field is generated by uniformly offsetting curves that radiate out from each support position and Boolean-Unioned where they overlap. Seed points are randomly distributed between each contour curve relative to surface area between the contours at a probability of  $0.0157/M^2$ , and further culled relative to local principal stress values at a rate of: 0.700. This results in a variable density of agents relative to surface area and structural stress distributions, seeding more agents in high deflection areas.

Each instance of the agent class contains an (x,y,z) coordinate for current position, (x,y,z) vectors for velocity and acceleration, and an integer for its current 'state'. Each timeframe an agent sums an acceleration vector with its current velocity and adds this vector to its current position to move in space-time. Position and velocity vectors are constrained to the ceiling mesh by finding their closest points on the mesh. The acceleration vector is calculated each frame by summing results from a series of behavioural methods (Figure 1) that includes: agent-to-agent methods (seek, avoid, align), agent-to-agent trail methods (seek trail, align to trail), agent-to-structural data methods (seek local highest stress values, align to local stress vectors), and agent-to-vertical supports methods (seek top of support, align to support). Each method calculates the distance to neighbours of a specified type (agent, trail, structural data or

vertical support) within a specified radius (R) and field of view (FOV). For each method, vectors are calculated only relative to neighbours located inside of the agent's R and FOV. Seek methods calculate a vector pointing towards the average position of neighbours while avoid methods calculate an inverse vector. Align methods average the orientation of neighbours. All methods determine a steering vector that is the difference between their calculated vector and the agent's velocity vector, resulting in a vector that causes the agent to turn towards or away from each influence. All methods return a single vector which is then unitized to ensure they can be proportionally scaled relative to one another to influence an agent's behaviour. Agents operate under two states which utilise different weightings of the above method calculations, commencing as state "1", and changing to state '2' when in range of a vertical support to prioritize seeking and aligning to vertical supports (Figure 1).

Each agent's motion trajectory over 100 timeframes is translated to a Nurbs curve. As agents were constrained to the ceiling mesh, a method was created to adjust the z-height of each agent trajectory curve knot below the ceiling, relative to adjacent agent curves and structural stress. The displacement vector is calculated as the difference in position (pos) to neighbouring agents (a.pos) scaled by the closest principal stress value divided by the number of agents in range (n) and weighted by a constant (C):

$$\text{pos.z} = \text{pos.z} - \left( \frac{\sum(\text{distance}(\text{a.pos to pos})) * \text{stress}}{n} * C \right)$$

A continuous mesh geometry is generated around the adjusted curves by first creating polysurfaces from a single rail sweep from cross-section curve profiles arrayed along each curve. Polysurfaces are then converted to a mesh and imported into Zbrush™ and merged into a single mesh using Zbrush's dynamesh and smooth commands. This method was selected over more accessible methods in Rhino3D™ such as isosurfacing due to its benefits in enabling greater control over cross-sectional geometry.

#### 4.2. STRUCTURAL, GEOMETRIC & EMBODIED CARBON ANALYSIS

Design outcomes were geometrically analysed (GA) for surface area and volume, and structurally analysed (SA) for deflection and principal stress. GA metrics were obtained using built-in Rhino3D™ methods for mesh analysis. Although SA can be performed on meshes using a shell analysis, it would be computationally prohibitive if integrated within high-iteration optimization routines. As such, a frame analysis was performed using Karamba3D™ that was computable in a fraction of time. To prepare a structural model for frame analysis, each agent trajectory curve was converted to a polyline and exploded into several line elements. A series of connective lines between trajectories was generated between nodes in proximity to one another at a range equivalent to mesh cross-section profiles, producing a network with connectivity comparable to the mesh result. This was then used for SA. To estimate embodied carbon, a high-strength steel re-enforced concrete (RC40/50) with 100kg/m<sup>3</sup> steel reinforcement and 30% fly ash was specified with an embodied carbon (ECO2) value of 330 kgCO<sub>2</sub>/m<sup>3</sup> (MPA The Concrete Centre, 2020). By utilising volumetric data obtained in GA, an embodied carbon value is estimated for design outcomes.

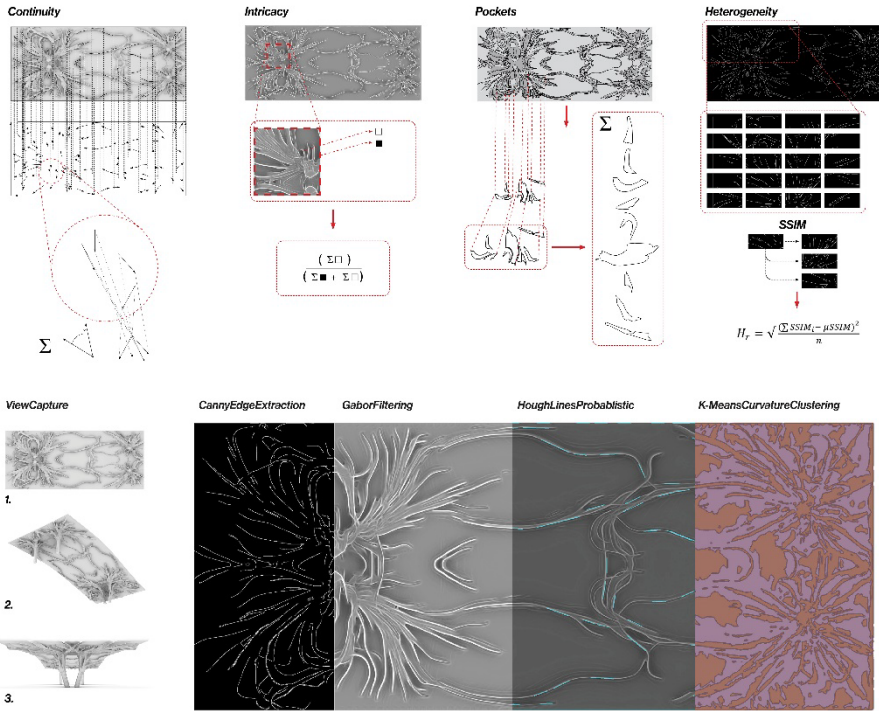


Figure 2. VCA. a) quantitative methods b) image processing methods applied to three camera views

#### 4.3. DESIGN VISUAL CHARACTER ANALYSIS (VCA)

To assess the visual character of design outcomes a methodology for image analysis was developed that involved viewport capture and image processing in Rhino3D™. For each design outcome, images at 1920x1080 resolution were captured using three specified camera views and Rhino3D's 'Arctic' display style. The computer vision framework OpenCV™ (OpenCV, 2021) was made accessible within Rhino3D by using GHPythonRemote™ to enable Python to be executable within Rhino3D's IronPython 2.7 GHPython components. A series of visual characteristics (including *intricity*, *heterogeneity*, *continuity*, and *surface recesses*) were established. To quantify these, specific OpenCV methods were selected that could generate new images that corresponded to each characteristic, and methods developed to extract and quantify data from each OpenCV image outcome. The specific methodologies are illustrated in Figure 2 and summarised below:

- *Intricity*: greater amounts of geometrical definition. A method was established to quantify edges in an image. A linear filter used for texture analysis, Gabor Filtering was selected due to its ability to highlight edges by identifying areas of highest difference in pixel value which are representative of edge curvature (OpenCV, 2021). In greyscale images, these are visualised as white pixels, whose quantity relative to total pixels in the image was quantified to describe a degree of intricacy.
- *Heterogeneity*: greater degrees of difference between different regions within an

image. The Canny edge detection method was utilised to first identify features in the geometry due to its ability to produce an image that depicts continuous edges in white on a black background through a multistep process that leverages spatial-frequency filtering and hysteresis thresholding (Davies, 2017, p. 136). A Structural Similarity Index (SSIM) Image Difference Comparison is then performed on the Canny image to evaluate the similarity between different tiled regions within the image (OpenCV, 2021). This returns a float for the standard deviation of SSIM distance evaluation between the image tiles and used as a value of Heterogeneity.

- *Continuity*: degree of alignment between edges in areas of high curvature. OpenCV's Probabilistic Hough Lines Transform method was utilised to create a series of line segments that approximate continuous edges within an image (OpenCV, 2021). As the Hough Transform is more effective if edge detection pre-processing is performed beforehand, a Canny Edge image was created and used as the input image. An average angle between Hough Line segments within proximity of Height/6 pixels to each other was calculated and averaged for all lines within the image to provide a metric for continuity.
- *Surface Recesses*: degree to which a surface geometry is separated into regions of varying depth. In lieu of the artic render, a coloured image was produced through the development of a custom Rhino3D Python script that coloured mesh vertices relative to local mesh curvature. Pockets are identified by segmenting the image into regions of different colour, thus segmenting by local minima/maxima of curvature in the surface. K Means Clustering of the image is used to categorise pixels of similar colour as a cluster (OpenCV, 2021). OpenCV contouring of pixel clusters enables the quantity of clusters to be identified as a metric for pockets.

#### 4.4. DESIGN MATRIX AND COMPARATIVE ANALYSIS

A matrix of pavilion designs was developed for four different spatial set-outs (A,B,C,D) that varied the number and spacing of columns, walls, and slab dimensions including an 8x8m 4-column grid, a 30x12m rectangular grid with symmetrical and asymmetrical column supports, and an option combining walls and columns. The MAD method was tested with three variations in ruleset values (labelled 1,2,3), together with an additional base condition (0) for each set-out that was explicitly modelled to have regular rectangular columns similar to Le Corbusier's Maison Domino (Figure 3c). VCA, GA, and SA metrics were developed for each design outcome. Metrics were re-mapped to a value between 0 and 1 and graphed (Figure 4) to support comparative analysis and to establish whether the method could be used to describe a multi-objective design fitness value.

### 5. Results and Discussion

The MAD method demonstrated successful adaptation to diverse spatial and structural conditions (Figure 3a). Minor differences in the rulesets had significant impact on outcomes, illustrating an expansive design space with varying degrees of geometric alignment and density. SA and GA results for all three rulesets performed similarly well in symmetrical spatial set outs A and B yet evidenced greater variability in

asymmetrical set outs C and D, indicating potential improvements could be achieved by inclusion of an optimization routine. While MAD and SA/GA methods were

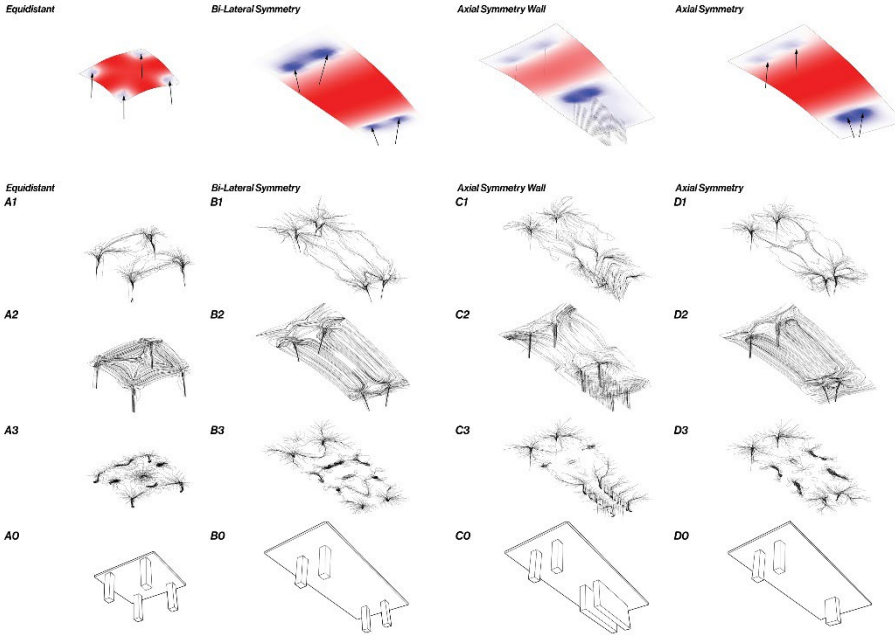


Figure 3. Matrix of design outcomes relative to user-specified column/wall and slab spatial organizations. a) structural loading conditions, b) agent trajectories, c) explicit base condition

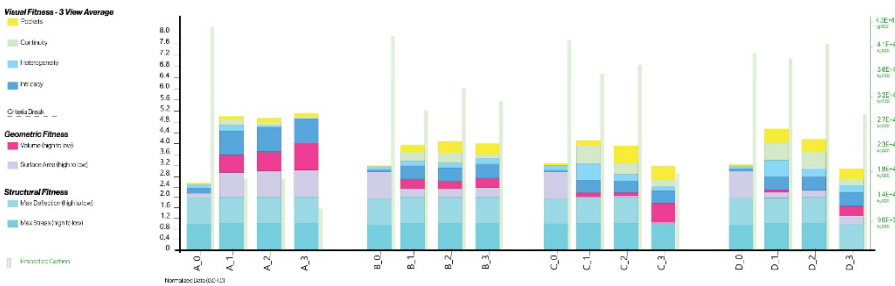


Figure 4. Comparative analysis of design matrix results

materially agnostic (with calibration to specific materials and performance requirements in future work), embodied CO2 (ECO2) was estimated using a high-strength reinforced concrete. ECO2 calculations derived from GA volumetric data demonstrated ruleset #1-3 MAD efficiency gains of 4-39% over base conditions. For set-out #A, a 23% reduction is equivalent to a saving of 4400 kgCO2.

As anticipated, all MAD rulesets scored substantially higher in VCA than base conditions. Evaluating VCA methods, Heterogeneity was validated by high-value results in asymmetrical set outs C and D, and ruleset #1's high values reflected a greater organisational diversity in its design outcomes. MAD results also demonstrated more

*intricacy* than base conditions yet remained relatively constant across all rulesets. This suggests that the quantification of Gabor filter images was not effective at describing *intricacy* of design outcomes, or, that all MAD rulesets generated similar levels of intricacy despite ruleset #3 consolidating detail in smaller clusters. *Continuity* was greatest in ruleset #1 and extremely low in base conditions as expected. K-Means Clustering was not as successful for evaluating recesses. Values were similar across all MAD rulesets, with the most heterogeneous results scoring lower than more homogenous outcomes. The number of clusters did not adequately describe the significance of features. As MAD Ruleset #1 outcomes have more distinguishable recesses at a range of scales compared to other rulesets, scale should be incorporated into the recesses evaluation method.

While MAD rulesets #1 and #2 scored higher in SA, ruleset #3 had superior GA results demonstrating the need for SA and GA feedback within an iterative optimisation process. The authors found MAD ruleset #1 exhibited preferable design outcomes that were formally continuous and heterogeneous (Figure 4). MAD ruleset #3 was least preferable due to a lack of heterogeneity resulting from dense clusters around supports. Results validate that selected visual character rules aligned to the author's aesthetic intentions yet highlight a conflict between author (or any user) preference, material, and structural optimisation. This demonstrates the importance of quantitative evaluation of visual characteristics to ensure that some aesthetic properties can be developed in relation to optimisation approaches. In assessment of the matrix of results, a suitable fitness function for a design optimization process might enhance intricacy(I), heterogeneity(H) and continuity(C) while reducing material volume (V) and structural deflection (D) and stress (S). This might be paraphrased as:

$$fitness = \frac{w_1I + w_2H + w_3C}{w_4V + w_5D + w_6S}$$

Whereby  $w_1$  through  $w_6$  are constants assigned by any user/designer to weight the influence of each parameter and could be adjusted to any designer's preferences.

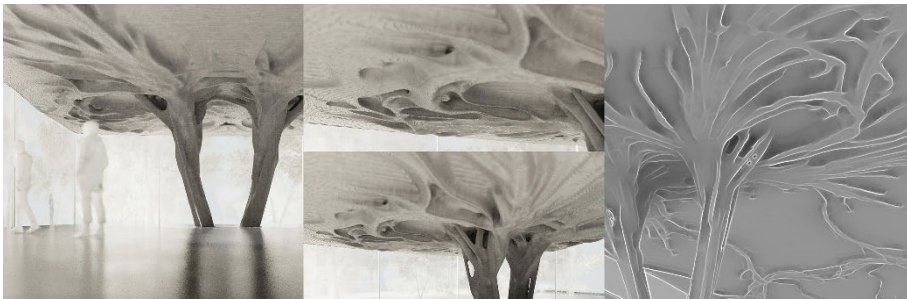


Figure 5. a) MAD outcomes adapt to structural and spatial constraints, while producing visual character that is quantitatively evaluated using b) VCA computer vision methods.

## 6. Conclusion

The research demonstrates a MAD method that integrates VCA methods together with SA, GA and, EC02 estimation. The MAD method demonstrated successful adaptation to diverse user-specified spatial and structural set-out conditions, while the use of

computer vision image processing techniques supported quantification of visual character alongside more readily quantifiable performance metrics for structural and material efficiency. Results illustrate the utility of the research but also highlight the necessity for inclusion of an optimization routine such as reinforcement learning to achieve results suited to material, and structural performance requirements. Given the challenges of establishing VCA methods and the limited criteria developed in this research, more work expanding VCA to encompass a broader set of generalisable aesthetic conditions would provide greater utility outside of this proof of concept. A survey paper of image-processing methods would be a good next step. Further development of the research will involve expansion of broader spatial/formal user design input, and incorporation of MEP and environmental performance considerations. AM architecture holds immense potential to reduce the material and environmental impact of building, while offering exciting opportunities for geometric design freedom. This research provides a means to develop topologically complex designs suited to AM methods together with feedback from structural, volumetric, and visual character metrics, providing a novel design approach that can be integrated with optimisation routines. It is hoped the research fosters a more holistic approach to design, correlating aesthetics with the material and structural efficiency of buildings, and thereby facilitating a reduction in the carbon footprint of architectural designs.

## References

- Alymani, A., Jabi, W., & Corcoran, P. (2020). Machine learning methods for clustering architectural precedents - classifying the relationship between building and ground. *In Anthropologic: Architecture and Fabrication in the cognitive age - Proceedings of the 38th eCAADe Conference*.
- Anton, A., Bedarf, P., Yoo, A., Reiter, L., & et al. (2020). Concrete Choreography: Prefabrication of 3D Printed Columns. *Fabricate Making Resilient Architecture*.
- Bergman, D. (2013). *Sustainable Design: A Critical Guide*. Princeton Architectural Press.
- Davies, E. R. (2017). *Computer Vision: Principles, Algorithms, Applications, Learning*. Elsevier Science.
- Jipa, A., Bernhard, M., Dillenburger, B., Meibodi, M., & Aghaei-Meibodi, M. (2016). 3D-Printed Stay-in-Place Formwork for Topologically Optimized Concrete Slabs. In K. Bieg (Ed.), *2016 TxA Emerging Design + Technology* (pp. 96–107).
- Mitchell, M. (2019). *Artificial Intelligence: A Guide for Thinking Humans*. Penguin Books Limited.
- MPA The Concrete Centre. (2020). Specifying Sustainable Concrete: Understanding the role of constituent materials.
- OpenCV. (2021). *About - OpenCV*. Open CV Team. <https://opencv.org/about/>
- Rehm, M. C. (2020). Other Experts: Disciplinary and Aesthetic Impacts of Artificial Intelligence. *Architectural Design*, 90(5), 14–21. <https://doi.org/https://doi.org/10.1002/ad.2606>
- Reynolds, C. W. (1987). Flocks, herds and schools: A distributed behavioral model. *ACM SIGGRAPH Computer Graphics*, 21(4), 25–34. <https://doi.org/10.1145/37402.37406>
- Sasaki, M., Ito, T., & Isozaki, A. (2007). *Morphogenesis of flux structure*. AA Publications.
- Shaji, A. (n.d.). *Understanding Aesthetics with Deep Learning | NVIDIA Developer Blog*. Retrieved November 11, 2021, from <https://developer.nvidia.com/blog/understanding-aesthetics-deep-learning/>
- Snooks, R., & Stuart-Smith, R. (2012). *Formation and the Rise of the Nonlinear Paradigm*. Tsinghua University Press.



# MACHINE LEARNING MODELING AND GENETIC OPTIMIZATION OF ADAPTIVE BUILDING FAÇADE TOWARDS THE LIGHT ENVIRONMENT

YUANYUAN LI<sup>1</sup>, CHENYU HUANG<sup>2</sup>, GENGJIA ZHANG<sup>3</sup> and JIAWEI YAO<sup>4</sup>

<sup>1</sup>*Department of Architecture, Qingdao City University*

<sup>2</sup>*School of Architecture and Art, North China University of Technology*

<sup>3</sup>*Department of Architecture, Tamkang University*

<sup>4</sup>*College of Architecture and Urban Planning, Tongji University*

<sup>1</sup>*lyuan93li@hotmail.com, 0000-0002-6761-8169*

<sup>2</sup>*huangchenyu303@163.com, 0000-0002-6360-638X*

<sup>3</sup>*r2328643@gmail.com*

<sup>4</sup>*jiawei.yao@tongji.edu.cn, 0000-0001-7321-3128*

**Abstract.** For adaptive façades, the dynamic integration of architectural and environmental information is essential but complex, especially for the performance of indoor light environments. This research proposes a new approach that combines computer-aided design methods and machine learning to enhance the efficiency of this process. The first step is to clarify the design factors of adaptive façade, exploring how parameterized typology models perform in simulation. Then interpretable machine learning is used to explain the contribution of adaptive facade parameters to light criteria (DLA, UDI, DGP) and build prediction models for light simulation. Finally, Wallacei X is used for multi-objective optimization, determines the optimal skin options under the corresponding light environment, and establishes the optimal operation model of the adaptive façades against changes in the light environment. This paper provides a reference for designers to decouple the influence of various factors of adaptive façades on the indoor light environment in the early design stage and carry out more efficient adaptive façades design driven by environmental performance.

**Keywords.** Adaptive Façades; Light Environment; Machine Learning; Light Simulation; Genetic Algorithm; Specific Instead of General; SDG3; SDG12.

## 1. Introduction

As the boundary between inside and outside of a building, building facades play an essential part within the broad scope of building performance. It has the potential to provide maximum comfort and a comfortable light environment for the interior space. With the popularity of large glass windows and curtain walls, the adaptive facade has become an important and sustainable method to control natural indoor light (Attia et

al., 2018). The traditional facade cannot respond to different light angles caused by the sun's movements and create constant visual comfort for the interior during daylight (Michael & Heracleous, 2017). At the same time, the adaptive facade can respond to changes over time and find a balance between different light evaluation indicators to keep a comfortable indoor light environment (Hosseini et al., 2019; Luna-Navarro et al., 2020).

On the basic definition of adaptive façade, the concept of adaption requires effective control, which could be distinguished as extrinsic and intrinsic control (Loonen et al., 2013). Extrinsic control is a computer-based system that reacts with input and output that decides the behaviour of adaptive facades. For instance, Shi et al. (2020) compare two types of façade typology for light and energy performance. However, the expected outcomes are difficult to predict due to the complexity of climate data and the difficulty of massive light stimulation, which results in the need for intelligent decision-making in the control of adaptive facades. In the architectural field, new computer-aided methods like genetic Algorithms (Carlucci et al., 2015; Rizi & Eltaweel, 2021) are introduced to face the multi-functionality and non-linear adaptations which can be transferred to comparative design. Adaptive façade needs an intelligent control system that could make decisions on collected information and existing requirements (Böke et al., 2019). To achieve the intelligent goal, machine learning is introduced to establish a new workflow for adaptive façade towards lighting performance.

## 2. Research Methodology

The software for light environment simulation often requires interdisciplinary knowledge like physics and programming skills that are usually beyond the designer's ability. Therefore, simulation for the adaptive facade in the early design stage is difficult to synergy. Designers usually use an exhaustive optimization method, which is impossible to respond to changing environments. A computer-aided design approach and machine learning can be combined in the early design stage, effectively improving the design process and achieving accurate optimization results.

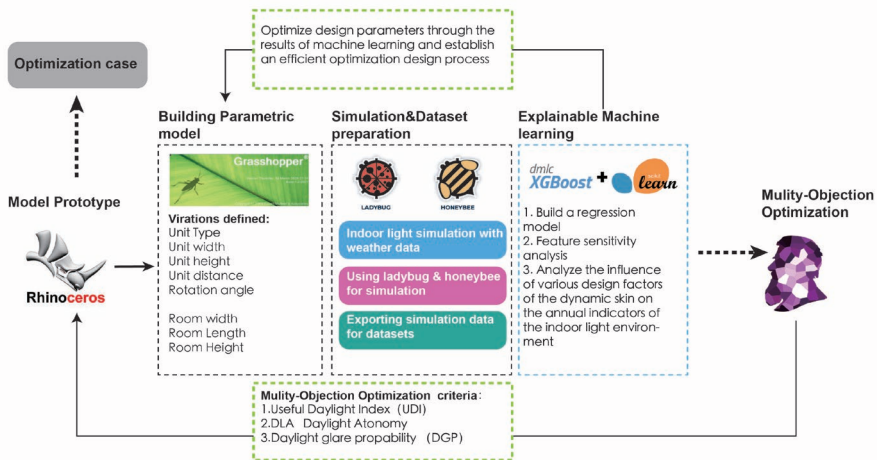


Figure 1. Research workflow

### 2.1. RESEARCH PATH

As shown in Figure 1, the research framework of this paper consists of four main parts. Based on domestic and international studies and cases, the first step is analyzing common dynamic skin design parameters, then establishing three parametric facades models based on Rhino/Grasshopper platform. Secondly, using simulation software such as Honeybee and Ladybug to evaluate the light environment and get indexes such as Daylight autonomy (DLA), daylight illuminance (UDI), and daylight glare probability (DGP) of the models. All the variations and indexes are exported to construct sample datasets by controlling the parametric models for batch simulations. The contribution of the dynamic epidermal parameters (unit length, unit width, unit distance, material transparency, different patterns, etc.) to the indoor light environment index under the corresponding light environment (weather data), and several regression models were developed to obtain the prediction models of different dynamic epidermal parameters under the corresponding light environment. Finally, a multi-objective optimization based on Wallacei X of genetic algorithm is performed to determine the optimal epidermal option under the corresponding light environment and establish the optimal operation model of the adaptive facade for light environment changes.

### 2.2. PARAMETRIC DYNAMIC SKIN ESTABLISHMENT

The adaptive facades involved in this study are unit modules, which form a whole through repeated arrays in horizontal and vertical directions. The independent units themselves are difficult to have a good effect in function and aesthetics, but the whole formed by the array can provide a good light environment for the room, and at the same time, can also constitute a certain sense of rhythm.

There are three main unit forms chosen in the paper: rectangular, triangular, and hexagonal. The size of the design unit is generally arranged in a way that there are 2-3 rows of one-floor height, and the building floor height is multiplied by the basic unit height, i.e.

$$H = n * h$$

where h is the height of the dynamic unit, H is the floor height of the building, and n is the multiplicative relationship. Different samples of the three parametric dynamic skins are shown in Figure 2.

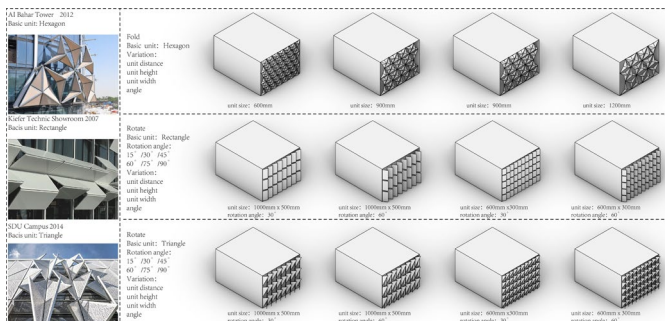


Figure 2. Three parametric prototypes

Ultimately, through simplification and for unification purposes, the dimensions of the room (face width, depth, floor height) and the dimensions of the parametric dynamic skin (unit distance, unit length, unit width, unit movable parameters) are used in this paper as potential influencing factors that affect the light environment indexes.

### 2.3. SIMULATION FOR INDOOR LIGHT ENVIRONMENT AND DATASETS PREPARATION

#### 2.3.1. Indexes for light environment

This paper uses the Ladybug and Honeybee plug-ins of the Rhino/Grasshopper platform to simulate the indoor light environment with a parametric dynamic skin. The adaptive facade can significantly improve the light environment near the glass curtain wall to meet the appropriate natural illumination and avoid glare during the daytime. Considering that office spaces with large depths are prone to uneven illumination, and small-sized residential spaces near glass curtain walls and floor-to-ceiling windows are often strongly lit and generate glare, this paper generates adapted dynamic skins for rooms of different sizes and performs batch light environment simulation by setting the room size as a variable. Based on the LEED and WELL building evaluation standards, the final light environment indicators such as all-natural lighting percentage (DLA), effective daylight illuminance (UDI), and daylight glare probability (DGP) are selected, as shown in Table 1.

Table 1. Light environment criteria

Criteria	Definition	Unit	Comfortable value
<b>DLA</b> (Daylight Autonomy)	The percentage of annual work hours during which all or part of a building's lighting needs can be met through daylighting alone.	%	320Lux
<b>UDI</b> (Useful Daylight Illuminance)	Useful daylight illuminance (UDI) is a daylight availability metric that corresponds to the percentage of the occupied time when a target range of illuminances at a point in space is met by daylight.	%	Autonomous UDI ( 500lux—2500lux)
<b>DGP</b> (Daylight Glare Probability)	DGP assesses the probability that an occupant will be disturbed by glare conditions using the vertical eye illuminance of an entire scene and its relationship to specific potential glare sources within that scene (2006)	%	<0.35

#### 2.3.2. Simulation of interior light environment

The light environment simulation is mainly based on Rhino/Grasshopper platform, using components of light environment simulation such as Ladybug and Honeybee to simulate light environments for indoor spaces with adaptive façades. The meteorological data were imported by reading EPW files through ladybug (Table 2), the selected location was Shanghai, and the imported climate-related data were weather, time (month and day), sky type, etc.

Table 2. Weather data

Input weather variations	Definition
dirRad	The hourly Direct Normal Radiation
diffRad	The hourly Diffuse Horizontal Radiation
skyType	CIE Sky Type [0] Sunny with sun, [1] sunny without sun, [2] intermediate with sun, [3] intermediate without sun, [4] cloudy sky, [5] uniform sky

### 2.3.3. Datasets preparation and correlation analysis

In this study, the preparation time of the dataset is relatively long. Batch simulations use the ladybug fly component, and the CPython component will export results. Each parametric prototype model generates 500 simulation cases, and the corresponding three light environment indicators will be exported to construct sample datasets for machine learning, the parametric model and simulation process is shown in Figure 3.

The correlation analysis of the datasets is an important part of the data exploration in the pre-modeling stage. By calculating the correlation between the characteristic values (independent variable: design parameters) and the target values (dependent variable: three light environment indicators), the correlation degree of the design parameters of the adaptive façade model to the light environment indicators can be derived, which in turn provides support for the designer's design decision.

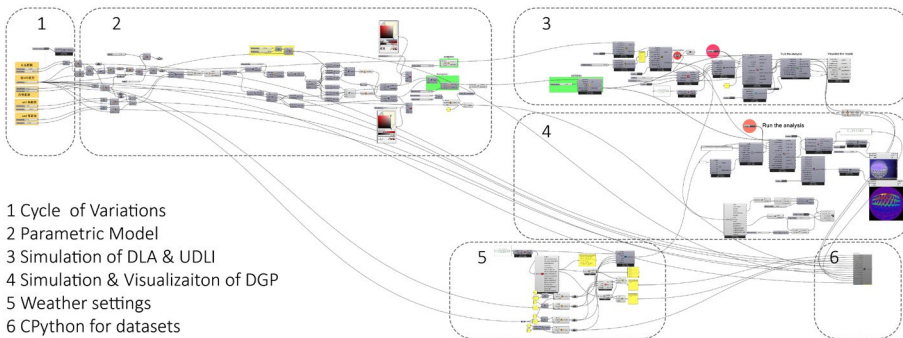


Figure 3. Grasshopper battery for one prototype

## 2.4. INTERPRETABLE MACHINE LEARNING

Current building environmental performance assessment is mainly based on the design "post-evaluation paradigm," which is difficult to influence the optimization of design solutions. The complexity and time cost of environmental performance simulation in the early stage of design makes it difficult to introduce environmental performance evaluation tools in the early decision-making stage. The introduction of machine learning, however, can accelerate the performance simulation and quickly give predicted results of environmental simulation in the early design phase when certain design parameters are uncertain. In this study, an integrated model XGBoost (Extreme

Gradient Boosting), which is an integrated machine learning algorithm based on decision trees with Gradient Boost as a framework for the prediction problem of unstructured data (images, text, etc.), was chosen, and the artificial neural network outperforms other algorithms or frameworks. However, when dealing with small to medium-sized structured data or tabular data, decision tree-based algorithms are now generally considered to be the best. XGBoost can be used for classification and regression tasks and provides an explanation of the importance of each feature at the algorithmic level, making it an important algorithm for interpretable machine learning. In this paper, we will use the three aforementioned approaches to establish a mapping of dependent variables representing model parameters to light environment metrics and to interpret the feature importance of the regression model, as well as to derive the extent to which the model parameters contribute to the indoor light environment. The main results will be given in section 3.

## 2.5. GENETIC ALGORITHMS AND MULTI-OBJECTIVE OPTIMIZATION

Wallacei (which includes Wallacei Analytics and Wallacei X) is an evolutionary multi-objective optimization and analytic engine. The multi-objective optimization algorithm Wallacei X is advantageous here because it allows selecting one or more optimal solutions from the machine learning prediction models. This evolutionary solver can consider several objective functions simultaneously to determine the optimum solution. In addition, the solver allows the user to store and save arbitrary data for each iteration of the design.

## 3. Results and discussion

### 3.1. TRAINING DATASETS

Light environment simulations were performed on the Rhino/Grasshopper platform using Ladybug and Honeybee components. The Shanghai epw data was used as the meteorological data. The initial settings were used in the light environment simulation: clear sky, sunlight, minimum 500 lux at the height of 0.85 from the floor on the working plane, occupancy schedule (8-16), sensor grid with a scale of 0.5\*0.5 m<sup>2</sup>, and no shadows or artificial light. For further simulation of the real climate, more weather variations are added. Batch simulations were performed using the ladybug fly component, and simulation results were looped and saved using the pandas to import into Cpython component. The section of the constructed dataset is shown in Table 3 below.

Table 3. Datasets preparation

	dirRad	difRad	skyType	room_width	room_lengh	room_height	unit_distance	unit_width	unit_height	unit_angle	DLA	UDLI_100_2000	DGP
0	813	144	2	12536.189430	7364.017132	5178.729810	273.757	1054.070972	1515.324367	0.879547	93.411429	65.717143	0.340175
1	0	64	5	8238.017257	12065.844960	5938.155484	751.331	576.496286	2560.174995	0.803791	74.825521	74.468750	0.247651
2	596	259	2	12939.845085	7767.672787	3697.581158	228.906	1098.921601	1605.025623	0.728035	80.656000	74.885333	0.297528
3	13	155	5	8641.672913	12469.500615	4457.006832	706.481	621.346915	2649.876252	0.652280	55.855392	69.372549	0.238960
4	771	124	3	13343.500741	8171.328443	5216.432506	184.055	1143.772229	1694.726880	0.576524	82.997596	75.463942	0.299081

First, the six model parameter variables and three light environment indicators were plotted in a bivariate scatter matrix to test whether there was a linear correlation between the characteristic variables and the response variables, as shown on the left of Figure 4, from which it was seen that several characteristic variables were significantly linearly correlated with the response variables, so the Pearson method should be used for correlation analysis of each variable.

The correlation coefficient between room\_length and DLA is -0.85, a negative correlation. The correlation coefficient between unit\_width and DLA is 0.46, a positive correlation. The correlation coefficients of room\_width, room\_height, and unit\_distance with DLI were 0.46, 0.4, and 0.49, respectively, all negatively correlated. In addition, the correlation coefficient between room\_width and DGP is 0.81, which has a positive correlation. From the above results, it is clear that. Room\_width, room\_length, room\_height, and unit\_width have more influence on the light environment indexes than other parameters, and the direction and degree of influence on the three light environment indexes are different.

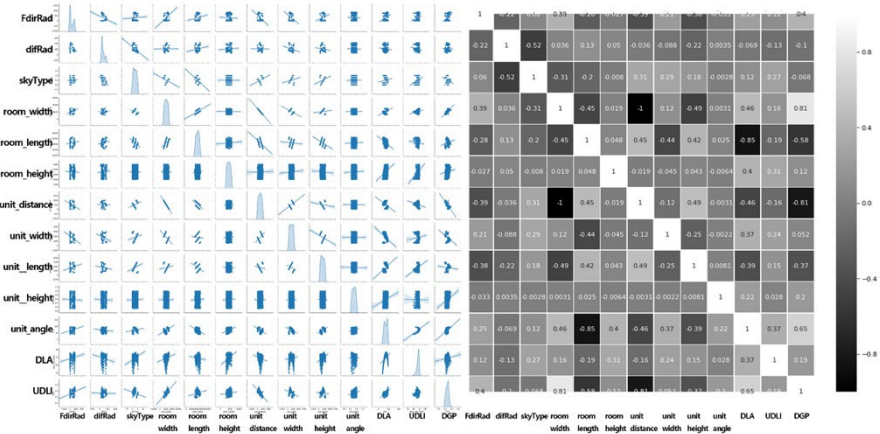


Figure 4. Bivariate scatter matrix and Pearson correlation coefficient matrix

### 3.2. XGBOOST MACHINE LEARNING MODEL

Permutation Importance is an algorithm that calculates the importance of model features. Feature importance refers to how much a feature contributes to the prediction. Some models, such as XGBoost, LR, decision trees, etc., can calculate feature importance directly. Assuming that the importance of a feature is to be studied, then the data in this column is disrupted, the data in the other columns are kept constant, and then the accuracy of the prediction of the regression loss is observed to change by how much. The feature importance is determined based on the amount of change, which is essentially similar to feature sensitivity analysis. The nine model parameters are trained using the Python language based on the scikit-learn machine learning tool for DLA, UDI, and DGP using the XGBoost model respectively, and are calculated using the feature importance function that comes with the model, the outcome is shown in figure 5.

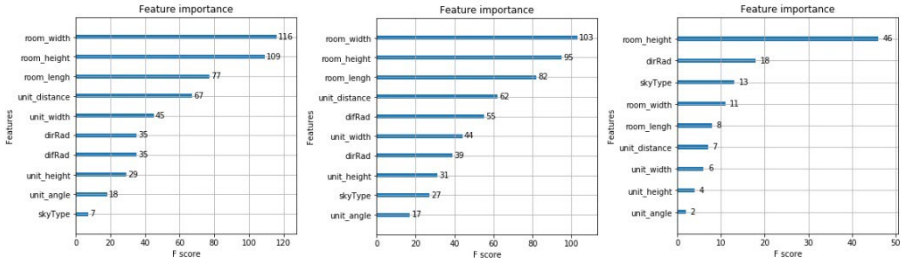


Figure 5. Feature Importance of prototype 2

With CPython in Grasshopper, the exported machine learning model can be imported into the Grasshopper platform, the relationship between the variation and light simulation outcome can be calculated fast and easily to read.

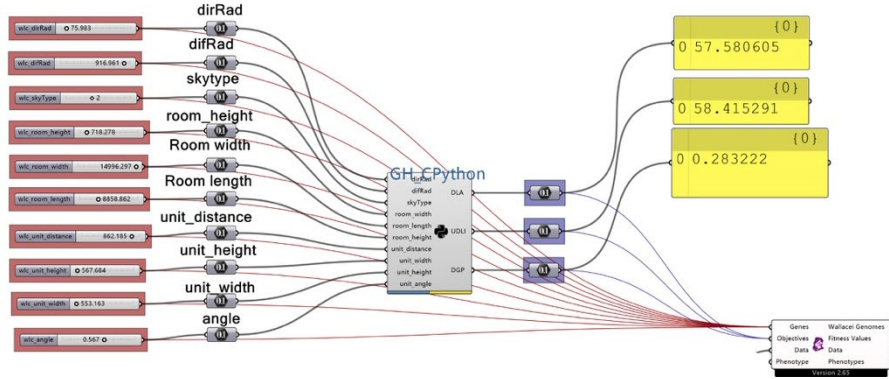


Figure 6. Machine learning model connected to Wallacei

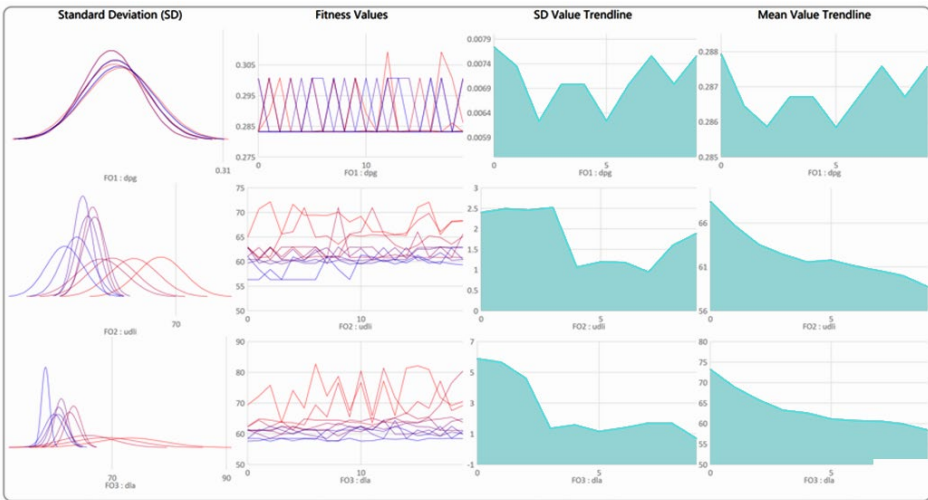


Figure 7. Wallacei analytic



### 3.3. GENETIC ALGORITHMS AND MULTI-OBJECTIVE OPTIMIZATION

By running the multi-objective evolutionary simulation for the main design problem, the setting for generation count is 10; each one contains 20 populations. In total, 200 genotypes with three fitness values per solution were produced. The evaluation of the multi-objective optimization process commenced with a set of analyses to examine how the evolutionary simulation addressed the fitness objectives and how successfully it performed in its entirety. The outcome is shown in figure 7.

## 4. Conclusions

This study proposes a workflow that uses parametric 3D modeling (Grasshopper), light environment simulation (Ladybug), and machine learning (XGBoost) modules combined with Genetic Algorithm (Wallace X) to optimize and select design cases from a large number of models corresponding to the light environment in the pre-design phase of adaptive facade design. The integrated design aims to find suitable adaptive facade designs that can effectively improve natural light glare in the room and provide sufficient illumination.

The workflow shows flexibility in the selection of sample evaluation sizes, especially saving time in the optimization process. This versatility also allows to analyze the final phenotype and plot the results in a fast way in a user-friendly interface. The genes and genomes information extracted from the selected phenotypes are then stored for subsequent replication and evaluation of the environmental targets. This is an important advantage of the implemented algorithm over other currently available algorithms that do not allow the storage of informative data.

In addition, using genetic algorithms with machine learning shows high efficiency for data-driven design towards light environment indicators (DLA, UDLI, and DGP) optimized. The current workflow allows building a database of different facade prototypes with machine learning predictive models, saving a significant amount of simulation time and responding to the climate timely to enhance indoor light comfort. However, the number of adaptive facade prototypes in this study is insufficient and could not be real-time visualized. In subsequent studies, the number of model prototypes will be increased, and a workflow that allows visualization feedback in real-time will be established. And this study can be extended to analyze different geographical locations, shapes of structures, and adaptations to specific latitudes and weather conditions, and further work should be carried out on more assessments.

## References

- Atia, S., Bilir, S., Safy, T., Struck, C., Loonen, R., & Goia, F. (2018). Current trends and future challenges in the performance assessment of adaptive façade systems. *Energy and Buildings*, 179, 165–182. <https://doi.org/10.1016/j.enbuild.2018.09.017>
- Böke, J., Knaack, U., & Hemmerling, M. (2019). State-of-the-art of intelligent building envelopes in the context of intelligent technical systems. *Intelligent Buildings International*, 11(1), 27-45. <https://doi.org/10.1080/17508975.2018.1447437>
- Carlucci, S., Cattarin, G., Causone, F., & Pagliano, L. (2015). Multi-objective optimization of a nearly zero-energy building based on thermal and visual discomfort minimization

- using a non-dominated sorting genetic algorithm (NSGA-II). *Energy and Buildings*, 104, 378–394. <https://doi.org/10.1016/j.enbuild.2015.06.064>
- Hosseini, S. M., Mohammadi, M., & Guerra-santin, O. (2019). Interactive kinetic façade: Improving visual comfort based on dynamic daylight and occupant's positions by 2D and 3D shape changes. *Building and Environment*, 165, 106396. <https://doi.org/10.1016/j.buildenv.2019.106396>
- Loonen, R. C. G. M., Trčka, M., Cóstola, D., & Hensen, J. L. M. (2013). Climate adaptive building shells: State-of-the-art and future challenges. *Renewable and Sustainable Energy Reviews*, 25, 483–493. <https://doi.org/10.1016/j.rser.2013.04.016>
- Luna-Navarro, A., Loonen, R., Juaristi, M., Monge-Barrio, A., Attia, S., & Overend, M. (2020). Occupant-Facade interaction: a review and classification scheme. *Building and Environment*, 177(March), 106880. <https://doi.org/10.1016/j.buildenv.2020.106880>
- Michael, A., & Heracleous, C. (2017). Assessment of natural lighting performance and visual comfort of educational architecture in Southern Europe: The case of typical educational school premises in Cyprus. *Energy and Buildings*, 140, 443–457. <https://doi.org/https://doi.org/10.1016/j.enbuild.2016.12.087>
- Rizi, Rana Abdollahi, and Ahmad Eltaweel. "A User Detective Adaptive Facade towards Improving Visual and Thermal Comfort." *Journal of Building Engineering*, vol. 33, no. June 2020, Elsevier Ltd, 2021, p. 101554, doi:10.1016/j.jobe.2020.101554.
- Shi, X., Abel, T., & Wang, L. (2020). Influence of two motion types on solar transmittance and daylight performance of dynamic façades. *Solar Energy*, 201(March), 561–580. <https://doi.org/10.1016/j.solener.2020.03.017>
- Tabadkani, Amir, et al. "Integrated Parametric Design of Adaptive Facades for User's Visual Comfort." *Automation in Construction*, 106, 102857. doi:10.1016/j.autcon.2019.102857.

# REINFORCEMENT LEARNING-BASED GENERATIVE DESIGN METHODOLOGY FOR KINETIC FACADE

SIDA DAI<sup>1</sup>, MICHAEL KLEISS<sup>2</sup>, MOSTAFA ALANI<sup>3</sup> and NYOMAN PEBRYANI<sup>4</sup>

<sup>1,2</sup>*Clemson University.* <sup>3</sup>*Aliraqia University.* <sup>4</sup>*Institut Seni Indonesia Denpasar.*

<sup>1</sup>*sidad@clemson.edu, 0000-0002-5883-381X*

<sup>2</sup>*crbh@clemson.edu, 0000-0002-4347-3890*

<sup>3</sup>*mostafa.waleed@aliraqia.edu.iq, 0000-0002-7494-8940*

<sup>4</sup>*dewipebryani@isi-dps.ac.id, 0000-0002-4931-7071*

**Abstract.** This paper presents a reinforcement learning (RL) based design method for kinetic facades to optimize the movement direction of shading panels. Included with this research is a case study on the Westin Peachtree Plaza in Atlanta, USA to examine the effectiveness of the proposed design method in a real-life context. Optimization of building performance has been given increased attention due to the significant impact buildings have on energy consumption and carbon emissions. Further, building performance is closely related to the “Sustainable Cities and Communities” mentioned in SDG11. Results show that the novel design method improved the building performance by reducing solar radiation and glare and illustrate the potential of RL in tackling complex design problems in the architectural field.

**Keywords.** Reinforcement Learning; Kinetic Facade; Generative Design; Design Methodology; SDG 11.

## 1. Introduction

With the increasing demand for building energy performance and occupant comfort, kinetic facades (building facades with kinetic elements) have attracted increased interest in recent years due to its ability to adapt to the external environment by actively changing itself. For instance, Al Bahr Tower in Abu Dhabi has a kinetic facade system that reduces 50% energy consumption for office spaces and 20% for the whole building (Karanouh and Kerber, 2015).

Many studies show that an optimized control system can improve the performance of kinetic facades (Loonen et al., 2013; Zhang et al., 2018; Smith and Lasch, 2016). However, expensive sensors and actuators in control systems significantly increase the facade cost (Attia et al., 2018). Improving the performance of kinetic facades through generative design creates higher requirements for the design method (Hosseini et al., 2019; Mahmoud and Elghazi, 2016).

The rapid development of artificial intelligence provides a new possibility for

optimizing kinetic facades. In the last 20 years, machine learning achieved considerable growth as a result of the rapid increase of computing power and the availability of large data sets (Russell and Norvig, 2016). Based on a reinforcement learning (RL) algorithm in machine learning, this paper presents a generative design method for kinetic facades. It improves the facade performance by optimizing the movement direction of each unit. There are two design objectives in this research: (1) Reducing the direct solar radiation heat on the building surface; and (2) improving occupants' visual comfort by reducing glare.

The following sections include the introduction of RL, including its theory, algorithms, and applications in the architectural field. Then the novel design system will be described based on the RL framework. A case study on the Westin Peachtree Plaza in Atlanta, Georgia is included to show a precedent for the design process alongside an experiment to verify the effectiveness of this design method. The last two sections discuss the experiment results and the contribution of this research.

## 2. Reinforcement Learning in Design

### 2.1. REINFORCEMENT LEARNING

Machine learning algorithms have multiple classes of learning: Supervised, Unsupervised, Semi-supervised, and Reinforced. The machine learning algorithm used in this study is RL. Compared with other algorithms, RL increases the efficiency of solving complex design problems with no training set and model-free algorithms (Sutton and Barto, 2018). Figure 1 shows the framework of the RL system.



Figure 1. Framework of Reinforcement Learning

The two indispensable subjects in this system are agent and environment. The agent can observe and implement action to the environment. After that, the environment can change its state based on the agent's action and give feedback to the agent. The agent learns how to maximize rewards through continuous interaction with the environment. In addition, the specific algorithm determines the way the agent learns from interactions.

### 2.2. Q-LEARNING ALGORITHM

The specific algorithm used in this research is Q-learning. Q-learning is a widely used model-free RL algorithm based on Markov Decision Process (MDP). Model-free MDP algorithms include Monte-Carlo Learning and Temporal-Difference Learning. Q learning uses the Q table to evaluate the pros and cons of different actions in a state. The learning process of Q-learning is the updating process of the Q table. Based on the Bellman equation, the updating method of the Q table in this research could be

described by the following equation:

$$\text{New}Q(s, a) = (1 - \alpha)Q(s, a) + \alpha[R(s, a) + \gamma \max_{a'} Q'(s', a')] \quad (1)$$

In this equation,  $Q(s,a)$  and  $\text{New}Q(s,a)$  represent the current and the updated  $Q$  value.  $\alpha$  is the learning rate of the model. A higher learning rate means less impact from former training.  $R(s,a)$  is the reward for taking that action in that state where  $\gamma$  represents the discount rate.  $\max_{a'} Q'(s', a')$  is the maximum expected  $Q$  value for the next state. With a lower  $\gamma$  value, the agent cares more about the reward in the current step rather than the expected future rewards.

### 2.3. APPLICATIONS IN DESIGN

Markov Decision Process (MDP) is the most common model in RL. The future state of the environment is only related to the present in MDP. It is often necessary to assume the environment and the agent follow the MDP model when applying RL in the design field due to the complexity of real-world design issues. The application of RL in the architectural field mainly focused on intelligent building control strategy and design generation. For instance, Zhang et al developed a Deep RL-based HVAC control framework, and the experiment result shows that this system reduces 15% of heating energy (Zhang et al., 2018). In addition, Smith and Lasch used RL in controlling adaptive facade systems through an Intelligent Adaptive Control framework (Smith and Lasch, 2016).

In the direction of design generation, Chang et al developed a design method that combines RL, parametric object modeling, and multivariate statistical approach to get better energy performance (Chang et al., 2019). Han et al proposed a design method for urban blocks combining deep RL and computer vision (Han et al., 2020). This method could optimize the performance and aesthetics of urban blocks. Additionally, Veloso and Krishnamurti generated spatial configurations to address specific spatial goals based on double deep Q-network (DDQN) and dynamic convolutional neural-network (DCNN) (Veloso and Krishnamurti, 2020).

## 3. Reinforcement Learning-Based Generative Design Method

### 3.1. DESIGN SYSTEM

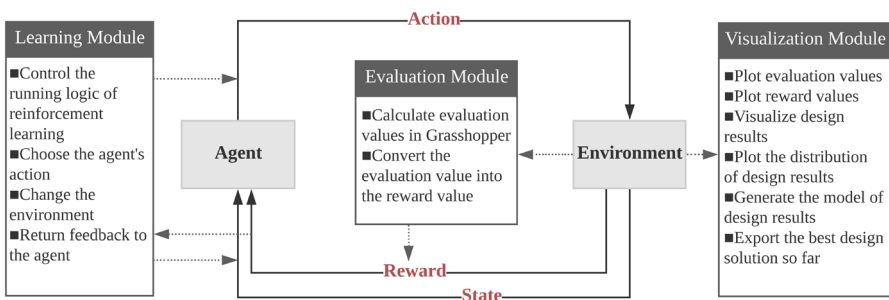


Figure 2. Design system framework

Based on the Agent-Environment framework of RL, this research proposed a design system as shown in Figure 2. Following the principle of modularity, this design system could be decomposed into three modules: Learning, Evaluation, and Visualization.

Learning module is the core of the design system and controls the logic of the RL process. Since the RL algorithm is Q-learning in this research, the learning module constructs a Q table based on the design issue. Further, the learning module chooses the agent's action and gives the corresponding feedback to the agent. After that, the learning module changes the environment based on the agent's actions.

Based on the optimization direction of this design method, the evaluation module gets the environment info from Python code and then inputs data into Grasshopper to calculate the reward reference value for corresponding optimization directions. After that, the evaluation module converts the reference value to the reward value.

The former two modules already form a workable RL system. This design system also constructed a visualization module to monitor the training process and present the design results. This module plots the change of evaluation values during the training process. In addition, this module visualizes the design results and prints the design solution with the best performance so far. It will help the designer to monitor the training process and select the final design solution.

### 3.2. DESIGN PROCESS

Figure 3 shows the detailed design process. Using Westin Peachtree Plaza as an example, the following sections describe the process from analyzing design requirements to getting the design solution.

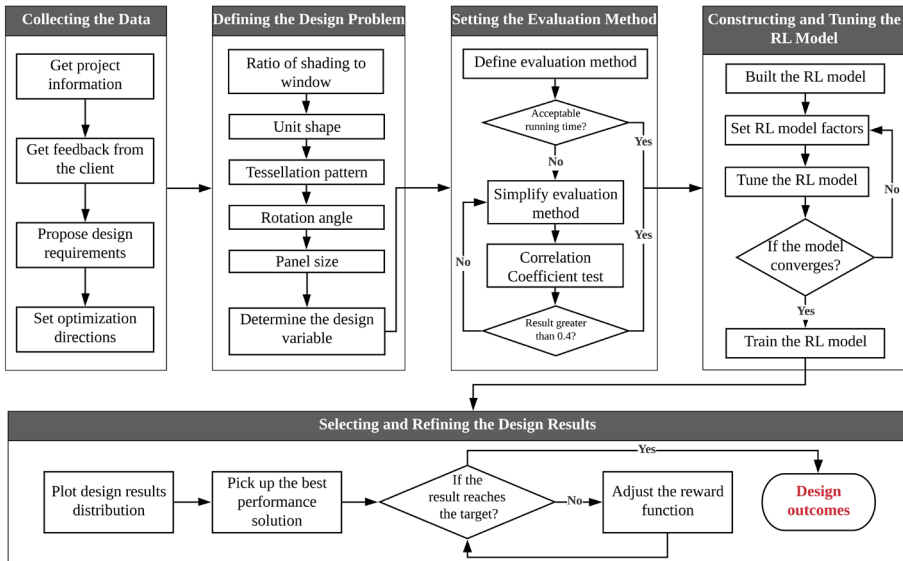


Figure 3. Design process of the RL-based design method

3.2.1. [Step 1]: Collecting the Data

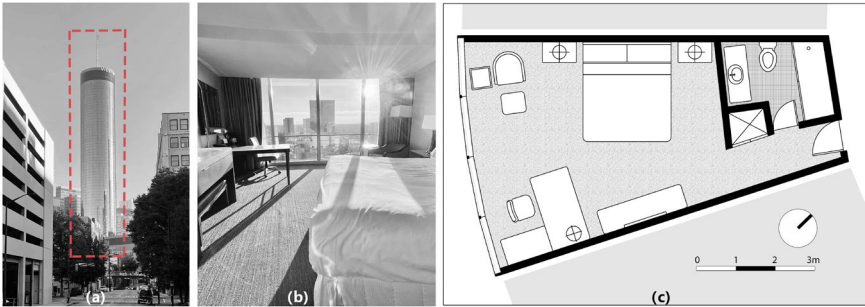


Figure 4. Project information. (a): Westin Peachtree Plaza (image taken by the author), (b): Glare in the guest room (image taken by the author), (c): Guest room plan

The first step in this design method is to collect the project information and propose design requirements. It provides a basis for the following optimization. This example is based on a guest room in Westin Peachtree Plaza which is oriented 35° west of south. It can be seen from Figure 4 that this room has a serious glare in the afternoon. So the design requirements in this example are less solar radiation and glare. Considering the shading efficiency and daylighting requirements, the kinetic facade system in this research adopts the operation mode shown in Figure 5.

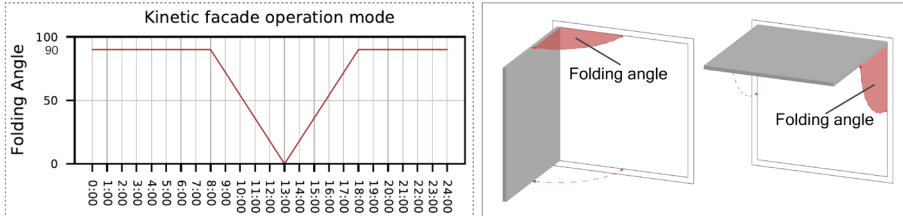


Figure 5. Kinetic facade operation mode

3.2.2. [Step 2]: Defining the Design Problem

Considering the variety of design decisions in kinetic facades design, it is hard to make all the design decisions based on RL. The design method needs to further specify the design problem for RL. As shown in Figure 6, the designer needs to make decisions from setting the covering rate of the facade panel to designing the facade pattern. Considering the form of the overall building and shading efficiency of the kinetic facade, the designer defines a kinetic facade pattern with a 100% shade to window ratio and 32 square panels. After that, the design problem in this example is how to determine the folding direction of each panel.

3.2.3. [Step 3]: Setting the Evaluation Method

Based on optimization directions of reducing the solar radiation and glare, this step needs to propose a corresponding evaluation method for these two optimization directions. The evaluation method is the basis of reward calculation, so this step is the

core step in this design method that bridges the building performance and RL. The setting of the evaluation method considers not only the accuracy but also the running time since there are massive calculations in training the RL model. To the consideration of efficiency and accuracy, this design method divides the time from 8 am to 6 pm on the summer solstice (since it has the longest daytime in a year) into ten units (each unit is one hour) and assumes the kinetic facade and solar is static in each time unit, to reduce the calculation amount.

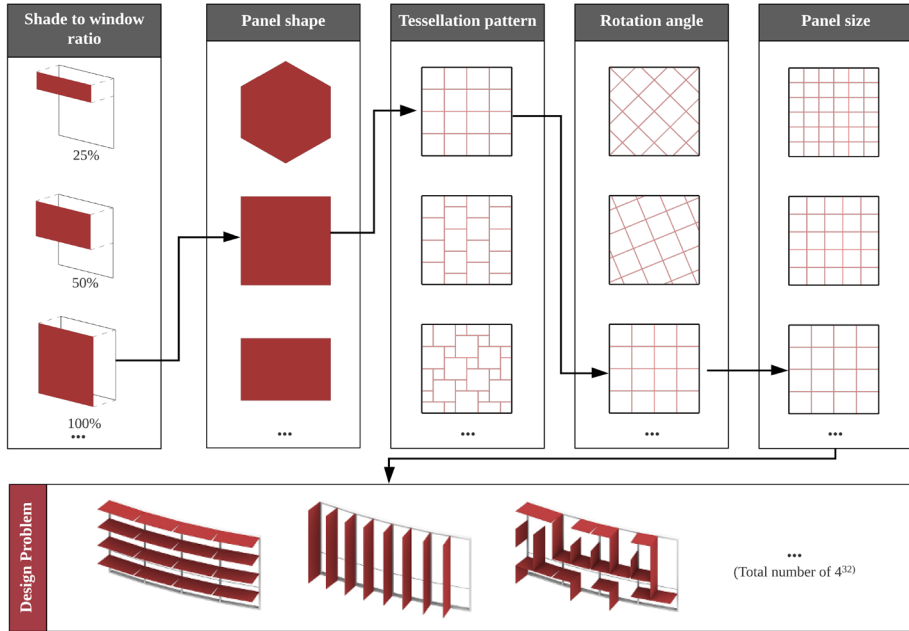


Figure 6. Design process of the facade pattern

For the solar radiation evaluation, the design method firstly calculates the position of shading panels in each time unit based on the operation mode and time (shown in Figure 7). After that, calculate unshaded areas based on the position of the shading panels. Then, project unshaded areas to the plane that is perpendicular to the solar ray and calculate the radiation heat based on the Direct Normal Irradiance (DNI) value.

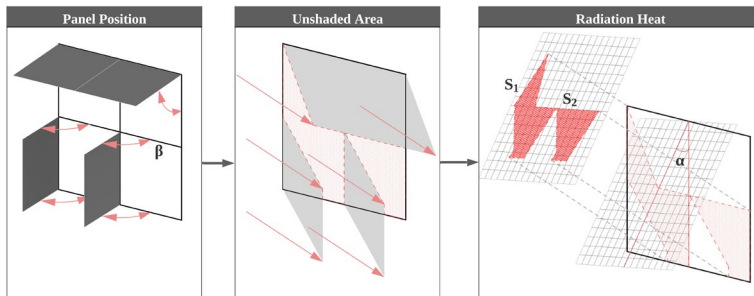


Figure 7. Solar radiation evaluation



For the glare evaluation, there is a widely accepted evaluation method for the glare in buildings called Daylight Glare Probability (DGP). However, the calculation of DGP based on simulation (such as using Honeybee plugin) takes a long time in RL training. Considering that direct sunlight will produce a serious glare in the room, this design method assumes the volume of sunlight in the room is positively correlated to the glare issue. To further test the accuracy of using sunlight volume as the glare evaluation criteria, this research constructed a Pearson Correlation Coefficient test with solar beam volumes and DGP values. This test was based on 100 randomly generated facades and the magnitude of the test in this research is 0.53 which means the solar beam volumes and the DGP values are high-degree correlated.

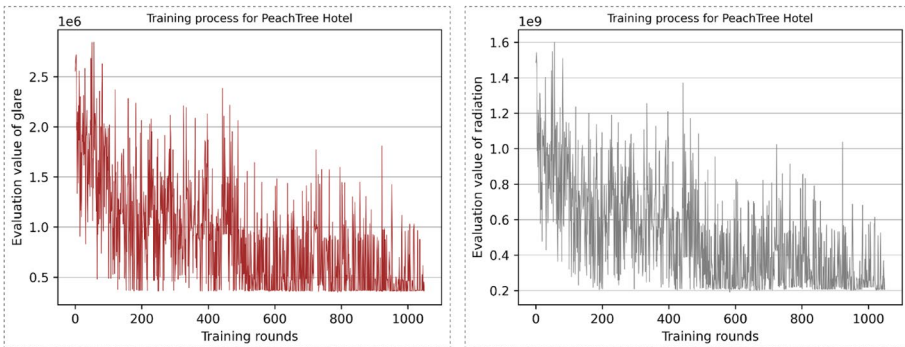


Figure 8. Training process of RL model

### 3.2.4. [Step 4]: Constructing and Tuning the RL Model

This step gathers all the information in the former steps and converts them into a feasible RL model. It starts from setting the reward calculation equation. This research uses the following equation to calculate the reward in each step.  $S_n$  and  $G_n$  are changes of radiation and glare evaluation values in each step, and  $d$  and  $g$  are factors of solar radiation and glare. For example, the reward will only be related to the glare evaluation value when  $d$  is equal to zero. The designer can use these two factors to adjust the preference of reducing solar radiation and glare in the training process.

$$R = d * S_n + g * G_n \quad (2)$$

Then, this step sets the critical parameters of the RL model and trains the RL model after several rounds of tuning. Figure 8 shows the training process of the RL model. It is worth noting that the agent has more stable performance and is more likely to get design results with less radiation and glare evaluation value (means better performance in blocking radiation and glare) with the increasing training rounds.

### 3.2.5. [Step 5]: Selecting and Refining the Design Results

In 1000 rounds of RL training, the design system generates many design solutions. This step firstly visualizes the distribution of these design results to provide a reference for the designer to select the final design result.

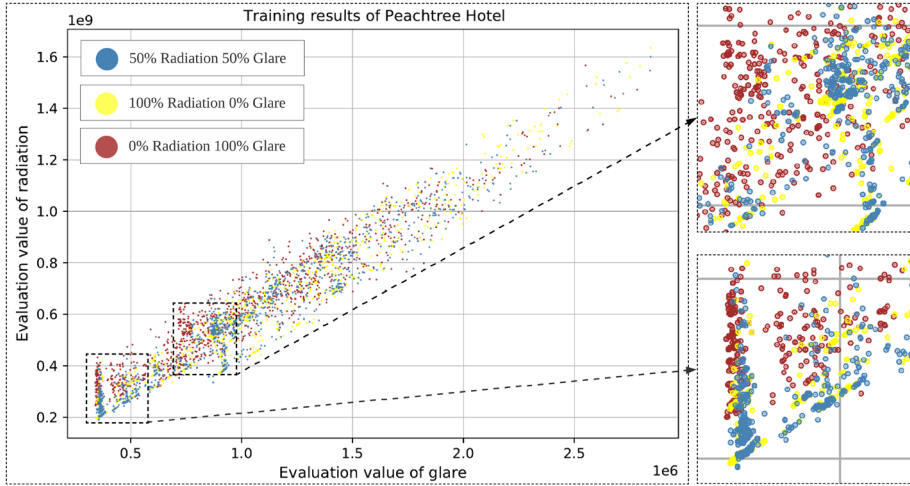


Figure 9. Distribution of design results

By changing the value of  $d$  and  $g$  in the former equation, the designer can redo the training and get new design results with the same system. For example, Figure 9 shows the distribution of design results with different compositions of rewards. A point closer to the x-axis means better performance in reducing radiation, and a point closer to the y-axis means better performance in blocking glare. In this example, the designer finally selected the design result closest to the origin, and Figure 10 shows the corresponding facade rendering at different times of a day.

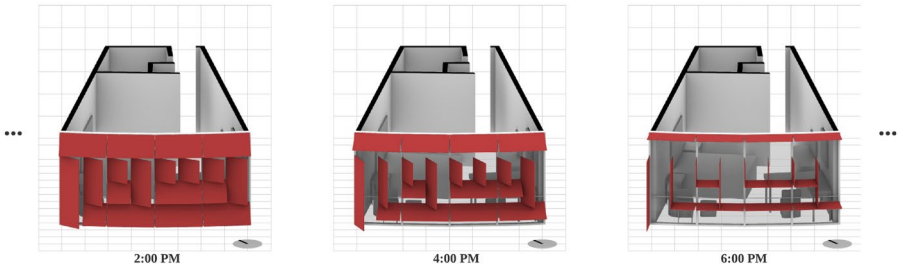


Figure 10. States of the final design result at different times of a day

#### 4. Experiment

To further verify the effectiveness of the RL-based design method, this research proposed an experiment with more sampling data and higher sampling density. This experiment is based on the real data of the Westin Peachtree Plaza guest room and the weather data in Atlanta. The independent variable in this experiment is the kinetic facade design. The dependent variables include the Direct Solar Radiation and Daylight Glare Probability (DGP).

This experiment tested the performance of different kinetic facade designs (include pure horizontal facade, pure vertical facade, RL generated facade, randomly generated

facade, and the original none shading facade as a base line) in 91 days from 5/7 to 8/5 (summer solstice is the midpoint) with a ten minutes time unit. The calculation of the solar radiation experiment is based on geometric calculation functions of Grasshopper in Rhino. In addition, the glare experiment is based on the Honeybee plugin in Grasshopper. Figure 11 shows the results for the solar radiation experiment and glare experiment.

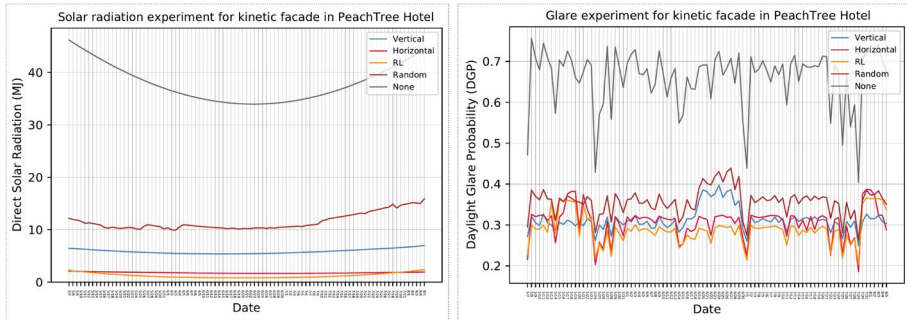


Figure 11. Experiment results for kinetic facades in Westin Peachtree Plaza

### 5. Discussion

As shown in Figure 12, this study analyzed the distribution of experiment results using boxplots. Further, it investigated the cumulative experiment results using bar charts. It is worth noting that the RL generated facade has the best performance in both the solar radiation and the glare experiment with the hybrid vertical and horizontal panels. Compared with the pure vertical facade, the RL generated facade reduced the cumulative solar radiation by 78.3% and the cumulative DGP by 7.1%. In addition, the RL generated facade reduced the solar radiation by 29.7% and DGP by 5.1% compared with the pure horizontal facade.

The rule of thumb in designing shading devices is that the horizontal louver is good at blocking solar radiation and the vertical fin is good at blocking glare, especially in west orientation. Experiment results show that the pure horizontal kinetic facade performs well in blocking solar radiation and glare in a Southwest orientation. In addition, the pure vertical kinetic facade has good performance in blocking glare but is not so good in blocking solar radiation.

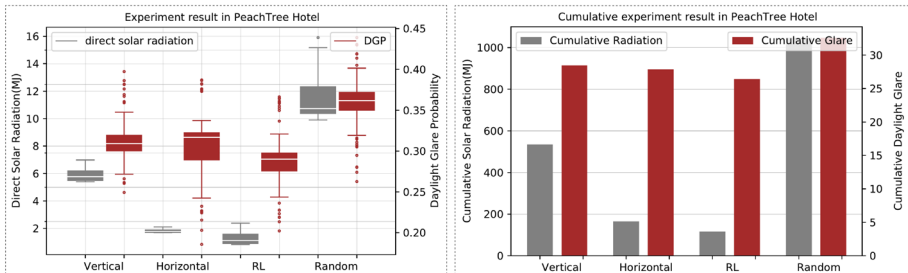


Figure 12. Left: Distribution of experiment results. Right: Cumulative experiment results

## 6. Contribution

The experiment results show that the kinetic facade generated by the novel design method has better performance than the facades generated by traditional design methods. Future research will further compare this RL-based design method with model-based design methods. Changing states to respond to different environments and design objectives, this method revealed more possibilities for adapting to future challenges reducing life-cycle carbon emissions.

We assert that this method improves the performance of kinetic facades in early design stages which could increase the efficiency of buildings with minimal energy consumption. It will further promote the realization of the "Sustainable Cities and Communities" described in SDG11. In addition, the design method based on Python and Grasshopper could reduce the learning curve for designers, which will help the spread of RL-based design methods. This research expands the field of generative design but also provides reference and inspiration for the application of RL in design.

## References

- Abdulmajid K. & Ethan K. (2015). Innovations in dynamic architecture. *Journal of Facade Design and Engineering*, 3(2):185–221.
- Hosseini, S. M., Mohammadi, M., Rosemann, A., Schröder, T., & Lichtenberg, J. (2019). A morphological approach for kinetic façade design process to improve visual and thermal comfort. *Building and Environment*, 153, 186-204.
- Mahmoud, A. H. A. & Elghazi, Y. (2016). Parametric-based designs for kinetic facades to optimize daylight performance: Comparing rotation and translation kinetic motion for hexagonal facade patterns. *Solar Energy*, 126, 111-127.
- Pedro V. & Ramesh K. (2020). An academy of spatial agents generating spatial configurations with deep reinforcement learning. *Proceedings of eCAADe 2020*.
- Richard S.S. & Andrew G.B. (2018). *Reinforcement learning: An introduction*. MIT press.
- Roel CGM L., Marja T., Daniel C., & Jan LM H. (2013). Climate adaptive building shells: State-of-the-art and future challenges. *Renewable and sustainable energy reviews*, 25:483–493.
- Shady A., Senem B., Taha S., Christian S., Roel L. & Francesco G. (2018). Current trends and future challenges in the performance assessment of adaptive facade systems. *Energy and Buildings*, 179:165–182.
- Smith, S., & Lasch, C. (2016). Machine Learning Integration for Adaptive Building Envelopes. *Proceedings of ACADIA 2016*.
- Soowon C., Nirvik S., Daniel C., & Perry Y. (2019). Multivariate relationships between campus design parameters and energy performance using reinforcement learning and parametric modeling. *Applied energy*, 249:253–264.
- Stuart J R. & Peter N. (2016). *Artificial intelligence: a modern approach. Malaysia*. Pearson Education Limited.
- Zhen H., Wei Y., & Gang L. (2020). A performance-based urban block generative design using deep reinforcement learning and computer vision. In *The International Conference on Computational Design and Robotic Fabrication*, pages 134–143. Springer.
- Zhiang Z., Adrian C., Yuqi P., Chenlu Z., Siliang L., & Khee P. L. (2018). A deep reinforcement learning approach to using whole building energy model for hvac optimal control. In *2018 Building Performance Analysis Conference and SimBuild*, volume 3, pages 22–23.

# PARASITE CITY: RETAINING THE INDUSTRIAL DISTRICT OF ALEXANDRIA, SYDNEY AS AN INTEGRAL PART OF URBAN REGENERATION

GUOYI CHEN<sup>1</sup>, SEUNGCHEOL CHOI<sup>2</sup>, MOHAMMED MAKKI<sup>3</sup>  
and JORDAN MATHERS<sup>4</sup>

<sup>1,2,3</sup>*University of Technology Sydney*

<sup>4</sup> *SJB*

<sup>1</sup>*guoyi.chen-1@student.uts.edu.au, 0000-0002-9985-7359*

<sup>2</sup>*seungcheol.choi@student.uts.edu.au*

<sup>3</sup>*mohammed.makki@uts.edu.au, 0000-0002-8338-6134*

<sup>4</sup>*jmathers@sjb.com.au, 0000-0003-3803-2131*

**Abstract.** Industrial lands are the most vulnerable urban typologies in areas undergoing urban regeneration. They are considered less adaptive to integrated residential typologies, and their legacies are threatened under fast gentrification. The goal of this paper is to explore a sustainable strategy to address the conflict between urban sprawl and industrial conservation in Alexandria, Sydney. Through the application of a sequential evolutionary simulation, the presented research proposes a potential mixed-use scheme to rejuvenate the existing industrial district of Alexandria in an integrative manner without necessitating its destruction. This paper provides a prototype of urban regeneration, optimised by a multi-objective evolutionary algorithm, that demonstrates the necessity of industrial integration in the pursuit of true mixed use urban typologies.

**Keywords.** Gentrification; Mixed-use; Urban Development; Sequential evolutionary simulation; SDG 9; SDG 10; SDG 11; SDG 12.

## 1. Introduction

The stresses imposed on existing cities to sustain the demand of population growth is ever-growing. Where the development of new cities attempts to respond to this demand; existing cities face unprecedented challenges of accommodating increased density within a finite space. One such city is Sydney, Australia; with a population expected to grow by 30% in the next 25 years (Greater Sydney Commission, 2018); the city's suburbs are experiencing a continuous cycle of urban regeneration, in which the old is replaced with the new. Much of this regeneration is through the development of mid to high-rise apartment towers in place of low density residential and brownfield sites (Newton et al., 2021). One such brownfield redevelopment area is the suburb of Alexandria which houses an industrial district that has existed for generations, historically serving as one of the city's largest industrial anchors due to its proximity to the airport and the central business district.



Figure 1. Existing Condition Analysis of Alexandria

To support Sydney's growing population, Alexandria has been identified as a key area for urban redevelopment, transitioning towards a high-density residential area over the next 30 years. Inevitably, to meet the demands for residential developments, Alexandria is threatened from going through a gentrification process, in which the industrial districts which have existed for the past century are to be erased to pave the way for new builds. The presented research (academic and industry collaboration) proposes an alternative urban regeneration strategy for the suburb of Alexandria that aims to retain the existing industrial district, minimise demolition and redevelopment within the site, and seek a sustainable way to allow the suburb to support growing numbers in the population. Additionally, through retaining industrial legacy and integrating it within urban regeneration, the potential environmental impact (in response to United Nations Sustainability Development Goals (SDGs)) on urban growth is significant. The research explores how the retention of industry avoids the negative impact of transportation times between industry and the community (SDG 9 and SDG 11); the existing location of industry, specifically with relation to the airport and the CBD, is optimal for maintaining primary transport lines, thus reducing vehicular carbon emissions (SDG 12); and finally, industry integration allowing for greater equity through a varied urban demographic (SDG 10).

Through the application of an evolutionary generative process, the paper proposes an urban growth strategy that maintains the urban structure of the existing and uses it as a basis for the development of new mixed-use typologies that 'parasitically' integrate within the suburb's existing morphology. The paper also builds on recent research (Randall et al., 2020) for the application of sequential evolutionary simulations to tackle complex design problems consisting of a high number of fitness objectives. The simulations address issues of network and spatial organisation, followed by the morphological distribution and relationships of various mixed-use typologies that leverage the existing urban fabric as the basis for the regeneration strategy.

## 2. Background and Context

### 2.1. IMPACT OF URBAN RE-DEVELOPMENT ON INDUSTRY

Urban re-development has left industrial zones vulnerable to demolition and relocation. Throughout Europe, space required for urban growth has overwhelmingly consumed land used for industrial and manufacturing services, creating an increased urban division between industry and 'mixed use', leading to a decrease in working opportunities and a negative impact on economic growth. The separation of industry from the city reduces the benefits from sustainable energy use, waste utilisation, creation of micro economic local networks and improved synergies by different urban typologies (Haselsteiner, 2020). Initiatives to integrate Industry within mixed use development (and thus creating true mixed-use typologies) have resulted in a renewed understanding of 'urban production', in which industry plays an integral role within the urban fabric and not decoupled from it (Cotter, 2012).

### 2.2. ALEXANDRIA INDUSTRIAL HISTORICAL BACKGROUND

Since the mid-19th century, industrial facilities have considerably contributed to Sydney's urban growth and technological advancement. During the 1920s-30s, Sydney's thriving industry expanded to the city's nearby suburbs, including Alexandria. Due to the cheaper land value, isolation from the population, and close proximity to the city and recently built Sydney Airport, various manufacturing plants were concentrated in this suburb. By 1943, more than 500 factories had occupied Alexandria, the peak for the suburb, with the post-war period marking a boom of Australia's industrial history. This has shaped Alexandria into a symbol of Australia's industrial past; informing the urban fabric through oversized property lots, broad roads and a multitude of car parks (City of Sydney 2014).

### 2.3. GENTRIFICATION IN ALEXANDRIA

Given the decline of secondary industry in the 1970s, the former industrial model failed to adapt to the shifting social pattern in Alexandria. Its industrial urban fabric started to vanish (Karskens and Rogowsky, 2004). The industrial facilities have since been suffering an inevitable decline. Meanwhile, the 'Inner city redevelopment plan' by the City of Sydney aims for the major re-development of several inner-city suburbs in Sydney, one of which is Alexandria, towards high-density residential suburbs over the next 30 years. As one of the closest suburbs to the Sydney CBD, Alexandria is experiencing an increasing demand for residential property due to the increase of Sydney's urban population. Clover Moore, the Mayor of Sydney as of writing, stated that "[The] Alexandria industrial site was chosen for its proximity to residential areas and public transport" (Gorrey, 2018). Thus, the residential area aims to propel local commercial, cultural, and recreational events. Alexandria's overwhelming gentrification, represented by the swell of land value and the rapidly growing number of apartments, is threatening the local industrial legacy that has been existing for the past century.

## 2.4. WHY CONSERVE THE INDUSTRIAL LEGACY

According to The Burra Charter (2013), Australian identity and experience can be illustrated by the place of cultural significance. The local industrial urban fabric is a record of Alexandria's history that archives the prosperity of manufacturing in Australia. Even today, many of those industrial facilities are still indispensable parts to the city, as they are still contributing to Sydney's urban services and logistics and providing a large number of working opportunities. By 2018, industrial land occupied 8% of Sydney however made up 19% of jobs (Hill 2018). Considering the irreplaceable value and current contribution of those industrial facilities, it is urgent to find a balance between the local housing demands, urban growth and legacy conservation according to the "principle of intergenerational equality" (Burra Charter 2013). Inspired by the ethos of "never demolish" advocated by Lacaton & Vassal, the presented research examines an alternative architectural solution to conserve Alexandria's industrial legacy.

## 2.5. THE POTENTIAL FOR INDUSTRIAL MIXED-USE

Industrial spaces are usually associated with urban characteristics (noise and pollution) not conducive to residential spaces. However, according to Howells and Openshaw (2021), this association has slowly changed through new innovations around sustainable materials, cleaner emissions and advanced technologies. Furthermore, new paradigms of what constitutes 'industrial' has been rapidly changing in recent years. Small businesses that support online marketplaces and homemade goods are emerging. These new aspects of industry are more cohesive with other uses, helping support a more diverse characteristic of mixed-use developments. This enables the possibility for a mixed-use development to be utilised as a hybridisation strategy between local industry and current residential, commercial and cultural demands. Whereas industry is not typically a component of mixed-use developments, this research suggests an alternate strategy that allows industry to remain, and be supported by various other uses to create a truly mixed-use urban fabric. A place that supports social sustainability, local business, economic growth and the culture of Sydney's industrial legacy.

# 3. Method

## 3.1. EVOLUTIONARY SEQUENTIAL SIMULATIONS

The experiment presented utilises sequential multi-objective evolutionary simulations using Wallacei X (Makki. 2018). Each simulation examines the site at a different scale and optimises for various fitness objectives accordingly. The aim of the simulations is to address the complexity of the site without the need to abstract the formulation of the design problem, allowing for each simulation to tackle a unique set of fitness functions at different scales, in which the output of the first simulation (network and plot divisions) translates as the input for the second simulation (block typology and programmatic distribution). Additionally, to achieve a better understanding of Alexandria's urban network, accessibility and spatial configurations, Space Syntax theory (Hillier and Hanson, 1984) is integrated within the simulation as both an analytic and generative tool to better understand the spatial and network distribution of the site,



as well as develop solutions that respond to integration, choice and centrality.

The pseudo code (Figure 2) presents both evolutionary simulations and their relationship to one another. Sequence 1 analyses the existing infrastructure of the site and its surrounding context. In order to minimise demolition and redevelopment within the site, industrial activities are to be retained through maintaining the industrial footprint, while introducing infrastructure above the existing built forms. The main objective of this sequence is to evolve a solution set that minimises interruption to the existing industrial activity, while maximising pedestrian integration within the existing circulation. Sequence 02 assigns new built forms above the existing industrial buildings based on the resulting phenotypes generated by the first sequence. This sequence analyses the phenotype and each program’s proximity to each other through a thorough analysis of the relationship of various mixed-use typologies and their impact on one another.

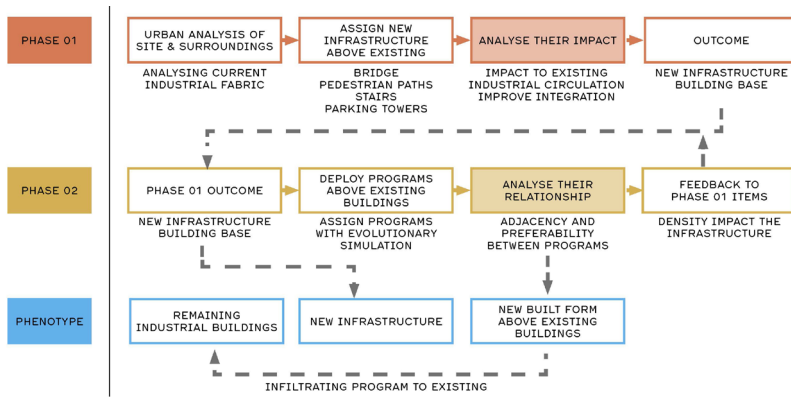


Figure 2. Pseudo Code (Evolutionary Workflow Diagram)

## 4. Experiment Setup

### 4.1. SEQUENCE 01

#### 4.1.1. Construction of the Superblock and Fitness Objectives

Figure 3 presents the various parameters that contribute to the construction of the phenotype for sequence 01. Various plots within the site are identified for the integration of new programmatic functions or parking towers, serving both existing industry as well as added typologies. The impact of the parking towers on the existing industrial vehicular circulation is assessed and reconfigured accordingly. Rooftops with an area above a certain threshold (in this case 1000sq.m) are identified and are divided and connected via a network of pedestrian paths and bridges. Finally, the upper-level network is connected to the ground level network through a series of vertical access points that maximise the integration value of both levels. The fitness objectives being optimised in the algorithm aim to maximise the integration between the existing network and the newly added network, while also optimising the distribution of parking spaces, bridges and pedestrian paths (Figure 4).



Figure 3. Construction of the Sequence 01 Parametric Definition

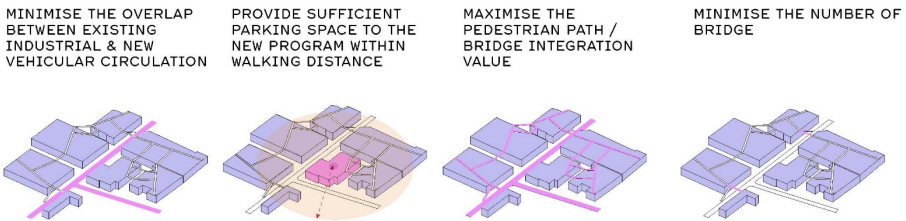


Figure 4. Fitness Objective of Sequence 01

## 4.2. SEQUENCE 02

### 4.2.1. Construction of the Superblock and Fitness Objectives

Sequence 02 uses the output from sequence 01 as the base phenotype and deploys mixed use block typologies that respond to the network generated in the first sequence. Rooftops are subdivided into building plots based on size and orientation, each one attributed with a use and associated typology that responds to the typology of its neighbours. This is informed by various parameters such as building height, proximity to parking towers and programmatic relationships at multiple scales (Figure 5). The fitness objectives identified and integrated within the algorithm for this sequence examine programmatic distribution and its influence on each building block defined in the urban fabric (Figure 6).

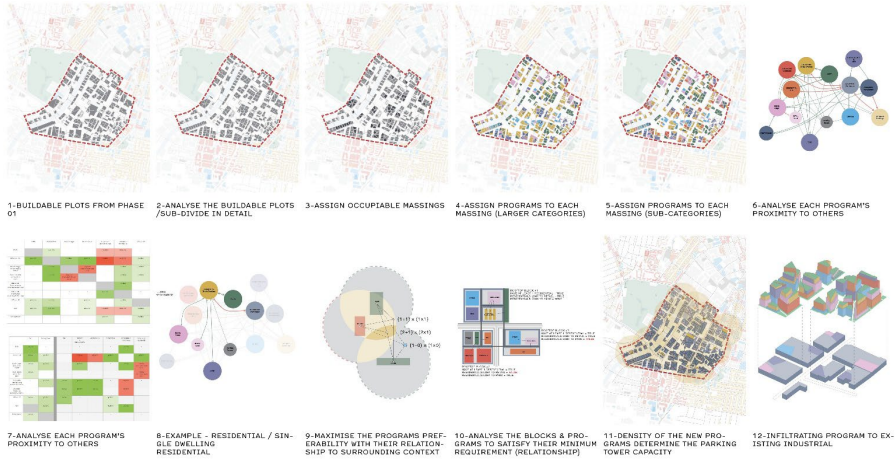


Figure 5. Construction of Sequence 02 Parametric Definition

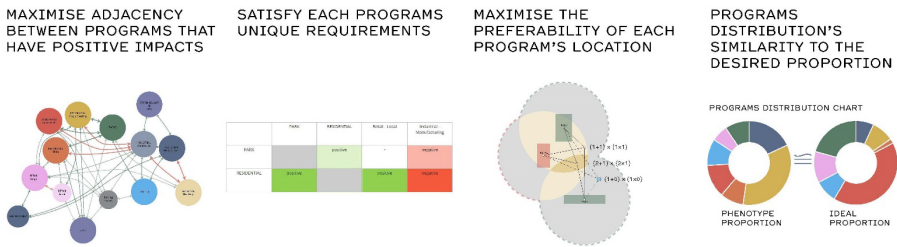


Figure 6. Fitness Objective of Sequence 02

## 5. Experiment Results & Selection

### 5.1. SEQUENCE 01

#### 5.1.1. Results and selection

The algorithm for sequence 01 generated a population of 1000 solutions (generation size of 20 and generation count of 50). This is primarily due to the large simulation runtimes resulting from the integration of space syntax analysis methods within the algorithm. From the 1000 solutions, 116 solutions formed the pareto front (the most non-dominated solutions in the population) which were selected for further analysis for the purposes of selection. The pareto front solutions were clustered (using hierarchical clustering) with a K-value of 9, with each cluster centre analysed against additional criteria to inform selection. The matrix presented in Figure 07 presents the performance of each solution against several metrics, including the distribution of parking towers, bridges, vertical circulation and the integration value of the street network (vehicular and pedestrian). Through ranking the cluster centres against these metrics, the top 3 performing solutions were selected for further detailed analysis, in which they were evaluated according to the key objectives of sequence 01, which is to improve the integration value of pedestrian paths and bridges, as well as minimise the disruption of

existing vehicular networks. The results identified solution ‘generation 11 individual 17’ as the most optimal against the evaluation criteria utilised (Figure 8). This solution is extracted as the primary input for sequence 01.

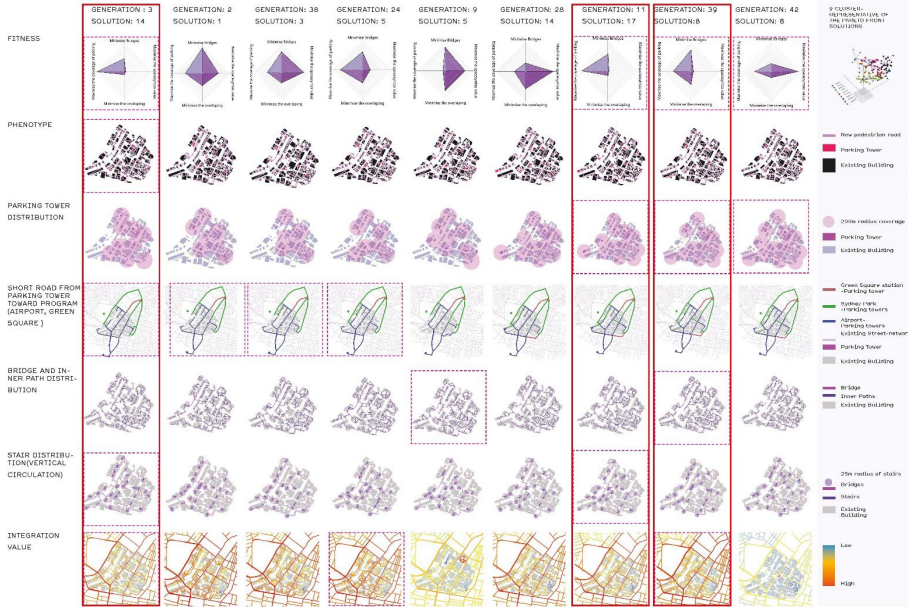


Figure 7. Simulation Matrix for Sequence 01' Selection

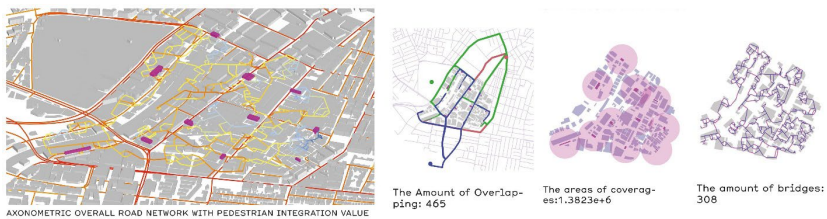


Figure 8. Selected Solution from Simulation 01 (Gen. 11 Ind. 17)

## 5.2. SEQUENCE 02

### 5.2.1. Results and selection

A lower runtime for sequence 02 allowed for a larger population size of 5000 solutions (generation size of 50 and generation count of 100), generating a pareto front of 172. A similar clustering approach is also applied in sequence 02 in which the pareto front is clustered into 16 clusters. Each cluster centre was assessed according to their fitness values (using a diamond chart) and 4 solutions were selected for further analysis according to their performance across all objectives. Each of the four solutions is analysed against additional metrics focusing on programmatic distribution and relationships, in which the proximity of different program types is assessed and

evaluated according to the degree of mixed-use typologies within each block. The results from the analysis are presented in Figure 9. The selected solution that most optimally met the programmatic requirements and relationships of the space was ‘generation 82\_individual 04’ (Figures 10 and 11).

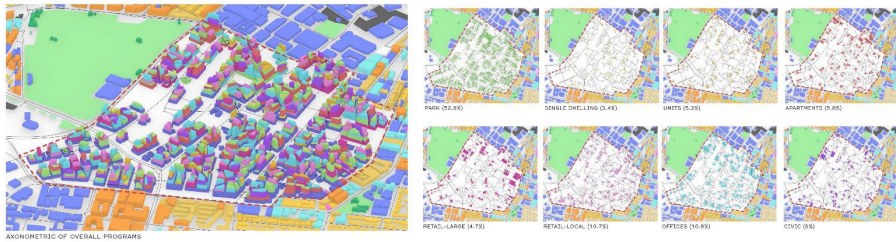


Figure 09. Selected Solution from Simulation 02 (Gen. 82 Ind. 04)



Figure 10. Existing Section (Top) and New Section (Down)

## 6. Conclusions/Discussions/Future Work



Figure 11. Final Axonometric (left) and Renders (right) of the Proposal.

The mix-use strategy presented above examines the potential advantages for integrating industry as part of mixed use urban fabrics. In suburbs such as Alexandria, where industry holds a historical significance, its integration within urban redevelopment initiatives is imperative to the retention of historical urban value. The presented experiment illustrates the potential integration of industry within the urban fabric through a multi-level urban model, in which industry is retained on ground level, while new urban development is integrated on upper levels, maintaining network and

spatial relationships between the activities taking place between the two. The use of sequential evolutionary simulations avoids the necessity to simplify the design problem, allowing each sequence to tackle different scales and typologies (network in one and block in another). However, as the process is sequential, the decisions and results applied in the first sequence holds greater significance to the results of the second sequence, as such great attention must be given to the formulation of the phenotype and chromosomes of the first sequence.

Future work that addresses various aspects of demographic structure and social relationships is required. Moreover, the growth and future adaptation of the redeveloped superblock is critical; questions that address the temporal scale of integrating industry with mixed use development, and its impact on growth patterns, requires further interrogation.

## References

- City of Sydney. (2014). *Industrial and warehouse buildings heritage study* [Ebook] (pp. 22-24). Retrieved 20 November 2021, from <https://www.cityofsydney.nsw.gov.au/heritage-guidelines-studies/industrial-sydney-industrial-warehouse-buildings-heritage-study>.
- City of Sydney. (2020). *City Plan 2036* [Ebook] (pp. 11-74). Retrieved 12 December 2021, from <https://www.cityofsydney.nsw.gov.au/planning-for-the-future>.
- Cotter, D. (2012). *Putting Atlanta back to work: Integrating light industry into mixed-use urban development*. Atlanta, Georgia: Georgia Tech Enterprise Innovation Institute.
- Gorrey, M. (2018). *Sydney's next late-night hot spot: Alexandria*. Smh.com.au. Retrieved 17 November 2021, from <https://www.smh.com.au/national/nsw/smash-repairers-and-gyms-is-this-sydney-s-next-late-night-hotspot-20181106-p50ebt.html>.
- Greater Sydney Commission. (2018). *A Metropolis of Three Cities-connecting people* [Ebook] (pp. 1-29). Retrieved 12 December 2021, from <https://www.planning.nsw.gov.au/Plans-for-your-area/A-Metropolis-of-Three-Cities>
- Haselsteiner, E., Grob, L. M., Frey, H., Madner, V., Laa, B., & Schwaigerlehner, K. (2020). The Vertical Urban Factory as a Concept for Mixed Use in Future Cities. *Shaping Urban Change—Livable City Regions for the 21st Century*, Proceedings of the REAL CORP.
- Hillier, B., & Hanson, J. (1984). *The Social Logic of Space*. Cambridge University Press.
- Hill, S. (2018). *Why Sydney needs to keep its sheds and industrial lands*. The Fifth Estate. Retrieved 7 December 2021, from <https://thefifthestate.com.au/columns/spinifex/why-sydney-needs-to-keep-its-sheds-and-industrial-lands/>.
- Howells, J., & Openshaw, G. (2021). *Why mixing residential and light industrial is good for our cities*, AECOM. Retrieved 12 December 2021, from <https://aecom.com/without-limits/article/why-mixing-residential-and-light-industrial-is-good-for-our-cities/>.
- ICOMOS Australia. (2013). *Burra Charter* (4th ed., pp.3).
- Karskens, G., & Rogowsky, M. (2004). *Histories of Green Square*. UNSW Printing
- Makki, M., Showkatbakhsh, M. and Song, Y. (2018), Wallacei: An evolutionary and Analytic Engine for Grasshopper 3D, Wallacei, [www.wallacei.com](http://www.wallacei.com)
- Newton, P., Newman, P., Glackin, S., & Thomson, G. (2021). *Greening the Greyfields New Models for Regenerating the Middle Suburbs of Low-Density Cities* (pp.1-49). Palgrave macmillan.
- Randall, M., Kordrostami, T., & Makki, M. (2020). The Taikoo Shing Superblock: Addressing urban stresses through sequential evolutionary simulations. In *25th International Conference on Computer-Aided Architectural Design Research in Asia: RE: Anthropocene, Design in the Age of Humans, CAADRIA 2020* (pp. 415-424). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA).

# A GENERATIVE DESIGN APPROACH TO URBAN SUSTAINABILITY RATING SYSTEMS DURING EARLY-STAGE PLANNING

OR MOSCOVITZ<sup>1</sup> and SHANY BARATH<sup>2</sup>

<sup>1,2</sup>*Technion, Israel Institute of Technology.*

<sup>1</sup>*or.moscovitz@campus.technion.ac.il, 0000-0001-7262-7302*

<sup>2</sup>*barathshany@technion.ac.il, 0000-0003-0776-7389*

**Abstract.** Sustainability rating systems (SRS) aim to guide decision-makers in the planning process by defining clear guidelines and metrics. Nowadays, this process usually requires further tasks and the involvement of multiple professional advisors that eventually increase planning complexity and lead to lower SRS implementation. In this paper, we explore generative urban models and multi-objective optimization of SRS metrics to potentially enhance SRS use in planning processes. Furthermore, we apply this framework to a case study that has not reached its SRS planning goals due to contradicting trade-offs between municipal and stakeholder objectives. The urban model reflects the stakeholder design requirements and constraints such as the desired floor area ratio (FAR), building types, and units' number while the SRS metrics act as optimization goals. As part of the process, we automate quantitative indicators from Israel SRS '360 Neighbourhood' to use them as optimization goals and to analyse their correlation and trade-offs. Through this process, we enable a generative exploration of high-performing design iterations relative to a chosen set of SRS goals. Such a framework can enhance the integration of verified sustainability goals in the planning process, thus informing the stakeholders of their decision trade-off's concerning SRS indicators in urban development.

**Keywords.** Sustainability Rating Systems; Generative Design; Multi-objective optimization; Urban Modelling and Simulation; SDG 11.

## 1. Introduction

Urban development is becoming increasingly complex and demanding concerning rapid urbanization. Increased building activity is needed to meet the demands of anticipated population growth, adding significantly to the existing challenges of achieving sustainable urban environments. Sustainability rating systems (SRS) play a critical role in meeting these challenges and achieving UN sustainable development goals for sustainable cities and communities. This paper investigates computational optimization techniques to enhance SRS use, therefore, increasing positive impact on sustainability. Sustainability rating tools primarily serve for the evaluation of buildings. The rapid growth of cities and the challenge to assess the built urban environment concerning sustainability benchmarks have focused research on developing tools and

assessment frameworks for urban design (Smith et al., 2016). Various methods and tools emerged in the search for a sustainable city that allows projects to display their environmental, economic, and social benefits to the local community in different planning stages of their development processes. These tools consist of frameworks with several indicators grouped into categories. While assessing and ranking the sustainability of urban developments, the tools also guide and encourage the design of sustainable informed, and high-performing communities throughout the planning process (Castanheira et al., 2014). The most established urban assessment methods are LEED-ND, BREEAM, CASBEE, and Green Star (Aspinall et al., 2012). Currently, SRS aims to help decision-makers in the planning process by defining clear benchmarks and guidelines. Yet, they also add further tasks, planning time, and the need to involve multiple professional advisors that eventually increase planning complexity and lead to lower SRS implementation (Yoffe et al., 2020).

Furthermore, During the planning process, sustainable, economic and social goals often contradict, and planning a scenario that demonstrates good trade-offs between those goals is challenging (Nagy et al., 2018). SRS metrics are built from both qualitative indicators and quantitative indicators. While qualitative indicators can be part of an administrative process or rely on expert knowledge, the quantitative indicators are analysed as numerical equations and can potentially be automated. SRS automation may save manual effort, time, and resources. In addition, it allows the integration of SRS in a generative design process as optimization goals using multi-objective optimization (MOO). Integrating LEED indicators in a generative process has been explored in the building scale (McGlashan et al., 2021). We expand the research to an urban scale that integrates municipal constraints with SRS goals. The proposed workflow can contribute to the framework of SRS due to its immediate and responsive qualities. It allows stakeholders to understand each design decision consequence, thereby enhancing informed decisions throughout urban planning evaluation, thus increasing process productivity. Another gap being addressed concerns the local SRS in Israel '360 Neighbourhood' in which each indicator is currently evaluated separately with no formal correlation. Here, we explore indicators correlation, revealing their relationships in the planning process and whether they contradict or support planning requirements. The proposed computational workflow allows the exploration of multiple iterations through high-performing design solutions relative to a chosen set of SRS goals. Applying such workflows at an urban scale enhances the integration of verified sustainability goals in the planning process and its potential correlation with the multiple stakeholders involved in the planning process.

### 1.1. RELATED WORK

Recent advancements in the tools available for designers pose new opportunities for measuring urban design performance. Furthermore, the increasing availability of tools and the reduction of computation time needed for analysis make performance indicators suitable for an optimization process that takes heavy computation resources (Natanian and Auer, 2020). Multiple procedural modeling techniques have been explored to assist urban design while saving time and resources (Koenig et al., 2019, Schmitt et al., 2008). Moreover, such models can be integrated with a MOO process to benefit urban design, using multiple inputs and metrics, showing correlation and



analysis of the design (Wilson et al., 2019). In urban design practice, such processes have been examined, optimizing a multi-block cluster for profitability and solar energy generation while maintaining developers' requirements and design constraints (Nagy et al., 2018). Design optimization research has introduced the RBFMOpt algorithm, a novel optimizer that includes a learning algorithm that constructs surrogate models as it runs to predict simulation results (Wortmann & Fischer, 2020). At the same time, integrated computational frameworks are explored to measure urban sustainability, using machine learning predictions integrating social, environmental, and economic metrics (Koenig et al., 2021). However, Not much research has integrated urban certification sustainability systems as part of a generative framework . In landscape design, a workflow for evaluating the performance of urban landscape ecological indicators in line with sustainability rating systems has been developed. The study uses Grasshopper and Python to translate the criteria into quantitative spatial metrics and demonstrate optimized biomass measurement (Yoffe et al., 2020). At the building scale, McGlashen also automates several metrics from SRS within a design framework and demonstrates how various goals and trade-offs can be optimized by a generative design procedure that seeks to improve certification scores and reveal indicators' relationships (McGlashen et al., 2021). In this context, it is our aim to develop such implementation of SRS within a generative design system at an urban scale.

## 2. Methodology

This paper describes the application of automated SRS indicators as part of a multi-objective methodology at an urban scale. A case study was selected for an urban renewal project on an existing 24,000 sqm area in Holon, Israel. The project requires an expansion from 276 residential units to 1000 units and a public school of 6500 sqm. The project was in an early-stage test-fit, having two early plans made. Together with the developer, we examined possibilities to integrate the SRS indicators score while maintaining the profitability and design principles of the existing plan as defined by the developer. First, we developed a design space that could yield different design scenarios that correspond to the municipal constraints of the planning context and answer the developer's requirements. The second stage automated the SRS indicators and evaluated each design option's sustainability and profitability performance relative to the project's goals. The third phase employs the SRS metrics used as objective goals during an optimization process.

### 2.1 THE DESIGN SPACE MODEL

A procedural model was created to generate a broad mixture of design options based on input parameters and variables that consider the constraints and requirements of the project. The program constraints define rules determined by local municipality regulations, while program requirements are the project's programmatic goals. The chosen case study featured a program that required a thousand residential units adjusted in three



Figure 1. Previous site plan

different building typologies: three to five towers and row or 'L' type low buildings. Program constraints included a predefined site boundary, a maximum height of forty floors for towers and nine floors for low buildings, a defined density and floor area ratio (FAR) of 5.6 and an addition of 6500 sqm for a public-school area. Finally, we evaluated the developers' previous plans (Figure 1) and integrated the design guidelines and typologies in our procedural model.

## 2.2. PROCEDURAL GEOMETRY GENERATION PROCESS

The procedural model parameters are based on dependant relationships and rules. The project brief which outlines design constraints and municipal requirements informs the selection of the parameters. Some parameters are defined by the designer as manual settings, while others function as dynamic variables and constantly change during the optimization process (Figure 2). Rhinoceros 3D and various parametric plug-in's in the grasshopper environment are used as they allow this kind of urban procedural modeling and provide many capabilities in urban design. (Koenig et al., 2019).

Parameters list	
Manual settings	Dynamic variables
Number of towers	Street network
Row building max. Floor	Public area location
Floor area ratio (FAR)	Tower location
Public area sqm	Plot length
Avg. floor height	Building types
Building depth	
Building setbacks	
Apt. length	
Buildings' Setbacks	

Figure 2. Parameter's list

### 2.2.1 Site boundary, Street network and public-school area

The site boundary line (see fig.3a) is set as the basis for the model and does not change during the process. According to the design principles, ten possible street networks and public-school location scenarios are planned and stored as a list, later to be employed as variables in the optimization process. Each street width is defined depending on its location and length, eventually splitting the boundary into blocks (Figure 3b).

### 2.2.2. Block subdivision to plots by length

The blocks are subdivided into lots according to their length by applying the Decoding Spaces tool kit (Koenig, 2017). Using the length as a variable enables the generation of multiple lots in various sizes at each iteration (Figure 3c).

### 2.2.3. Subdivision of building typologies and apartments

Building typologies are defined in advance with stakeholders to include desired formal characteristics appropriate to the project. Both building and apartments are defined by their front length and depth and are manually set according to the typology design (Figure 3d). For this case study, two types of low-rise row buildings and two types of high rises buildings are defined.

2.2.4. *Building's volume*

Building height is an outcome of the desired density calculated as FAR and their typology coverage (Figure 3e). Each plot sqm is divided by the FAR value, which gives the number of the approximate floor for each building. Due to the nature of the design problem in this case study, low buildings height parameters derive from the manual setting parameters while the high-rise height compensate the missing floors to reach the desired FAR.

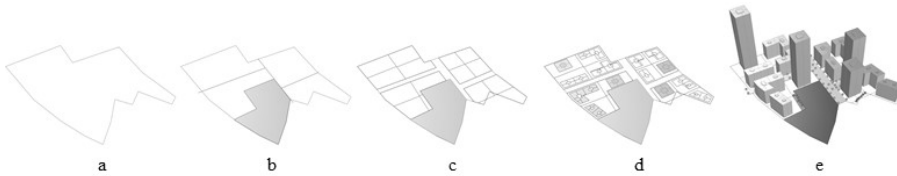


Figure 3. The urban procedural model generation stages

2.3 SRS INDICATORS AUTOMATION

In the local context of this research, we use indicator description, metrics and relative scoring from the analog Israeli evaluation tool for urban design and sustainability, '360 ° neighborhood' (ILGBC, 2019). For this case study we selected to automate five indicators based on their potential conflict with the planning requirements and their dependency on urban form (Figure 4). Moreover, the '360° neighborhood' indicator descriptions enables to categorize them by the three sustainability pillars: social, environmental, and economic (Koenig et al., 2021). While the social and environmental are from the SRS metric, the economic metric derives from the developer's programmatic requirements. For example, we aim to achieve the required density while preserving the needed "building sun rights" or achieving the developer-required towers while standing at the "affordable housing" apartments from tower percentage requirements. Later in the process, indicators are evaluated separately and act as objectives for the MOO process.

SRS Indicators list							
Indicators	Affordable Housing(AH)	Density	Walkable Streets(WS)		Buildings Sun Rights (BSR)	Housing Mix(HM)	
Value threshold	a. Less than 20% of units from towers	a. 5 units per 1000 sqm	a. Less than 5m Front setback 55% to 70% = 2 70% & more = 3	b. 80% & more of 1:1.5 building street ratio	a. 4 hours sun exposure in 50% roofs surface	a. 25% of small apartments and the average unit is less than 100 sqm	b.75% three types apartments
Score	1	5	2 - 3	1	1	1	2

Figure 4. Table of the SRS chosen indicators, their value threshold and score.

2.3.1 "Affordable housing " (AH) 1 point

"Affordable housing " attempts to ensure residential units within the neighborhood that are accessible to the entire population. This metric requires that the percentage of units from 'high-rises' does not exceed 20% of the total housing units, as '360° neighborhood' considers them less affordable.

### 2.3.2 "Density" - 1- 5 point

"Density" promotes liveability, walkability, and transportation efficiency, reducing distance travelled (Koenig et al., 2021). In '360° neighborhood' density calculations include all planned and existing buildings within the project boundary. First, the project will be at least one and a half times larger than the minimum residential density requirements. The second requirement demands that any residential area be at a density of five dwelling units per thousand sqm, and above ten for maximum points.

### 2.3.3 "Walkable streets" (WS) 1-4 points.

"Walkable streets" promotes walking by providing safe, appealing, and comfortable street environments. It requires at least 55% of the block length to have a facade setback less than 5m (Figure 5). and a minimum front façade to a front plot ratio of 5.5:10 or more than 7:10. Furthermore, it requires at least 80% of all the blocks length within the project to have a minimum building-height-to-street-centreline ratio of 1:1.5 (Figure 6)

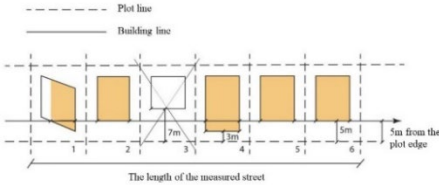


Figure 5. Setback requirement diagram from '360° neighborhood'

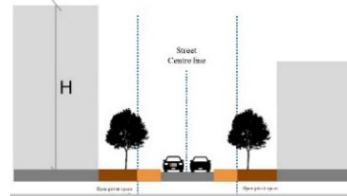


Figure 6. building-height-to-street-ratio diagram from '360° neighborhood'

### 2.3.4 "Buildings Sun right"(BSR) 1 point

"Buildings Sun rights" refers to channelling solar radiation to illuminate buildings and generate renewable energy. It requires that at least 90% of the buildings' roofs have 50% surface with four hours sun exposure time between 09:00-15:00 (Figure 7). The solar radiation studies were carried out using the plug-in Ladybug within Grasshopper.

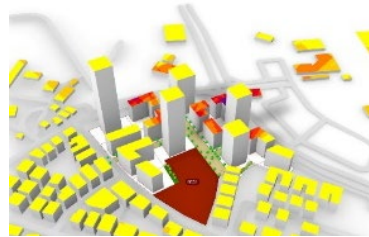


Figure 7. BSR, sun radiation analysis

### 2.3.5 "Housing mix" (HM)2-3 points

"Housing mix" includes apartments of different sizes and allows for a choice of accommodation from different socio-economic backgrounds and needs of populations. It requires that the average housing unit size in the project be up to 100 sqm, and at least 25% of the housing units will be small apartments 30 - 80sqm. The second requirement demands that at least 75% of the buildings in the project will include apartments of at least three different sizes, while one of the sizes must be small.

## 2.4 MULTI OBJECTIVE OPTIMIZATION

Whether architectural problems are well composed as MOO remains a discussion in architectural design optimization research. However, MOOs are recommended when necessary to understand trade-offs between conflicting objectives (Wortmann & Fischer, 2020). The algorithm we used in this study is the Radial Basis Function Multi-Objective Optimization (RBFMOpt) inside the "Opossum" plug-in for grasshopper. RBFMOpt is a novel, machine learning-related MOO algorithm that potentially is more efficient than evolutionary MOO algorithms like NSGA-II and HypE, popular MOO algorithms in design communities (Wortmann, 2017). The optimization trial consisted of 1000 iterations resulting in 128 possible solutions (figure 8) with an above-average SRS ranking. The optimization objectives are maximizing each indicator's points score. This way, the different points perform as weight in the optimization process.

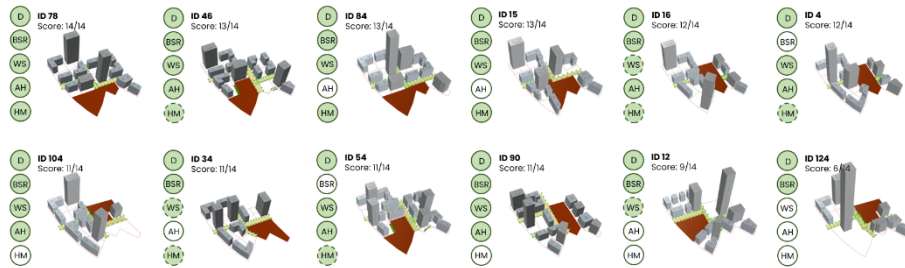


Figure 8. Possible scenarios, their SRS ranking and score. Full green dot - full indicator score, Dashed line - partial score and white dot - zero score.

## 3. Simulation Result and Analysis

Given that SRS has a holistic approach, and each indicator potentially can affect the optimization result, all indicators should be integrated to comprehend their behavior during the optimization process. However, the simulation aimed to test the integration of the SRS scores in a multi-objective optimization process and resulted in several behaviors indicating the experiment's success in our case study. Initially, we selected the high-ranking scenarios, compared them, and examined the relationship between their design qualities and the potential trade-off. This process highlighted recurrent patterns in several designs that can potentially become discussion topics in the planning process. For example, achieving the "Affordable housing" point was rare due to its conflict with the developer requirement of multiple towers. In our analysis, this was mitigated through the location of the public open space, which demonstrated its importance to the developer. Second, we produced a Spearman correlation matrix (figure 9) to understand the relationships between each SRS indicator to confirm our hypotheses. Moreover, the correlation matrix gave us a deeper understanding of the results as it is challenging to understand when the indicator influence comes from its given weight or its conflicting nature with the project's constraints and requirements.

As density plays a critical value in urban planning, we chose to use the FAR as an input rather than a goal in the optimization. The outcome of this choice is that all iterations answer SRS required density and the developer FAR requirements. Moreover, this makes the optimizing process clearer than using building floor numbers as input and FAR as an optimization objective. We used Spearman Rs value to show positive (blue) and negative (red) correlation, while the circle size shows if the correlation is strong, moderate, or weak. "Buildings Sun right" has a moderate correlation with the tower height resulting in returning solutions of public area located in the south of the plan and towers in the north part of the plan, possibly due to the shadow of towers affecting the solar radiation. "Walkable streets" has a moderate positive correlation with "Affordable housing" and negative on solar radiation on rooftops. Building height correlated positively with one "Housing mix" requirement and negatively with the other. Considering the requirement context, we assumed our tower typologies contained many large apartments and should be changed in future iterations. Also, some possible outcomes received identical SRS point score from different indicators prioritisation. This helps determine the influence of each indicator on the design and can be discussed in the planning process with the multiple stakeholders (Figure 10).

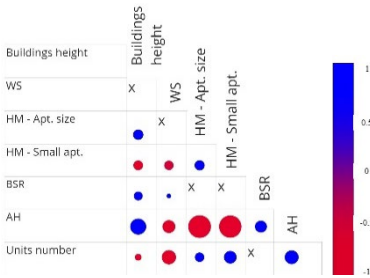


Figure 9. Spearman Rs values  
Correlation matrix

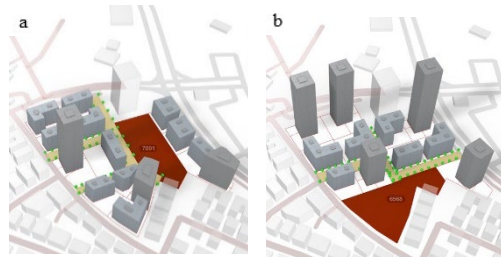


Figure 10. Design outcomes with the same SRS score  
prioritising different indicators. 9a AH, 9b BSR

#### 4. Discussion and future work

There are significant challenges concerning the integration of generative models with SRS. First, the design space of the model cannot be unbiased or completely variable (Wilson et al., 2019, Nagy, 2018). While a more flexible model can generate unplanned scenarios, creating high-performing solutions requires more objectives and constraints. On the other hand, a less flexible and well-defined model will generate achievable solutions, but it could exclude possibilities desirable to some stockholders. As SRS metrics perform as guidelines in the planning process, they should be defined more as thresholds within the planning goals of the optimizing process. Therefore, one of the challenges is creating a well-structured design space to guide a successful generative process. Nonetheless, the simulation results have shown that automating SRS indicators can guide a procedural model and reveal relevant sustainable and economically scenarios for advising decision-makers. While saving planning time and manual work, this framework can lead to a higher SRS implementation in stages of early urban planning, therefore, potentially increasing positive impact on sustainability.

As described in the previous chapters, in this case study we selected to automate 5 out of 19 quantitative indicators from the local '360° neighborhood' SRS, selecting the indicators that are conflicting with the project requirement and related to urban form. Integrating only part of SRS indicators in this process is not ideal given the holistic approach of SRS. Based on our initial results, we assume that with further developments most SRS indicators within '360° neighborhood' could be automated and integrated into the computational framework described in this paper. However, further study will examine how this affects the optimization process, results readability, and iterations number required to reach optimal solutions. The simulation results established conclusions concerning our case study and the success of this process. Nevertheless, further evaluation is required to establish insights on both '360° neighborhood' and possibly other SRSs. Comparing indicators from numerous SRSs in a similar generative process could reveal correlated differences and relations.

## 5. Conclusion

This paper presented the integration of SRS's indicators in a generative design workflow at an urban scale through the case study of an urban renewal project in Israel. The proposed computational workflow allows the exploration of multiple iterations through high-performing design solutions relative to a chosen set of SRS goals. By applying such workflows on an urban scale, we enhance the integration of verified sustainability goals in the planning process and its potential correlation with the multiple stakeholders involved. A notable advantage of the framework is that it can act as a 'discussion table' in planning meetings for the evaluation of existing and future planning scenarios that are inclusive to multiple stakeholders and are driven and informed by verified sustainability rating systems. Future steps should examine further case studies and address municipal challenges concerning SRS implantation through the proposed framework within current planning processes. Generative design tools are currently transforming the designer's role from designing through plans to creating design spaces that can be explored and optimized by computational systems. Therefore, we believe that the integration of automated verified sustainability indicators within a generative process can be used as a platform to enhance future sustainability in a multi-stakeholder urban planning process.

## References

- Aspinal, S., Sertyesilisik, B., Sourani, A., Tunstall, A. (2012). How Accurately Does Breeam Measure Sustainability? *Creative Education*. 03(7B).  
<https://doi.org/10.4236/ce.2012.37B001>.
- Castanheira, G., Bragança, L. (2014). The Evolution of the Sustainability Assessment Tool SBToolPT: From Buildings to the Built Environment. *The Scientific World Journal* 2014. 491791. <https://doi.org/10.1155/2014/491791>
- Elshani, D., Koenig, Reinhard., Düring, S., Schneider, S., Chronis, A. (2021). Measuring sustainability and urban data operationalization. In *26th International Conference on Computer-Aided Architectural Design Research in Asia: Projections, CAADRIA 2021* (pp.407-416). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA).

- Halatsch J., Kunze A., Schmitt G. (2008). Using Shape Grammars for Master Planning. *Design Computing and Cognition '08*. Springer, Dordrecht, pp. 655-673.  
[https://doi.org/10.1007/978-1-4020-8728-8\\_34](https://doi.org/10.1007/978-1-4020-8728-8_34)
- ILGBC. (2019) '360 Neighbourhood' SRS. Retrieved December 4,2021, from <http://www.nd360.org/on>
- Koneig, R., Miao, Y., Schneider, S., Vesely, O., Bus, P., Bielik, M., Abdulmawla, A., Dennemark, M., Fuchkina, E., Aichinger, A., Knecht, K. (2019). DeCodingSpaces Toolbox for Grasshopper: Computational analysis and generation of STREET NETWORK, PLOTS and BUILDINGS.
- McGlashan, N., Ho, C., Breslav, S., Gerber, D., Khan, A. (2021) Sustainability Certification Systems as Goals in a Generative Design System, In *Proceedings of the 2021 Symposium on Simulation for Architecture and Urban Design, SimAUD 2021* (pp.102)
- Miao, Y., Koenig, R., Knecht, K. (2020). The Development of Optimization Methods in Generative Urban Design: A Review. In *Proceedings of the 2020 Symposium on Simulation in Architecture and Urban Design, SimAUD 2020*.
- Nagy, D., Villaggi, L., Benjamin, D. (2018). Generative Urban Design: Integrating Financial and Energy Goals for Automated Neighborhood Layout. In *Proceedings of the 2018 Symposium on Simulation for Architecture and Urban Design, SimAUD 2018* ,25 (pp. 1–8).
- Natanian, J., Auer, T. (2020). Beyond Nearly Zero Energy Urban Design: A Holistic Microclimatic Energy and Environmental Quality Evaluation Workflow. *Sustainable Cities and Society*, 56,102094.  
<https://doi.org/10.1016/j.scs.102094>
- Roudsari, M.S. and Pak, M. (2013) Ladybug: a parametric environmental plugin for grasshopper to help designers create an environmentally conscious design. In *13th Conference of International Building Performance Simulation Association. IBPSA 2013* (pp. 128-3135).
- Smith, RM and Bereitschaft, B (2016). Sustainable urban development? Exploring the locational attributes of LEED-ND projects in the United States through a GIS analysis of light intensity and land use, *Sustainability*, 2016, 8(6), 547.  
<https://doi.org/10.3390/su8060547>
- Wilson, L., Danforth, J., Davila, CC., Harvey, D. (2019). How to generate a thousand master plans: a framework for computational urban design. In *Proceedings of the 2019 Symposium on Simulation in Architecture and Urban Design, SimAUD 2019* (pp. 1–8.).
- Wortmann, T., (2017). Opossum-introducing and evaluating a model-based optimization tool for grasshopper, In *22th International Conference on Computer-Aided Architectural Design Research in Asia: Protocols, Flows and Glitches, CAADRIA 2017*, (pp. 283–292). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA).
- Wortmann, T., Fischer, T. (2020). Does architectural design optimization require multiple objectives? A critical analysis. In: *25th International Conference on Computer-Aided Architectural Design Research in Asia: RE: Anthropocene, CAADRIA 2020* (pp. 365-374). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA).
- Yoffe, H., Plaut, P., Fried, S., Grobman, Y. J. (2020). Enriching the Parametric Vocabulary of Urban Landscapes. A framework for computer-aided performance evaluation of sustainable development design models. In *Anthropologic – Architecture and fabrication in the cognitive age. The 38th eCAADe conference, ECCADE 2020*, 1, (pp.47–56).



# AN ALTERNATIVE MODEL FOR URBAN RENEWAL: A GENERATIVE APPROACH TO THE (RE)-DEVELOPMENT OF XIAN VILLAGE

LING KIT CHEUNG<sup>1</sup>, ZHITAO XU<sup>2</sup>, PEI CHEN<sup>3</sup> and MOHAMMED MAKKI<sup>4</sup>

<sup>1,2,3,4</sup>*University of Technology Sydney.*

<sup>1</sup>*lingkit.cheung-1@student.uts.edu.au, 0000-0001-5040-8686*

<sup>2</sup>*Zhitao.Xu@student.uts.edu.au, 0000-0002-0541-6780*

<sup>3</sup>*Pei.Chen-1@student.uts.edu.au, 0000-0003-2677-3565*

<sup>4</sup>*Mohammed.Makki@uts.edu.au, 0000-0002-8338-6134*

**Abstract.** The impact of urban renewal, specifically in countries experiencing rapid urbanisation due to population growth, has resulted in the erasure of urban culture and heritage in favour of repetitive homogeneity that has been synonymous with 20th century modernist planning models. One such region experiencing this rapid urban renewal is the Guangzhou region in southern China. The presented experiments examine Xian Village in Guangzhou, a culturally rich urban tissue currently experiencing redevelopment, and proposes an alternative model for urban renewal, employing a bottom-up approach to urban growth through the use of a multi-objective evolutionary model; presenting a model that integrates historic and existing urban characteristics adapted to future development plans.

**Keywords.** China, Guangzhou; Xian Village; Village in the City; Urban Renewal; Cultural and Heritage Preservation; Multi-Objective Evolutionary Algorithm (MOEA); SDG 10; SDG 11; SDG 13.

## 1. Introduction

The complexity of urban form, coupled with a rapidly changing climate and an exponentially growing population, has highlighted the demand of understanding this complexity through bottom-up approaches rather than top-down systems (Batty, 2008), in which the rationalisation of the urban fabric is achieved through local rules in favour of global order (Weinstock, 2010). Culture and heritage plays a vital role when examining this urban complexity, a relationship discussed by many throughout the 20th century; from Gustavo Giovannoni, who equated the value of urban heritage to that of urban development (Giovannoni, 1913), to Patrick Geddes, who emphasised the understanding of the city in both its growth and life (Geddes, 1915), to Lewis Mumford, who emphasised the relationship between people and the city (Mumford, 1935), and to Jane Jacobs, who was one of the first voices against the ‘city as a machine’ (Jacobs, 1961). However, environmental changes, specifically related to population growth in countries experiencing rapid urbanisation (e.g. China, India and

South Africa), have led to a growing pattern of ‘cultural erasure’ and ‘urban replacement’ throughout the 21st century (Kiruthiga and Thirumaran, 2019). Cities that can no longer ‘adapt’ to environmental stresses often prompt the demand to rebuild, erasing the old and replacing it with what is deemed suitable for near future. In doing so, the historical significance of culture and heritage of the existing is discarded.

This paper examines these phenomena by reimagining Xian Village, a ‘village in the city’ in China undergoing this culture erasure, through the design of an urban tissue that retains the urban characteristics of the existing village yet integrates it to the surrounding city currently experiencing rapid urban renewal. In addition, the presented experiment examines three UN SDGs; firstly, addressing the inequality between the inhabitants of Xian village and those surrounding it (SDG 10); secondly, through a more sustainable approach to urban living by integrating design goals that support the end users of the city, with a primary focus on pedestrian activities (SDG 11); and finally, climate response through utilising environmental and climatic design goals and analytic criteria as primary drivers in the conducted experiment (SDG 13).

## 2. Research and Background

### 2.1. ERASURE OF CULTURE IN FAVOUR OF NEW DEVELOPMENT

In China, especially the southern regions, rapid urbanisation has created a phenomenon called ‘village in the city’ (ViC). They are formed when villages located in the suburbs are surrounded by rapid urban redevelopment, frequently transformed into small pockets of under-developed superblocks (Liu, 2019). These ViCs are infamous for their dense, compact and oftentimes illegal structures due to population demands. They are regarded by some as “an obstacle to the transformation of a modern metropolis, an eyesore in a well-planned city” and are slowly being demolished and re-developed (Nanfang Daily, 2000). In the last two decades, many ViCs in the Guangzhou region have experienced redevelopment, in which heritage-rich urban fabrics have been replaced by modernist urban models. One such example is ‘Xian Village’; an urban tissue that represents the struggle between a decentralised, bottom-up urban fabric failing to adapt to the speed of growth and demands of urbanisation, inevitably leading to its destruction in favour of top-down, centralised urban form.

### 2.2. XIAN VILLAGE

Xian Village, settled over 800 years ago, is one of the remaining ViCs in the Guangzhou region, and thus a critical urban tissue to examine. Analysis into the village is presented through two scales, the macro (i.e. the relationship of the village to its surrounding context), and the micro (i.e. the inner workings of the village itself). In the macro scale, the rapid urbanisation of the city around Xian village has generated various challenges. The village is surrounded by 4 arterial highways with only 2 pedestrian walkways, cutting off accessibility to the village and limiting pedestrian movement between the village and the surrounding blocks (Figure 1). The contrast between the building typologies of the village and the surrounding blocks is vast, revealing a high disparity of density and functionality within the urban form (Figure 1). Most importantly, the village represents centuries of cultural and historical value, a

trait not retained by the surrounding context. In the micro scale, Space Syntax centrality analysis (Hillier and Hanson, 1984) of the village reveals that despite the seemingly chaotic streetscape, most of the paths were well utilised by the village’s inhabitants. The majority of the inner streets are pedestrian-only streets due to their width, with commercial shops and street vendors located on either side and cantilevering residential buildings overhead, limiting the skyview factor from street level. The narrowness of the network creates fire and safety risks, as well as minimal sun exposure to the street. The building typologies within the village, primarily a result of increasing population numbers, allowed for increased levels of interaction between the inhabitants (Figure 1).

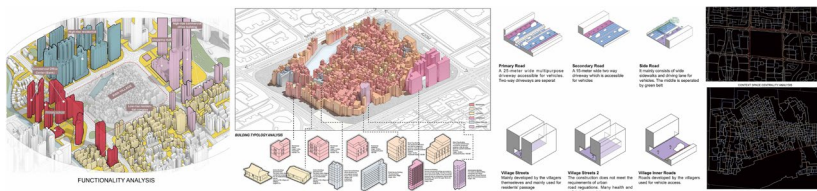


Figure 1. Xian Village typology and network analysis

Throughout the village, there are two primary urban typologies that hold significant value to the inhabitants; ancestral halls and local markets. Regarding the former, there are 5 ancestral halls located within the village, these are Chinese traditional ceremonial halls where ancestors are worshipped and serve as the cultural focal point of the village. Regarding the latter, markets are one of the lifelines of the village and play a vital role in its urban heritage. However, the lack of usable space within the village forced vendors to occupy the streets regardless of the availability of a storefront, resulting in a lack of regulation throughout the village.

### 3. Method

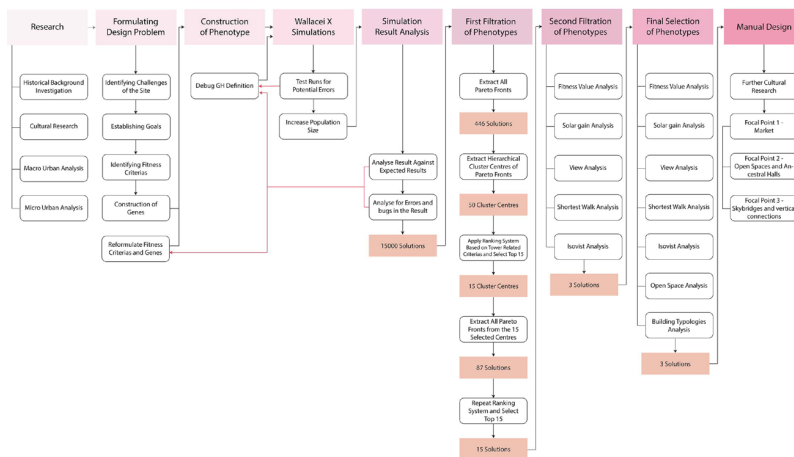


Figure 2. Pseudo Code (workflow methodology)

The presented experiment utilises a Multi-Objective Evolutionary Algorithm (MOEA) to generate an urban superblock that addresses Xian Village's high population density, while maintaining its existing cultural elements, heritage buildings and the connectivity between the residents as a community. The experiment is presented through 4 key stages (Figure 2); formulation of the design problem, in which the chromosomes, fitness objectives and phenotype are developed; running the MOEA and debugging (requiring several rounds of reformulation); analysis of the results outputted by the algorithm; and finally, analysis of the selected solution.

## 4. Experiment Setup

### 4.1. CONSTRUCTING THE PHENOTYPE

The phenotype is a reconstruction of the Xian Village superblock, with an emphasis on various characteristics such as cultural heritage, population density, building typologies, street network and site context. To retain the existing features within the superblock and adopt new urban features, the construction of the phenotype is divided into four key stages: Identifying podiums, towers, buildings and bridges.

The phenotype sits within the existing site boundaries (560mx440m), which is divided into a 15m by 11m grid, mimicking Xian Village's original scale. The parameters (chromosomes) defining the phenotype's morphology, and controlled by the algorithm to create variation, are the following (Figure 04): The size of each cell in the grid (by manipulating the corner points of each cell in the X and Y); the size and location of podiums (by clustering cells together); location of towers and their rotation in response to the ancestral halls; height of buildings on the podiums; size, orientation and shape of ground level buildings (cells are divided, rotated and merged); defining street network (by identifying routes between ancestral halls); location and size of lakes (according to proximity to ancestral halls); heights of ground level buildings according to relationship to ancestral halls; location of bridges between podiums, as well as between towers to each other, and towers to the context (i.e. the site outside the village).

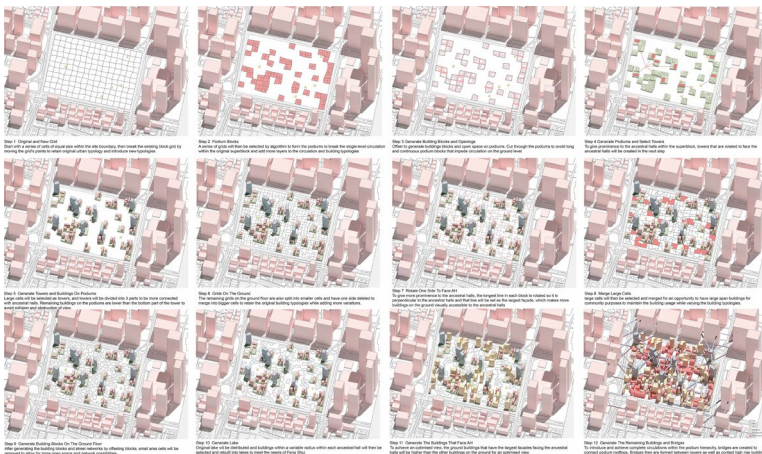


Figure 3. Phenotype Construction Process

There are additional parameters ‘hardcoded’ (i.e. not controlled by the algorithm) into the phenotype’s construction; these include the deletion of cells with an area that falls below a certain threshold and the division of tower heights into three parts and rotating each part accordingly (lowest part, which is residential, does not rotate; middle part, offices, rotates towards the closest ancestral hall; and the top part, cultural, is oriented towards the ancestral hall) (Figure 3).

#### 4.2. FITNESS OBJECTIVES

In addition to the chromosomes defining the phenotype’s morphology, to generate an urban tissue that addresses the demands of population growth, retention of cultural heritage, improved social spaces (public space and solar access) and greater connectivity to the external context, 4 fitness functions have been identified to drive the MOEA: 1. Maximise open space on ground and podium rooftops; 2. Maximise population density; 3. Maximise views to ancestral halls and lakes; 4. Maximise solar exposure on ground level and on building facades.

#### 4.3. SIMULATION SETTINGS

The MOEA used for the experiment is the NSGA-2 algorithm developed by Kalyanmoy Deb (Deb et al., 2000) within the software Wallacei (Makki et al., 2018), a plugin for Grasshopper 3D. The simulation was run on a high-end consumer-grade PC, running with Windows 10 Home 64-bit (Version 21H2 / DirectX 12), AMD Ryzen 9 5900X 12-Core Processor (24 CPUs threads)- 4.20 GHz with 32.0 GB of RAM, Asus TUF GAMING X570-Plus and NVIDIA GeForce RTX 3080 graphics card.

### 5. Experiment Results and Selection Process

#### 5.1. ALGORITHM RESULTS

The Algorithm ran a population comprised from 300 generations with 50 individuals in each generation, totalling 15000 solutions with a simulation runtime of 30 hours and 37 minutes. Analysis of the results and charts outputted by the simulation highlights that all objectives improved in mean value throughout the algorithmic run. Whereas objectives 2 and 3 (population density and views to ancestral halls) demonstrated greater convergence when compared to objectives 1 and 4 (open space and solar gain),

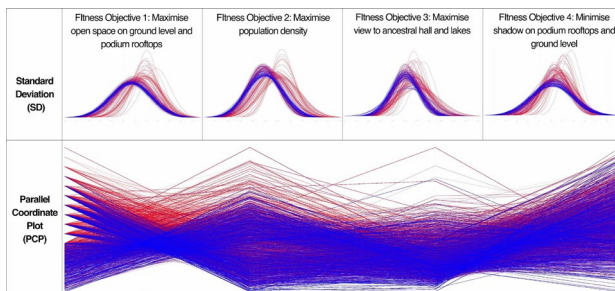


Figure 4. Algorithm Results (Top: Standard Deviation Graph, Bottom: Parallel Coordinate Plot)

analysis of the results through the parallel coordinate plot presents lower conflict between objectives 2 and 3, allowing the algorithm to potentially favour these two objectives over the others (thus resulting in greater convergence). From the 15000 solutions generated by the algorithm, 446 comprised the pareto front. These pareto solutions were extracted and analysed further in the selection process presented in the following sections (Figure 4).

## 5.2. SELECTION PROCESS

The MOEA resulted in a large number of pareto front solutions, thus necessitating a thorough selection process to ‘filter’ through the 446 pareto front solutions and select one. As such, a three-stage selection process has been applied to the pareto front solutions, with each stage analysing the solutions through different scales and criteria.

### 5.2.1. First Selection Matrix

Item #	40-0	70-2	73-45	77-49	83-48	93-13	104-17	104-34	112-43	117-24	124-21	125-3	139-33	217-47	268-48
Towers	7	7	5	7	6	5	6	7	6	6	5	5	7	7	7
Sky Bridges	16	12	7	11	11	8	11	12	10	7	6	5	13	11	11
Lowrise (large) buildings (0.5 pt)	2	N	N	1	2	6	3	6	N	4	6	6	4	1	N
^ (-0.5 pt per building if over 3)						-1.5		-1.5		-0.5	-1.5	-1.5	-0.5		
North to South connection (1 pt)	Y	Y	N	Y	Y	Y	Y	N	N	N	N	N	Y	Y	Y
East to West connection (5 pt)	Y	Y	N	N	N	N	N	Y	N	N	N	N	N	N	Y
15 adjacent grids empty of bridges (-5 pt)				-5				-5	-5		-5	-5	-5	-5	
Score	30	25	7	19.5	19	15.5	14.5	16.5	21	9.5	7.5	6.5	17.5	19.5	24

Figure 5. Results of first selection matrix

In the first round of selection, the pareto front solutions were clustered using hierarchical clustering (average linkage) into 50 clusters (using the 4 fitness values for each solution). The solution at the centre of each cluster (which in this case is determined to be the representative of the cluster) is selected and ranked based on tower-related criteria (number of towers, sky bridges, low-rise and large buildings, north to south connections and east to west connections (the latter is prioritised as it connects the two neighbouring CBD blocks) and consecutive blocks without bridges overhead). Each solution is given a new ranking and the top 15 cluster centres are selected. The process is then repeated to all the solutions that belong to these 15 cluster centres (a total of 87 solutions), and scored once more, selecting the top 15 solutions for the second round of selection (presented in Figure 5).

### 5.2.2. Second Selection Matrix

In the second round of selection, the 15 selected solutions were analysed further, in which 5 of the 15 solutions were culled as they did not demonstrate sufficient connectivity to the external context of the urban tissue (these are highlighted in grey in Figure 5). The remaining 10 solutions are analysed according to additional criteria; these include their fitness diamond chart (i.e. the relationship of the four objectives for each solution), phenotypic characteristics, solar exposure on building facade and low-rise building roofs, solar exposure on ground and podium rooftops, views to ancestral halls, shortest walk between ancestral halls and to site boundary, isovist occlusivity and

isovist min.radial using decoding spaces plugin (CPlan, 2021) (Figure 6). Each solution was given a new ranking, and the top 3 solutions were selected.

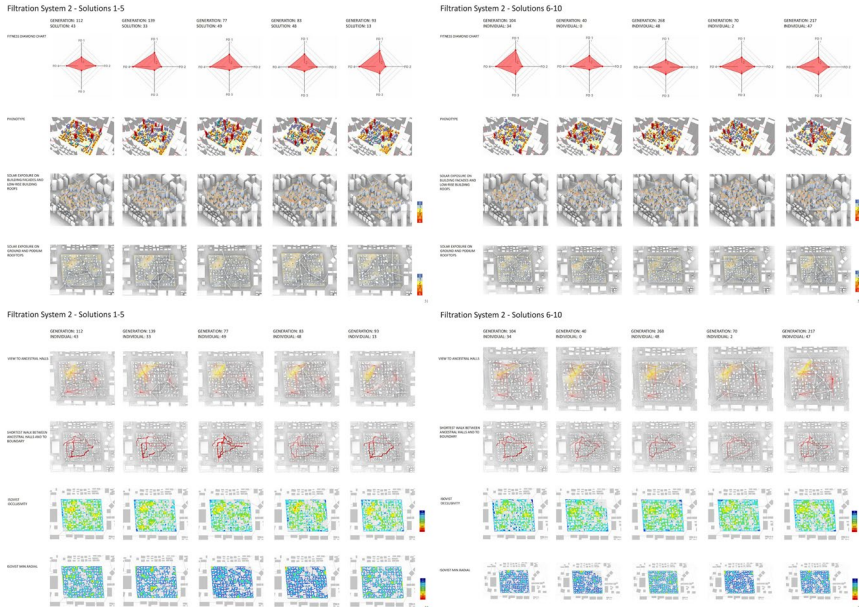


Figure 6. Results of second selection matrix

### 5.2.3. Third Selection Matrix

In the final round, the 3 selected solutions were analysed further with the following criteria: Total number of podium open space, total area of podium open space, total area of ground open space, total solar gain on ground, podiums and building facades, view to ancestral halls, shortest walk between ancestral halls and to site boundary, isovist occlusivity and isovist min.radial, population density)

The results of the third selection round present solution 2 (gen.70/ind.2) as the most favourable solution (Figure 7). Solution 2 demonstrates greater solar gain (on streets, podiums and building facades), as well as improved views on ground level throughout the urban tissue (as per the results of the occlusivity analysis). Pedestrian access on street level between ancestral halls as well as between public spaces performs better, and so does the connectivity to the site boundary (on ground level) and to the external context (through bridges on upper levels). As such, solution 2 was selected as the most optimal solution for the design objectives and assessment criteria defined in the simulation, and analysed further at 3 different scales to better understand the architectural urban impact of the selected solution.

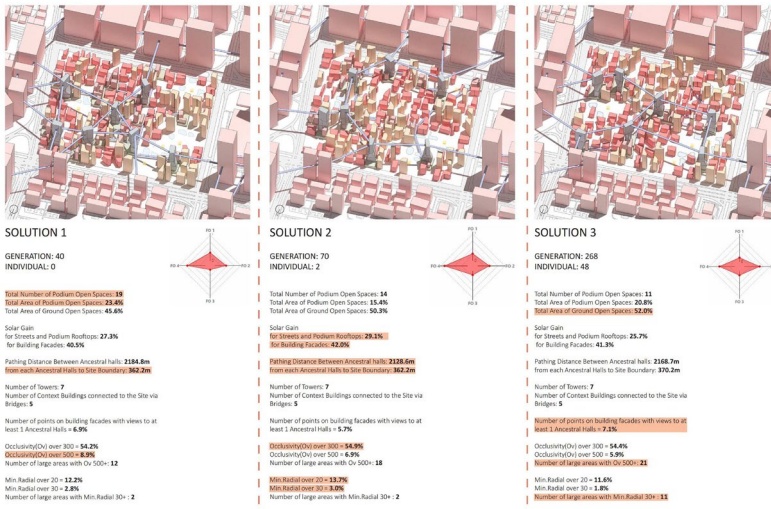


Figure 7. Results of third selection matrix. Solution 2 (middle) selected as the most optimal.

## 6. Selected Solution Analysis

The two vital cultural typologies of Xian Village, ancestral halls and markets, were further developed to illustrate a more detailed urban character of the selected solution. Three focal points, at different scales, are selected to demonstrate the qualitative aspects of the cultural typologies and their impact on the users of the space.

### 6.1. FOCAL POINT A - NEW MARKET SPACE

Focal Point A addresses the markets and street vendors of Xian village. The lack of store fronts and defined market space results in a haphazard occupation of the street. To maintain the flexible character of the street vendors, an adaptive module is integrated within the market space allowing for improved order while maintaining flexibility and growth (Figure 8). The module allows for the occupation of both public open space as well as narrow streets throughout the village. Ramps on the podium façades generate greater interaction between the ground floor and podium roof top (Figure 8), allowing for easy access between the two (both physical and visual).

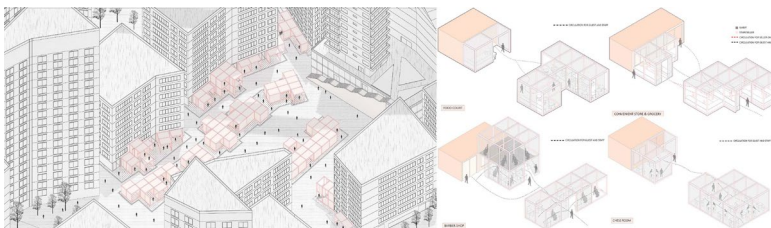


Figure 8. Focal point A: perspective and top view



## 6.2. FOCAL POINT B-ANCESTRAL HALLS OPEN SPACE

Focal point B explores the interaction between Ancestral Halls and lakes and the division of open spaces to cater to different functions. Open space throughout the village is defined into three categories: public squares, green spaces, and recreational areas near community building clusters. Selected open spaces near the site boundary are allocated as collection points for public transport, as well as car parks, allowing for the majority of the internal streets of the site to be pedestrian only (Figure 9).

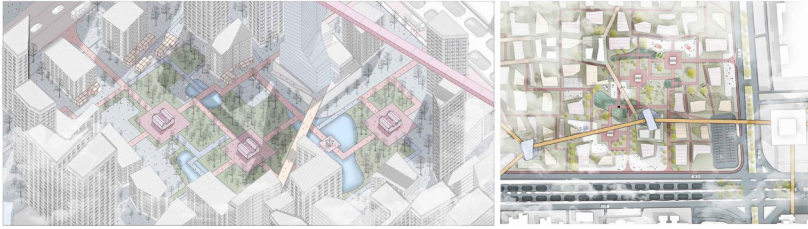


Figure 9. Focal point B: perspective and top view

## 6.3. FOCAL POINT C-SKY BRIDGES

The last focal point explores the vertical connections between sky bridges and buildings, and the edge conditions of the site and context. Figure 10 illustrates the circulation hierarchy of ground, podium bridges and tower bridges. A detail of the interaction between the podium bridges and the buildings is also illustrated. The point of contact between the two is converted into a public space allowing for more accessibility and interaction; equally activating both the ground level and bridge level.

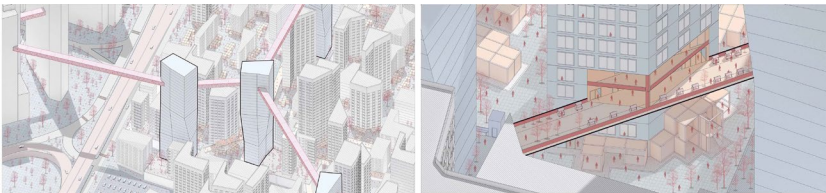


Figure 10. Focal point C: perspective view and detailed bridge to building relationship

## 7. Discussion and Conclusion

The presented research examines the increasing trend in countries experiencing rapid urbanisation of replacing culturally rich and historic urban tissues with modernist urban models. Through the use of evolutionary methods, there is an emphasis on the utility of bottom-up systems that address local conditions and relationships between various morphological characteristics within an urban patch, in which the complexity of city design is approached through localised decisions (in the presented experiments, the chromosomes and genes that define the phenotype) that reflect the end user of the space. At every stage of the experiment presented (phenotype construction, analysis of results, selection and filtration and local analysis of the selected solution), the aim of retaining the village's cultural urban characteristics while integrating it with its urban

environment was the primary driver throughout.

In the analysis of the three focal points in the selected solution, the evolved urban tissue demonstrated improved public awareness of ancestral halls through improved accessibility (both physical and visual) throughout. Open space on both ground and podium level with increased solar access contributes to one of the key cultural characteristics of the site, the markets, while the activation of the bridges connecting podiums and external context allows for greater integration to the surrounding site.

The experiment puts forward an alternative model of urban redevelopment that examines urban heritage within a superblock; identifying key elements of cultural significance and optimising it through a MOEA followed by a thorough selection mechanism, there were key limitations however, both in the formulation of the computational setup as well as the resulting solution. The criteria implemented in the first round of selection was unsuccessful in detecting solutions that lacked integration to the surrounding site, indicating the necessity for a reformulation not only of the design problem, but also the selection criteria as well. Although the selected solution addressed the various criteria used for its assessment, further work is required to address issues surrounding energy consumption (at different scales) throughout the superblock, as well as improved integration between the context and superblock at street level. Where the integration of bridges addressed this on upper levels, the positioning of buildings throughout the perimeter of the selected solution limits its integration to the surrounding. Finally, future research on the quantification of cultural and social characteristics, and its impact on urban form, is required to strengthen the presented model in order to continue to address and reverse urban cultural erasure.

## References

- Batty, M. (2008), The Size, Scale, and Shape of Cities, *Science*, 319, 769–771.
- Computational Planning Group (2021), *DeCoding Spaces Toolbox*, <https://toolbox.decodingspaces.net/>
- Deb, K., Agrawal, S., Pratap, A., & Meyarivan, T. (2000, September). A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. *In International conference on parallel problem solving from nature* (pp. 849-858). Springer, Berlin, Heidelberg.
- Geddes, T. (1915), *Cities in evolution: An introduction to the town planning movement and to the study of civics*. William and Norgate.
- Giovannoni, G. (1913), Vecchie città ed edilizia nuova, *Nuova Antologia XLVIII* (995): 449-472. (translated).
- Hillier, B. and Hanson, J. (1984) *The Social Logic of Space*. Cambridge University Press.
- Jacobs, J. (1961), *The Death and Life of Great American Cities*. Vintage.
- Kiruthiga, K., Thirumaran, K. (2019), Effects of urbanization on historical heritage buildings in Kumbakonam, Tamilnadu, India, *Frontiers of Architectural Research*, 8(1), pp. 94-105.
- Liu, Y. (2019), Village in the City. In Orum, A. M. (Eds.), *The Wiley Blackwell Encyclopaedia of Urban and Regional Studies* (pp. 1–8). John Wiley & Sons.
- Makki, M., Showkatbakhsh, M. and Song, Y. (2018), Wallacei: An evolutionary and Analytic Engine for Grasshopper 3D, *Wallacei*, <https://www.wallacei.com/>.
- Nanfeng Daily. (2000), *When city meets villages*, Nanfeng Daily, (translated).
- Weinstock, M. (2010), *The Architecture of Emergence: The Evolution of Form in Nature and Civilisation*. Wiley.

# OPTIMIZING DESIGN CIRCULARITY

*Managing complexity in design for circular economy through single and multi-objective optimisation*

FREDERICK PETER ORTNER<sup>1</sup> and JING ZHI TAY<sup>2</sup>

<sup>1,2</sup>*Singapore University of Technology and Design*

<sup>1</sup>*peter\_ortner@sutd.edu.sg, 0000-0002-3376-7852*

<sup>2</sup>*jingzhi\_tay@mymail.sutd.edu.sg, 0000-0002-4822-704X*

**Abstract.** This paper advances the application of computational optimization to design for circular economy (CE) by comparing results of scalarized single-objective optimization (SOO) and multi-objective optimization (MOO) to a furniture design case study. A framework integrating both methods is put forward based on results of the case study. Existing design frameworks for CE emphasize optimization through an iterative process of manual assessment and redesign (Ellen MacArthur Foundation, 2015). Identifying good design solutions for CE, however, is a complex and time-consuming process. Most prominent CE design frameworks list at least nine objectives, several of which may conflict (Reike et al., 2018). Computational optimization responds to these challenges by automating search for best solutions and assisting the designer to identify and manage conflicting objectives. Given the many objectives outlined in circular design frameworks, computational optimisation would appear a priori to be an appropriate method. While results presented in this paper show that scalarized SOO is ultimately more time-efficient for evaluating CE design problems, we suggest that given the presence of conflicting circular design objectives, pareto-set visualization via MOO can initially better support designers to identify preferences.

**Keywords.** Design for Circular Economy; Computational Optimisation; Sustainability; Design Optimisation; SDG 11; SDG 12.

## 1. Introduction

Design for circular economy (CE) is an increasingly urgent concern for built environment and product design, with recent calls to action from the United Nations Environment Programme among other international agencies (UNEP, 2019). Currently many frameworks, strategies and metrics exist for CE design, without a single comprehensive or universally accepted method (Saidani, 2019). Furthermore, evaluation of design for CE requires quantitative evaluation of many criteria, making precise evaluation for many design variants a time-consuming process. To support

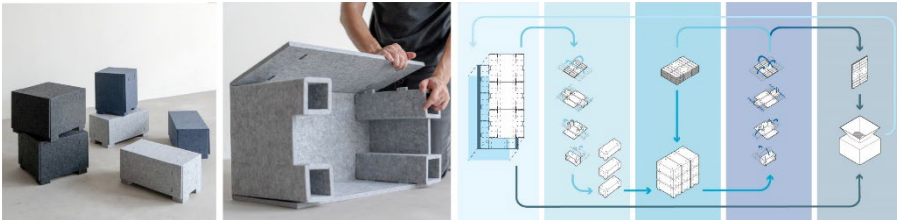
designers in identifying better design options for CE, computational tools are needed to automate quantitative evaluation of CE metrics, and support search among many options to find designs which demonstrate the highest performance while meeting subjective concerns of the designer. Responding to this need is circular design optimization, an area of research that asks how computational tools may support search for optimal circular designs and management of complexity of circular design process.

No definitive methodology to optimize a design for CE currently exists, with the choice between multi-objective (MOO) and scalarized single-objective optimization (SOO) representing a prominent open question. While the many criteria required to assess a design for circularity suggest that MOO may be most appropriate, the computational cost of MOO in comparison to SOO provides a counterargument against this default position (Wortman, 2020). Furthermore, prominent circular design frameworks rank their circularity criteria according to hierarchies of impact, raising the question if scalarized SOO with weights ranked according to these hierarchies may offer a viable means of identifying good results (Reike et al., 2018).

In this paper we propose a framework to apply computational optimization to design for CE, with focus on efficiency of identifying good results and supporting understanding of conflicting objectives. We demonstrate this framework via optimizing a case study of an existing CE furniture design. We compare application of SOO or MOO at specific points in the design process and put forward a method of applying each as appropriate. We employ statistical analysis and novel visualizations to expand on the results and support selection of design variants.

## 2. Methods

### 2.1. DESIGN FOR CIRCULAR ECONOMY: CASE STUDY AND PARAMETRIC MODEL



*Figure 1. Images of design for CE furniture case study: range of prototypes produced (left), single-material assembly process (centre) and illustration of product lifecycle: fabrication, assembly, reuse and storage, disassembly, recycling (right).*

In this paper we test optimizations for criteria related to CE on a parametric model of an existing furniture design. The furniture design was developed by the authors as a modular table/podium system and incorporates CE design strategies including the use of pre-recycled material, quick assembly and disassembly and single material without glues or fasteners (Figure 1). In developing and fabricating the design the authors found that several interlinked design challenges were difficult or impossible to resolve through manual design process. These problems were translated into four quantifiable

objectives for computational optimization (Section 2.2).

A pre-requisite for computational optimisation is developing a fully computational parametric model, with clear inputs and outputs, from the original furniture design. The inputs used include six dimensional parameters that control significant features of the design and two material selection parameters that permit testing for different material and multi-material designs (Figure 2). These eight inputs are transformed into geometry and subsequently evaluated for each objective and inputted to the optimization algorithm.

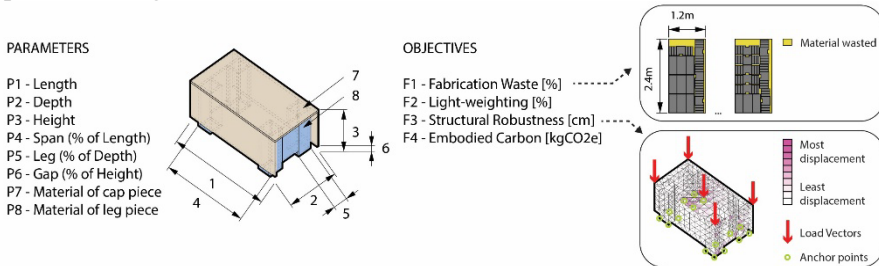


Figure 2. Description of six dimensional parameters, two material selection parameters and four objective functions in the parametric model

## 2.2. DEFINING EVALUATION CRITERIA

To optimize the four key CE objectives described in Section 2.1, we implemented four respective methods of evaluation. Below we describe the objectives, their relationship to a widely used 9R framework of 'r-imperatives' for circular economy and the method of evaluation. The 9R framework encompasses objectives for: 0) Refuse; 1) Reduce; 2) Re-use; 3) Repair; 4) Refurbish; 5) Remanufacture; 6) Re-purpose; 7) Re-cycle; 8) Recover; 9) Re-mine (Reike et al., 2018).

Objective 1 – Fabrication Waste (F1), aims to minimise waste of materials during fabrication. We evaluate the percentage of waste generated from original stock of materials by simulating the fabrication of ten quantities of the design using bin packing algorithms (Figure 2). This objective also sets a constraint to our optimisation problem as the size of each assembly must be within the size of the stock material sheets. This objective is defined relative to 'Reduce' from the 9R framework.

Objective 2 – Light-weighting (F2), aims to minimise materials used for a given size of design. This is done by taking the ratio of volume of materials used to the overall volume of the rectangular bounding-box enclosing the design. This objective is also defined relative to 'Reduce' from the 9R framework.

Objective 3 – Structural robustness (F3), aims to minimise structural displacement under a vertical load of the weight of an average human body when applied evenly across five points on the design (Figure 2). We conducted this finite element analysis using Karamba for Rhino/Grasshopper. Material properties such as specific weight, Young's modulus, shear modulus and yield strength are included as attributes to the materials. This objective is defined relative to 'Reuse' in the 9R framework.

Objective 4 – Embodied carbon (F4), aims to minimise environmental impact by minimising the Global Warming Potential (GWP) of a design variant, evaluated as

kgCO<sub>2</sub>e / kg of material. GWP values are drawn from published sources of identical or close analogue materials (Secchi et al., 2016; Vervia Inc, 2019). This objective is only useful when used in conjunction with F2 to prevent the model from simply favouring the smallest possible designs. This is the third objective defined relative to ‘Reduce’ from the 9R framework.

With these four objectives and their method of evaluation defined, we can implement both multi-objective and single-objective optimization for the parametric model and compare the results.

### 2.3. OPTIMISATION METHOD

Our optimization model is built using the open-source python library, PYMOO (Blank, 2020). It provides libraries for easy implementation of a selection of SOO and MOO algorithms with a suite of benchmark problems for testing. We interface between the python-based optimization model and our parametric model using RhinoCompute, a Rhino software development kit. RhinoCompute allows us to call Rhino and Grasshopper functions within Python. While this combined framework is not yet common in comparison to existing optimization plugins for Grasshopper (Vierlinger, 2013), it permits access to open-source libraries and potential for faster computing with parallel solutions.

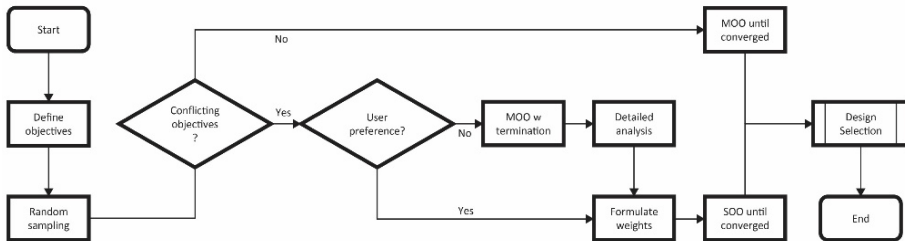


Figure 3. Flowchart for optimisation framework

The MOO problem is defined using the input parameters, evaluation functions (Figure 2) and a set of upper and lower boundaries. We modified an existing definition for the fast and elitist many-objective optimization algorithm, NSGA-II, to be used for our mixed-integer program. Finally, we set the algorithm to terminate after a fixed number of evaluations for comparison between results.

The SOO problem is defined similarly except for scalarizing the four objectives using weights. The weighted-sum approach is often preferred due to the clarity and reliability of optimising with a single dimension (Wortmann, 2020). Although there are many approaches to mathematically translate preferences into weights, we adopt a ranking method by allocating a weight of one for least important objectives and higher integer weights for more important objectives (Marler & Arora, 2010). We normalize each objective before weighting to capture the relative improvement to the performance.

Figure 3 shows the framework proposed by the authors, identifying when SOO and MOO are of greatest value in optimization for CE design. The framework is supported

by the results of the case study as presented in Section 3 below, including relative speed of convergence. While the framework is developed to target design for CE, it may be applied to design optimization more generally.

We employ five methods to visualize our results. First, we compare objective correlation matrices to preliminarily identify conflicting objectives. Secondly, with results from MOO we visualize the pareto-front in objective space to better understand trade-offs between conflicting objectives. Thirdly, we show the entire objective space using a pairwise plot. Fourth, we visualise convergence by plotting the hypervolume of each result set. Finally, we visualize a set of variants selected from the optimization process as outputs for the designer.

### 3. Results and Discussion

In this section we provide results from four sets of experiments, a random sample, a MOO run, and two SOO runs. First, by comparing correlation matrices derived from each set, we discuss the impacts of how the range of data distribution affects our understanding of conflicting objectives. Next, we analyse pairs of conflicting objectives by visualizing the specific pareto sets. Third, in conjunction with the previous method, we discuss findings from visualizing the full set of results in an exhaustive pair-wise plot of the entire objective space. Finally, we select best performing design variants from the different experiment sets as a summary of the optimization process.

#### 3.1. UNDERSTANDING TRADE-OFFS WITH CORRELATION

In a first step we seek to understand the presence of conflicting objectives in our model by creating a correlation matrix. The correlation matrices in figure 4 show the bivariate correlation between all pairs of objectives, indicating with negative numbers when improvement in one objective is likely to result in a decrease in another variable. The presence of negative correlation among objectives is relevant to optimization in general as it makes scalarized single-objective optimisation more complicated. Results presented in figure 4 are from four sets of experiment sets used consistently in this paper.

Our results show strong negative correlation between F2 and F3, with smaller negative correlations between F2 and F4 and F3 and F4. Trade-offs between 'reduce' objectives (F2, F4) and 'reuse' objectives (F3) are not unexpected as additional material can improve strength of the design while decreasing material-use efficiency. More complex is the trade-off between F2 and F4, which indicates that different definitions of the 'reduce' objective are at times contradictory.

Title	RDM1				MOO1				SOO1				SOO2			
# Sampled	1750				1750				1750				1750			
	F1	F2	F3	F4	F1	F2	F3	F4	F1	F2	F3	F4	F1	F2	F3	F4
F1	1.00				1.00				1.00				1.00			
F2	0.05	1.00			-0.03	1.00			0.58	1.00			0.56	1.00		
F3	-0.06	-0.38	1.00		-0.28	-0.42	1.00		-0.38	-0.44	1.00		0.03	-0.26	1.00	
F4	-0.03	-0.26	-0.18	1.00	-0.14	-0.41	-0.23	1.00	-0.22	-0.14	-0.28	1.00	-0.27	-0.38	-0.18	1.00

F1: Fabrication Waste  
 F2: Volume Ratio  
 F3: Total Displacement  
 F4: Embodied Carbon

RDM1: Solutions from a random sample set  
 MOO1: Solutions from multi objective optimization  
 SOO1: Solutions from optimizing with a weighted sum of [2,2,1,2]  
 SOO2: Solutions from optimizing with a weighted sum of [1,1,3,1]

Figure 4. Correlation matrices from four sets of experiments

Comparing the four correlation matrices drawn from 1) the results of a random sample set; 2) MOO1; 3) SOO1 and 4) SOO2 gives us a preliminary understanding of the complex manner in which correlations vary within the objective space (Figure 4). MOO1 has produced a relatively broad sampling of the objective space with results that are similar to the random sample. The two scalarized SOO runs, in contrast, show very clear deviations from random sample, MOO and from each other. In our results correlation between pairs of objectives demonstrates varying sensitivity to the distribution of the dataset. This implies that depending on the distribution of the objective targeted, designers will encounter different trade-offs among CE objectives.

Analysing similarly complex objective spaces, while unfamiliar in design fields, is addressed in statistics and data analytics. Sensitivity analysis is carried out to find how each independent variable affects other dependent variables, by applying techniques like clustering analysis and change point detection. Application of these methods exceeds the scope of the current paper, but we note that further research may help clarify the extent to which sensitivity analysis could assist in identifying best results in design for CE.

### 3.2. UNDERSTANDING TRADE-OFFS BY COMPARING PARETO FRONTS

Having identified pairs of conflicting objectives related to CE via the correlation matrix, we investigate trade-offs between these pairs of objectives in greater detail by generating two-dimensional visualizations of pareto solution sets from a multi-objective optimization. We chose the three objective pairs with highest correlation from figure 4 for further analysis. In figure 5, we plot the objective pairs of F2-F3, F2-F4 and F3-F4 using a novel visualization method.

To support the user's understanding of pareto-sets figure 5 combines a scatterplot of a 2D pareto set with a broader frame of reference provided by 1D plots of all objectives at full range. In the plot of F2-F3, we see the pareto solutions are close to converging. Even though they form a pareto front, these points represent a small fraction of the entire distribution for light-weighting and structural robustness, such that we can assume that all solutions shown on the 2D plots are almost equally well-performing. To further investigate these solutions for variability, the user could then use these subsets of points to plot another pareto front for the remaining objectives of



fabrication waste and embodied carbon.

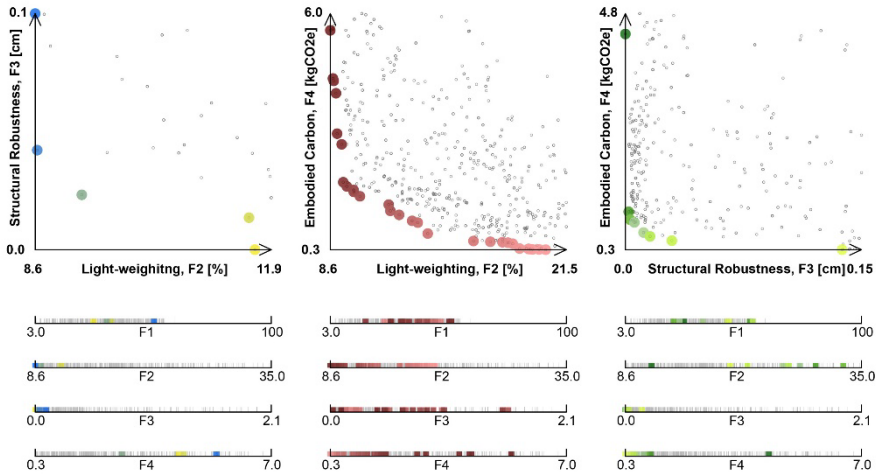


Figure 5. Unique Pareto front from each comparison of objectives. 2D plot shows dominated points in grey and non-dominated points in a colour gradient. The four 1D plots show where each point on the corresponding Pareto front lies on the full distribution of each objective.

### 3.3. COMPARING OBJECTIVE SPACE EVALUATION FOR SOO & MOO

With understanding of conflicting design objectives related to CE established through correlation matrix and Pareto-set visualizations (Section 3.1), we next employ scalarized Single-objective Optimization to target ‘best’ design variants within preferred areas of the objective space. Below we share results from two scalarized SOO runs which represent a hypothetical designer’s preference for a more reusable product (SOO2), or a product with better reduction of waste (SOO1). SOO1 targets ‘reduce’ by giving higher weightings to fabrication waste (F1), light-weighting (F2), and embodied carbon (F4). SOO2 targets ‘reuse’ by giving a higher weighting to structural robustness (F3). In this section we compare the relative benefits of SOO and MOO in identifying good circular design solutions based on effectiveness of objective space exploration and speed of convergence.

A first comparison between MOO and SOO highlights differences in objective space exploration by plotting all solutions from all evaluations in a set of 2-dimensional plots showing all pairs of objectives (Figure 6). In these results, design variants found by MOO cover a much wider range than those from SOO. Within the broad range of MOO results, the SOO results appear as punctual or linear clusters, converging at one, or in some cases two, small areas around local optima. The effect of changing the scalarized SOO weightings is visible in their separate clustering, visualized in red and green (Figure 6); with SOO1 prioritizing reduction in material, and SOO2 prioritizing reusability. Histograms for each column reinforce this finding, indicating the number of results within a given range for an objective: SOO1 and 2 demonstrate either one or two clear peaks in each histogram (Figure 6). MOO in contrast shows a normal distribution of results around a generally better area for all objectives.

In a second comparison of results, we look at how ‘quickly’ the two methods converge toward a solution over the course of a set number of evaluations. To make this comparison we chart the change in hypervolume. As the optimization algorithm converges on a solution set, the rate of change in the objective space slows, resulting in the plateauing of the hypervolume curve. Comparing the shape of the convergence curve between the two methods gives us a comparative indication of how well they can identify a clearly defined solution set. We note that comparing maximum hypervolume values between methods, is not pertinent to the question of convergence as the two optimization models have intrinsically different ranges.

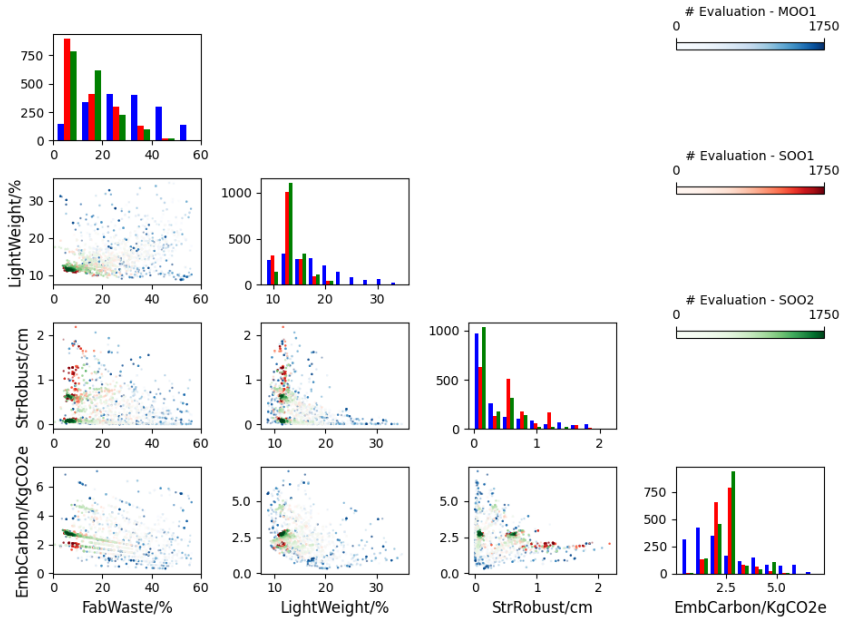


Figure 6. Pairwise plot of the objective space of all three result sets. Results from the two SOO sets are plotted by deconstructing the weighted sum back into individual objective scores. The histograms along the diagonal correspond to objective stated in the same column.

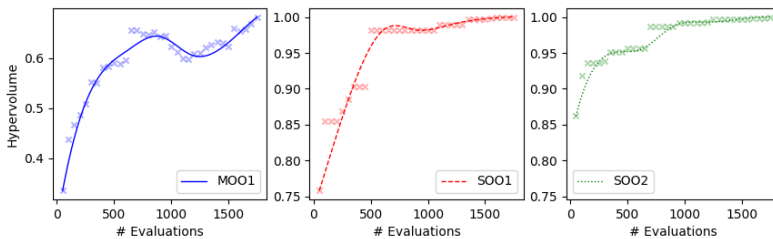


Figure 7. Evaluating for convergence using Hypervolume

Figure 7 shows that multiple runs of SOO with different weights result in a plateau around 1250 evaluations. MOO applied to the same model, however, does not converge toward a solution even after 1750 evaluations. These results suggest that while scalarized SOO will produce a clearly defined solution set relatively rapidly for our model, MOO will not converge toward a single solution over the same time frame. This result is pertinent to the question of design for CE: if no clear solution is available after a reasonable period, some users may find the method to be unsuccessful. When designers have established their understanding of the problem well-enough to assign weights, scalarized SOO is more efficient than MOO.

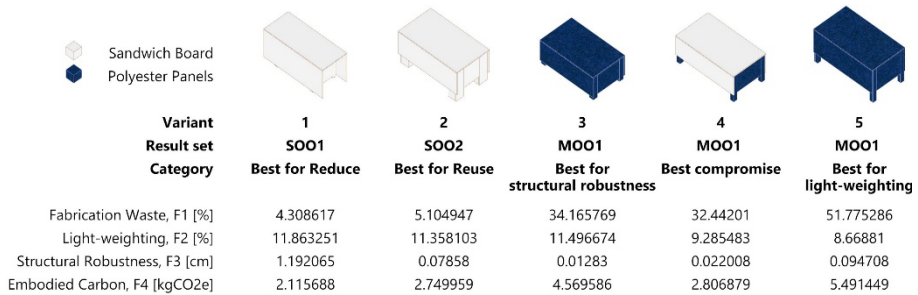


Figure 8. Summary of proposed variants

To draw further conclusions from our experiments, we tabulate a set of good results from the three experiments conducted. From the two SOO experiments, we select the single best performing results. From the MOO experiment we present in figure 8 a range of options pulled from the pareto-set for the pair of objectives with the largest negative correlation: light-weighting (F2) and structural robustness (F3). Variants 3, 4 and 5 shown in figure 8 are drawn from the two extreme ends and the middle, or ‘utopia’ point of the pareto curve (Figure 6).

The best variants drawn from SOO show clear advantages over the range of variants drawn from MOO. The design variant drawn from SOO1 (prioritizing the three ‘reduce’ objectives) is better performing in ‘reduce’ objectives than all other presented results. While not unexpected, the result is clear and reinforces the value SOO has brought to the case study. SOO2 (prioritizing robustness) produces a result that, while scoring highly in robustness, is not as high as two of the variants identified by MOO. Despite being slightly less optimal in this criterion, the best variant from SOO2 displays better performance in at least two other criteria in comparison to the MOO variants: showing the extent to which SOO can identify ‘across the board’ good results even when weightings skew toward specific preferences. The MOO results, in contrast, show mid-range variants between the two SOO options, with few visually distinguishing features. They are also notably poorly performing in fabrication waste (F1), reinforcing the inadequacy of making a design selection from a single 2D pareto front in a complex multi-dimensional objective space.

#### 4. Conclusion

This paper presents an example of how optimization can be used to address complexity presented by design for CE. With four objectives drawn from the 9R

framework, we observe conflict between multiple pairs of objectives, with substantial variation in negative correlation depending on which portion of the objective space is analysed. In response to this complexity, we propose a methodology that assists designers to understand the objective space, clarify preferences and then apply optimization. Random sampling is used to observe overall trade-offs between objectives, MOO is applied to understand nuances of trade-offs using the proposed visualization methods. To the informed designer, SOO then shows decided advantages in performance and clarity of results.

There remain aspects in optimizing design for CE to be improved. More research is needed to formulate quantifiable objectives, covering a wider spectrum of established CE design strategies. More studies are needed to understand conflicting objectives with the help of computational optimisation, explaining general rules and conflicts in CE. More broadly, there is a need for sensitivity analysis methods that can help designers narrow in on specific areas of CE objective space. Methods that balance between targeted and global design search will be of great value to support quantitative definition of designer's preferences and identification of best design results for CE.

## References

- Blank, J., & Deb, K. (2020). Pymoo: Multi-Objective Optimization in Python. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2020.2990567>
- Ellen MacArthur Foundation. (2015). *Circularity Indicators: An approach to measuring circularity*. Retrieved from [https://www.ellenmacarthurfoundation.org/assets/downloads/insight/Circularity-Indicators\\_Methodology\\_May2015.pdf](https://www.ellenmacarthurfoundation.org/assets/downloads/insight/Circularity-Indicators_Methodology_May2015.pdf)
- Marler, R., & Arora, J. (2010). The weighted sum method for multi-objective optimization: New insights. *Structural and Multidisciplinary Optimization*, 41, 853–862. <https://doi.org/10.1007/s00158-009-0460-7>
- Reike, D., Vermeulen, W. J. V., & Witjes, S. (2018). The circular economy: New or Refurbished as CE 3.0? — Exploring Controversies in the Conceptualization of the Circular Economy through a Focus on History and Resource Value Retention Options. *Resources, Conservation and Recycling*, 135, 246–264. <https://doi.org/10.1016/j.resconrec.2017.08.027>
- Saidani, M., YANNOU, B., Leroy, Y., Cluzel, F., & Kendall, A. (2019). A taxonomy of circular economy indicators. *Journal of Cleaner Production*, 207, 542–559. <https://doi.org/10.1016/j.jclepro.2018.10.014>
- Secchi, S., Asdrubali, F., Cellai, G., Nannipieri, E., Rotili, A., & Vannucchi, I. (2016). Experimental and environmental analysis of new sound-absorbing and insulating elements in recycled cardboard. *Journal of Building Engineering*, 5, 1–12. <https://doi.org/10.1016/j.jobte.2015.10.005>
- UNEP. (2019, October 11). *Circularity platform: Circularity framework*. UNEP - UN Environment Programme. Retrieved from <http://www.unep.org/circularity>
- Vervia Inc. (2019). *Environmental Product Declaration: EzoBord PET Accoustic Panels*. Vervia Inc. Retrieved from [https://cdn.scs-certified.com/products/cert\\_pdfs/SCS-EPD-05286\\_EzoBord\\_022619.pdf](https://cdn.scs-certified.com/products/cert_pdfs/SCS-EPD-05286_EzoBord_022619.pdf)
- Vierlinger, R. (2013). *Multi Objective Design Interface*. Master Thesis. Institut für Hochbau und Technologie. <https://doi.org/10.13140/RG.2.1.3401.0324>
- Wortmann, T., & Fischer, T. (2020). Does Architectural Design Optimization Require Multiple Objectives? *Proceedings of the 25th CAADRIA Conference*. CAADRIA.

# PL-SYSTEM: VISUAL REPRESENTATION OF PATTERN LANGUAGE USING L-SYSTEM

JUNAH YU<sup>1</sup> and DEEDEE MIN<sup>2</sup>

<sup>1,2</sup>*Department of Interior Architecture Design, Hanyang University.*

<sup>1</sup>*abcdmi0115@hanyang.ac.kr; 0000-0002-6765-4813*

<sup>2</sup>*mindeed@hanyang.ac.kr; 0000-0002-6651-5133*

**Abstract.** Pattern Language provides simple and conveniently formatted solutions to complex design problems ranging from urban planning to interior design for community-led inclusive designs. Despite the intention, the concept has been more widely adopted by computer science professionals. One possible reason is the lack of visualization, making it difficult to be used interactively by the non-professionals. To overcome the issue, we aim to integrate the patterns from Pattern Language into the L-system to visualize the paper architecture into geometric forms. Specifically, we implement Pattern L-System (PL-System), which generates diagrammatic floor plans using design rules based on the pattern languages. We first made analogical comparisons of the concepts and grammar structure between Pattern Language and the L-system. Next, we defined a geometric interpreter for drawing diagrammatic floor plans using turtle graphics, which consist of geometrical rules for putting shapes together. Then, we selected three patterns and reinterpreted them for visualization using strings of turtle graphics letters that determine the turtle's movements for the geometric representation of walls, columns, and doors. From this research process, we learned that Modular L-system opens up the possibility for the visualization of the patterns in Pattern Language.

**Keywords.** Pattern Language; L-System; Diagrammatic Floor Plan; Turtle Graphics; Geometric Forms; SDG 11.

## 1. Introduction

Christopher Alexander's Pattern Language Theory (1977) was originally created for design purposes. Pattern Language contains rules for how human beings interact with built forms reflecting different lifestyles, customs, and behaviours (Salingaros, 2014). This characteristic of pattern language allows it to be applicable for comprehensive design language. In fact, research was conducted that evaluates the applications of patterns from Pattern Language for universal design, including accessible, green, and public space (Lee, 2015; Iwańczak & Lewicka, 2020).

While Pattern Language is useful conceptually in architectural design, they present a lack of usage in architectural design and planning in practice mostly because they are

semantic and abstract, lacking in objectivity. From these characteristics, two issues arise when incorporating Pattern Language in design practice. The first issue is about the ambiguity of network structure patterns connections and the interrelationship between the patterns. In order to resolve this issue, a sequential approach has been proposed (Porter et al., 2005; Ribeiro, 2020). The second issue is about the lack of visual aid. As pointed out by Ribeiro (2020), visual aids such as photos or illustrations are essential when adopting Pattern Languages into a design process. However, this problem is left largely unsolved. Therefore, in this study, we tackle the visualization problem to increase the applicability of Pattern Language into design practice for all-embracing community designs.

To achieve this aim, we incorporate the concept of L-system to Pattern Language. In particular, we focus on creating a floor plan diagram using a Modular L-System with rules based on the patterns from Pattern Language. In order to achieve this goal, we first made analogical comparisons of the concept, characteristics, and grammar structure between Pattern Language and the L-system. By observing the two concepts theoretically, we were able to identify that they have different grammatical structures, which serves as a limitation when coupling the two. While the L-system is a parallel rewriting system and recursive, Pattern Language is serial and has a network structure. To overcome this limitation, we incorporated the concept of the Modular L-system proposed by Cieslak et al. (2011) as a strategy to develop an L-system using separate modules to represent different aspects. After, we defined a geometric interpreter for drawing floor plan diagrams using turtle graphics. Turtle graphics consist of geometrical rules for putting matter together. Turtle graphics can perform a geometric interpretation of the new L-string in each derivation step upon request. Next, in order to incorporate Pattern Language into the system, we selected three patterns applicable for rules from Pattern Language. The patterns were reinterpreted and represented using turtle graphics letters. At this stage, we were able to create modules that would become the new definition of the pattern using Turtle graphics letters. Finally, we implemented the PL-System using Processing language. From this research, we aim to answer whether L-System is applicable for opening up the possibility of extending the use of Pattern Language in the architecture field by visualizing semantic patterns into geometric forms facilitating the use for comprehensive community design by the non-professionals.

## **2. Literature Review**

### **2.1. PATTERN LANGUAGE**

Christopher Alexander has been a leading pioneer of academic research on the architectural and urban design since the early 1960s (Galle, 2020). Alexander addressed the problem of capturing recurring problems and their solutions in the context of civil architecture in the 1970s. In his work, Alexander struggled with the need to document and share architectural knowledge, which could be easily applied by his fellow architects. The theory of Pattern Language is a design theory for deriving 253 patterns from physical environments with design problems and solutions. According to the theory, because good spaces have regular and timeless patterns, professionals or ordinary people can design new spaces by modifying or combining

the patterns. Unlike the original intention of the Pattern Language, it has been applied to mostly computer science fields, including user interface configuration and evaluation of web services (Pautasso et al., 2016).

### 2.2. L-SYSTEM

In 1968, biologist Aristid Lindenmayer proposed a string-rewriting system that can model simplified plants and their growth process. This theory is now known as L-Systems (Hansmeyer, 2003). An L-System is a recursive and formal grammar system that rewrites collections of characters into new strings based on a certain set of rules. By iterating a number of times over the string, the character arrangement changes. Figure 1 illustrates the rewriting of a string by a set of rules. With the L-System, minimal inputs can create a complex output.

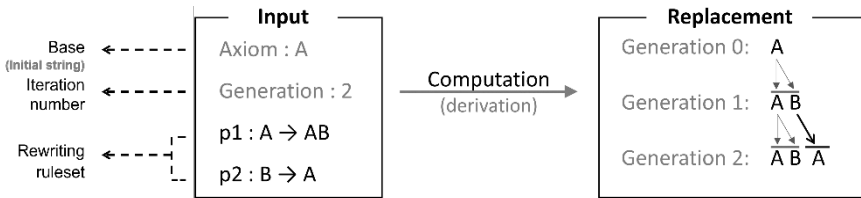


Figure 1. Rewriting string of L-system

#### 2.2.1. Geometrical Interpretation of Strings: Turtle Graphics

In 1986, in order to model higher plants, Prusinkiewicz (1986) focused on the geometrical interpretation based on a LOGO-style turtle. In turtle graphics, a fictive turtle follows movement commands that correspond to individual alphabets and symbols of a string. The turtle’s path is thus a visual interpretation of the string. F in turtle graphics means to move forward a step while drawing a line, + means to turn right, and - means to turn left at assigned angles. Figure 2 shows an example of the geometric interpretation of the L-System using turtle graphics. The initial string, rule, and rotation angle parameters produce a form that resembles a Sierpinski Triangle:

- Initial string: F
- Rule:  $F \rightarrow F + F - F + F$
- Rotation angle:  $60^\circ$

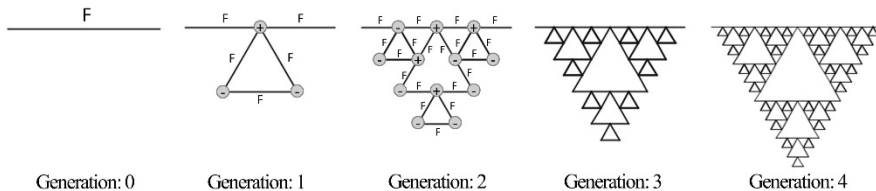


Figure 2. Sierpinski triangle

### 2.2.2. Modular L-System

Cieslak et al. (2011) proposed a strategy to develop a modular L-system based on the use of separate modules to represent different parts. As not every alphabet of the L-system needs to correspond to a turtle command, new characters can be used that are simply replaced by groups of other symbols and alphabets (Hansmeyer, 2003). For example, with the rewriting rule of  $C \rightarrow F+F+F+F$  where F implies a movement forward and + implies a right turn, a new character C could be defined to create a square, as illustrated in Figure 3.

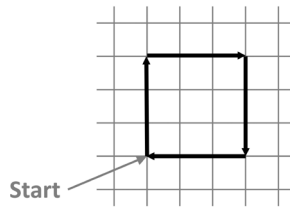


Figure 3. Example of modular form

## 2.3. PATTERN LANGUAGE AND L-SYSTEM

In Section 2, we reviewed the characteristics of the Pattern Language and the L-System. Pattern Language is useful for interpreting complex socio-spatial considerations through a simple building-block format, which makes this content accessible to non-professionals (Dawes & Ostwald, 2017). Meanwhile, using the L-system with a set of rules, simple initial objects can be successively replaced, resulting in a more complex final object (Hansmeyer, 2003). Through this process, L-systems can generate two-dimensional and three-dimensional geometries efficiently that otherwise would have taken an extended amount of time (Serrato, 2005). As discussed earlier, these two concepts have different grammatical structures—one being serial with a network structure and the other one being parallel with recursive structure—serving as a limitation when coupling Pattern Language and the L-system. Despite this limitation, the use of the L-system should be considered, given that, it is highly scalable and portable, consisting of scripts that can be parsed and easily rewritten. By expanding the logic to include independent parameters, we can gain a high degree of control over the produced form (Hansmeyer, 2003).

## 3. System Architecture

### 3.1. SYSTEM FLOW

The implemented PL-System largely consists of three modules. One is the turtle graphics module for visualization, another is the pattern definition module for the extraction of patterns from Pattern Language, and the third module is the rule production module for the definition of the drawing rules according to the patterns extracted. These three modules are linked into an iterative system flow, as shown in Figure 4. First, the system initializes the parameters of turtle graphics with start length, start thickness, and rotation angle. Next, the system creates patterns using the inputs by



the user, which are based on the patterns from Pattern Language. This process of adding patterns can be iterative as long as the user desires. When the patterns are set, the rule production module is set up. Using the rule production module, rewriting of the rules are iterated until the user's needs.

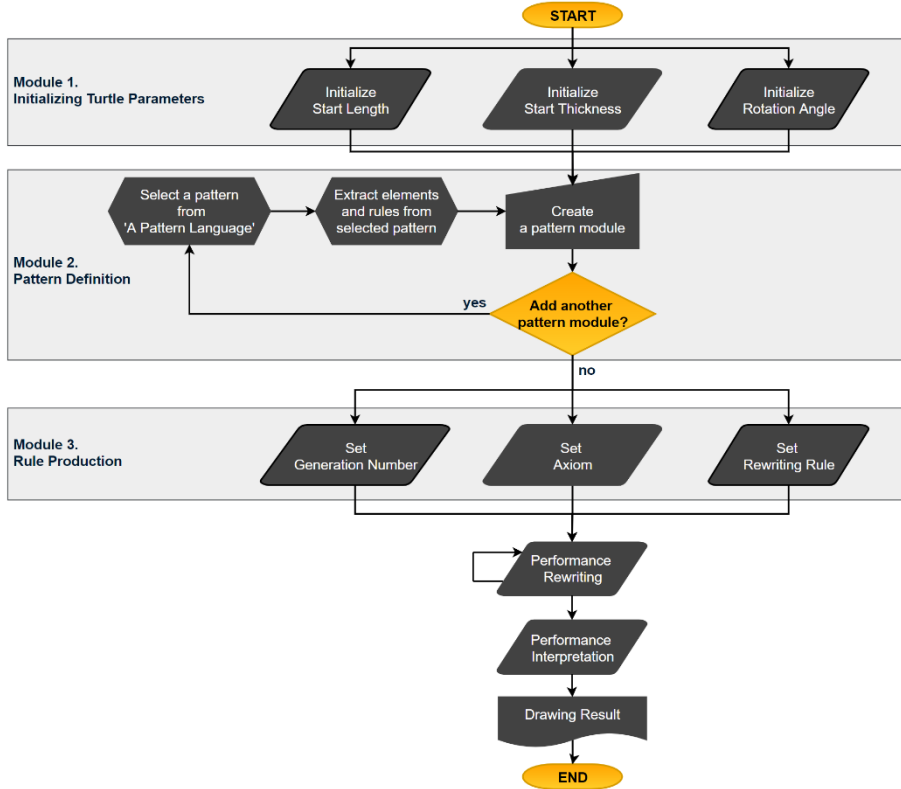


Figure 4. The System architecture of the PL-System

### 3.2. INITIALIZING TURTLE GRAPHICS

As mentioned above, the system initializes the parameters of turtle graphics using the user inputs of start length, start thickness, and rotation angle. Each of the inputs is used to define the turtle commands, as summarized in Table 1. For example, the wall element is represented using the character w, and its command is to move forward while drawing a line with the user-defined parametric length and thickness. Or, for the corner element, the character + commands the turtle to turn right at a user-defined angle. Using these turtle commands, the patterns from Pattern Language can be visually represented.

Table 1 Turtle commands

Elements	Characters	Commands	Parameters
----------	------------	----------	------------

door	f	move forward	Length
wall	w	move forward while drawing a line	Length, Thickness
column	r	move forward while drawing a rectangle	Width
comer	+	turn right	Angle
comer	-	turn left	Angle
-	[	push state	-
-	]	pop state	-

### 3.3. PATTERN DEFINITION

Among the patterns from Pattern Language, we chose three patterns specifically related to the layout of a floor plan consisting of walls, columns, and doors. Then from the semantic information provided by Pattern Language, we redefined the patterns using the characters of the turtle graphics, which provides pattern modulation. Then the modulated pattern can be instantiated as a pattern module and be stored in the system, and be used later in the rewriting process of PL-System.

In detail, we selected the patterns 196, 216, and 212, which are Corner Doors, Box Columns, and Columns At The Corners, respectively. For the redefinition process, we extracted the related problems and their corresponding solutions. Table 2 represents our redefinition results:

Table 2 Redefinition of patterns

Patterns	Problems	Solutions	Redefinition	Pattern Modules
196. Corner Doors	"If the doors create a pattern of movement which destroys the places in the room, the room will never allow people to be comfortable" (Alexander, 1977, p.904).	"In most rooms, especially small ones, put the doors as near the corners of the room as possible" (Alexander, 1977, p.905).	Place door(d) next to corner (+ or -)	A(1)=f+ A(2)=f- A(3)=+f A(4)=-f
216. Box Columns	"Columns feel uncomfortable unless they are reasonably thick and solid...[and] must be easy to connect to foundations, beams and walls" (Alexander, 1977, p.1013).	"A column which has all these features is a box column" (Alexander, 1977, p.1015).	Place box shape(r) columns	B = r
212. Columns at The Corners	"The necessities of the drawing itself change the plan, make it more rigid, turn it into the kind of plan which can be drawn and can be measured" (Alexander, 1977, p.990).	"On your rough building plan, draw a dot to represent a column at the corner of every room..." (Alexander, 1977, p.993).	Place column(B) next every corner (+ or -)	C(1)=B+ C(2)=B- C(3)=+B C(4)=-B

### 3.4. RULE PRODUCTION

After the pattern definition, rules can be produced for each of the pattern module IDs and the parameters initialized in turtle graphics (start length, start thickness, and rotation angle). The rules at this stage of the system aim to diagram floor plans by mainly converting the input parameters that describe the floor plan morphology into L-system parameters. For the rule production, three distinct inputs are required, which are the initial string, one or more substitution rules, and the number of times to perform the rewriting operation.

The rewriting process is divided into two parts. In the first rewriting process, the final sentence is returned through the rewriting process by applying the rewriting number, axiom, and ruleset input by the user. The system performs the rewriting process by setting the returned final sentence as an axiom of the new rewriting process and pattern modules as rulesets. During the second rewiring process, the system visualizes the final sentence derived through two rewriting processes on the display area of the floor plan through turtle interpretation.

## 4. System Design

### 4.1. USER-SYSTEM INTERACTION

As shown in Figure 7, the PL-System receives the turtle parameters, pattern module, and the production rules from the user to perform rewriting. Then after the calculation, the graphical interpretation of the patterns along with the axiom and the rewriting rules are displayed to the user.

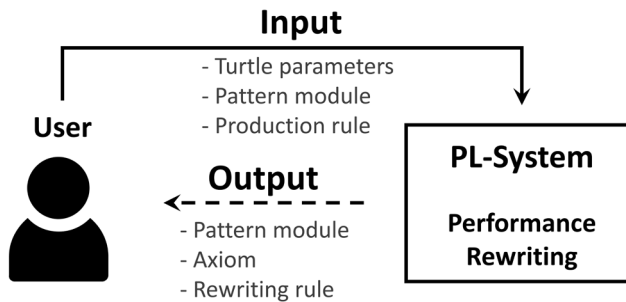


Figure 7. User-system interaction

### 4.2. GUI (GRAPHICAL USER INTERFACE)

In order for the abovementioned user-system interaction to occur with high understandability, a well-designed GUI is necessary. Therefore, as illustrated in Figure 8, the developed PL-System enables users to control the system through three different distinctive inputs and outputs for each of the modules: the pattern module, production

rule, and turtle graphics. For example, for the pattern module (Area A), pattern ID and the pattern components are received from the user. Then, these two inputs are combined, saved, and represented as a list of pattern modules (string) in Area E. As for the production rule input (Area B), the number of generations, axiom (string), and rulesets is received. Then the geometric interpretation of axiom and the rewriting rule is displayed in Areas F and G, respectively. The advantage is that the user can visualize the geometric shapes for the axiom and rewrite rules in real-time. Finally, for the turtle graphics input (Area C), the starting length, thickness, and rotation angle are received from the user. Integrating the data received from Area A and B, the system performs the rewriting and visualizes the floor plan accordingly in Area D.

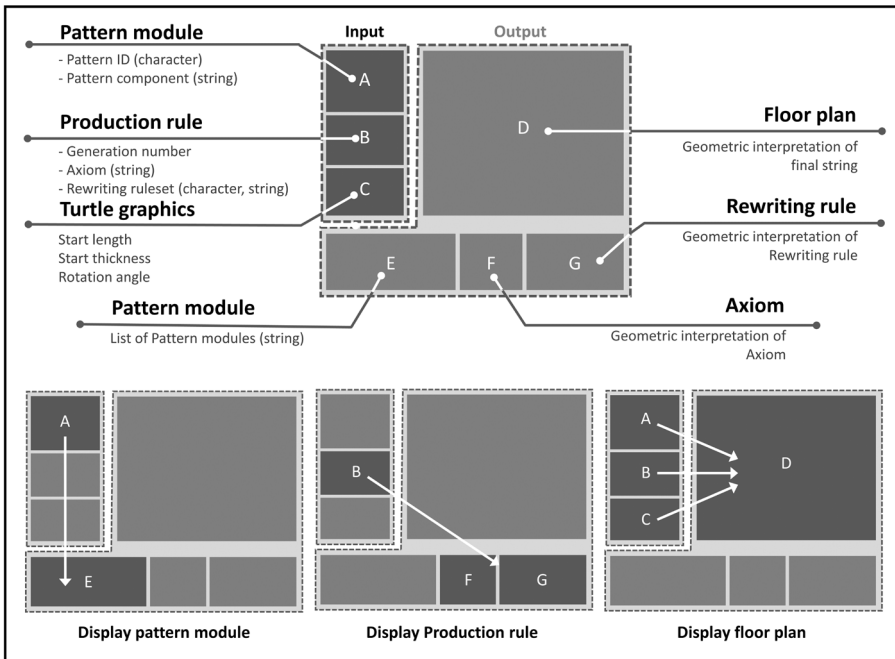


Figure 8. GUI layout

### 4.3. SYSTEM DEMONSTRATION

To demonstrate how our PL-System can graphically visualize a user-defined pattern based on the production rule derived from the redefinition of Pattern Language, we used all the pattern modules mentioned in Table 2 and set the generation number as two. Then, the axiom was defined as (w+w), which commands to draw a wall at a given length, turn at a given angle, and draw another wall at the same length as shown in Figure 9 inside the axiom display Area F. Then, the axiom can be elaborated into a form illustrated in the display Area of D using the user-given production rule and turtle parameters where the start length is 100, thickness is 20, and the rotation angle is 90°.

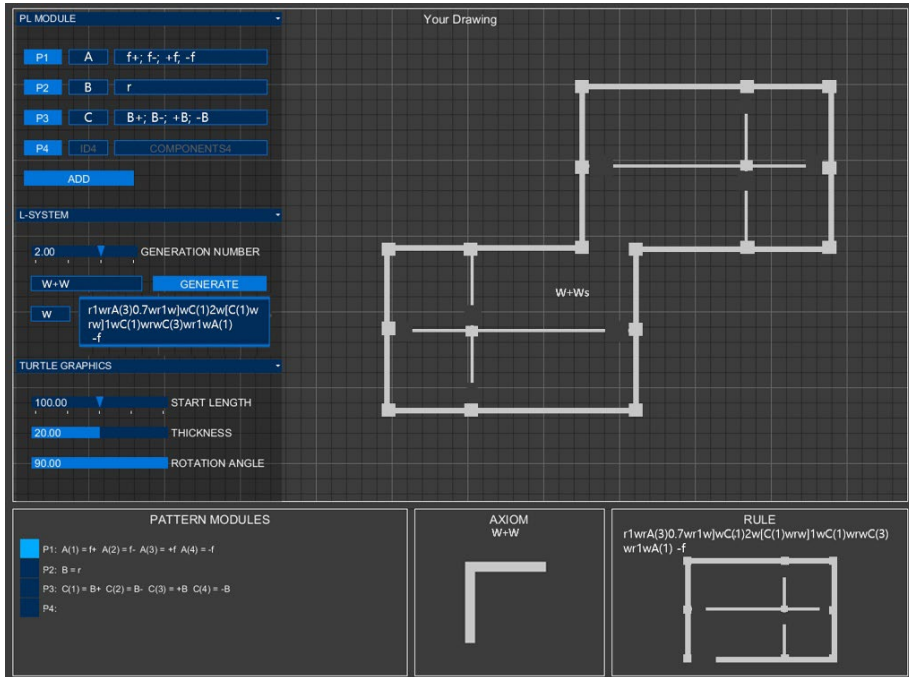


Figure 9. Demonstration of PL-System

## 5. Discussion

This paper proposes the possibility of integrating the L-System for the graphical interpretation of the patterns from A Pattern Language. The fundamental advantage of the L-System is that minimal inputs can create a spatially complex output. By expanding the logic to include independent parameters, we can gain a high degree of control over the produced form. According to the original intention of the Pattern Language Theory, users should be able to transform the given patterns for their own use. Along with this flexibility, Pattern Language has high scalability, applicable for participatory and inclusive sustainable design for small to large scale design projects. However, because the patterns are semantic, the non-professionals have difficulty transferring the knowledge into practice. In response, we proposed a PL-System that is able to redefine patterns with characters and create a diagrammatic floor plan

composed of walls, columns, and doors based on the generative concept of the L-System.

The PL-System proposed and developed in this paper meets the theoretical requirement of Pattern Language, which is the transformability with a high degree of control over the produced form but at the same time, deriving many patterns under certain parameters. The use of the PL-System in this paper is tested using the simple drawing of architectural elements such as walls and columns. By visualizing the patterns into geometric forms and using the recursive functions, we expect that the system will aid non-architects to understand the relationship between architectural elements and spaces visually when designing a public space for sustainable community. Further research will focus on the verification of the usefulness and usability of the system through user tests with both professionals and non-professionals.

## References

- Alexander, C. (1977). *A Pattern Language: Towns, Buildings, Construction*. Oxford university Press.
- Cieslak, M., Seleznyova, A. N., Prusinkiewicz, P., & Hanan, J. (2011). Towards aspect-oriented functional–structural plant modelling. *Annals of Botany*, 108(6), 1025-1041.
- Dawes, M. J., & Ostwald, M. J. (2017). Christopher Alexander's A Pattern Language: analysing, mapping and classifying the critical response. *City, Territory and Architecture*, 4(1), 1-14.
- Galle, P. (2020). Christopher Alexander's Battle for Beauty in a World Turning Ugly: The Inception of a Science of Architecture? *She Ji: The Journal of Design, Economics, and Innovation*, 6(3), 345-375.
- Hansmeyer, M. (2003). L-systems in architecture. *Distinguishing Digital Architecture in 6th Far Eastern International Digital Architecture Design Award, Taiwan*. Retrieved December 06, 2021, from <https://www.michael-hansmeyer.com/l-systems>.
- Iwańczak, B., & Lewicka, M. (2020). Affective map of Warsaw: Testing Alexander's pattern language theory in an urban landscape. *Landscape and Urban Planning*, 204, 103910.
- Lee, H. G., & Lee, J. H. (2015). A Study on the Spatial Characteristics Analysis of the Urban Public Space, applying a Pattern Language. *Journal of the Korea Academia-Industrial Cooperation Society*, 16(8), 5608-5618.
- Pautasso, C., Ivanchikj, A., & Schreier, S. (2016, July). A pattern language for RESTful conversations. In *Proceedings of the 21st European Conference on Pattern Languages of Programs* (pp. 1-22).
- Porter, R., Coplien, J. O., & Winn, T. (2005). Sequences as a basis for pattern language composition. *Science of Computer Programming*, 56(1-2), 231-249.
- Prusinkiewicz, P. (1986, May). Graphical applications of L-systems. In *Proceedings of graphics interface* (Vol. 86, No. 86, pp. 247-253).
- Pinho, D. (2020). *Effective Communication in Agile Teams: A Pattern Language*. Masters Dissertation, Universidade do Porto.
- Salingaros, N. (2014). *A theory of architecture part 1: Pattern language vs. form language*. *Archdaily*. Retrieved December 06, 2021, from <http://www.archdaily.com/488929/a-theory-of-architecture-part-1-pattern-language-vs-form-language>.
- Serrato-Combe, A. (2005). Lindenmayer systems–experimenting with software string rewriting as an assist to the study and generation of architectural form. In *SIGraDi 2005: Proceedings of the 9th Iberoamerican Congress of Digital Graphics, Lima – Peru* (vol. 1, pp. 161-166).

# **Artificial Intelligence and Machine Learning in Design**





# WHERE WILL ROMANCE OCCUR

*A New Prediction Method of Urban Love Map through Deep Learning*

ZHIYONG DONG<sup>1</sup>, JINRU LIN<sup>2</sup>, SIQI WANG<sup>3</sup>, YIJIA XU<sup>4</sup>, JIAQI XU<sup>5</sup> and XIAO LIU<sup>6</sup>

<sup>1</sup>*South China University of Technology.*

<sup>2,3</sup>*Tongji university.*

<sup>4</sup>*Shenzhen University.*

<sup>5</sup>*National University of Singapore.*

<sup>6</sup>*University of Southern California.*

<sup>1</sup>*ar\_dongzhiyong@mail.scut.edu.cn, 0000-0003-0958-2986*

<sup>2</sup>*jadya.lin@gmail.com, 0000-0002-4642-7953*

<sup>3</sup>*siqi\_wang@tongji.edu.cn, 0000-0003-0250-4619*

<sup>4</sup>*xuyijia2017@email.szu.edu.cn, 0000-0002-9620-3480*

<sup>5</sup>*e0709291@u.nus.edu, 0000-0002-0408-2432*

<sup>6</sup>*xliu7399@usc.edu, 0000-0001-8643-0009*

**Abstract.** Romance awakens fond memories of the city. Finding out the relationship between romantic scene and urban morphology, and providing a prediction, can potentially facilitate the better urban design and urban life. Taking the Yangtze River Delta region of China as an example, this study aims to predict the distribution of romantic locations using deep learning based on multi-source data. Specifically, we use web crawlers to extract romance-related messages and geographic locations from social media platforms, and visualize them as romance heatmap. The urban environment and building features associated with romantic information are identified by Pearson correlation analysis and annotated in the city map. Then, both city labelled maps and romance heatmaps are fed into a Generative Adversarial Networks (GAN) as the training dataset to achieve final romance distribution predictions across regions for other cities. The ideal prediction results highlight the ability of deep learning techniques to quantify experience-based decision-making strategies that can be used in further research on urban design.

**Keywords.** Romance Heatmap; Generative Adversarial Networks; Deep Learning; Big Data Analysis; Correlation Analysis; SDG 11.

## 1. Introduction

### 1.1. RESEARCH BACKGROUND

Urban design is essentially about place making, the creation of a distinct identity and memory for a place is an important attribute of successful urban design (Buchanan, 1988; Adams & Tiesdell, 2012; Loring, 2014). Romance, as a special kind of place semantics, is often closely tied to the place identity and memory. Norberg-Schulz (1980) regarded romance as one of the four types of man-made environment and considered it to be inseparable from the urban morphology. Romance places may serve as carriers for urban social and economic activities that are critical to urban life. A typical example is “Made in Marxloh” in Duisburg, Germany, the highly romantic street in Europe evolved from a poverty-stricken area, where romantic place making has directly led to rising rents and vitality (Tissink, 2016). As a result, romance should be a more prominently cultivated feature in the construction of urban space, and supporting romance and love should be a goal of the urban designer (Zukin, 1998). However, with the huge impact of the information flood brought by the mobile internet on urban life, romance seems to be “hidden” in cities (Rahimi et al., 2017). People living in cities are increasingly expressing romantic relationships on the Internet, so data from the Internet is becoming more and more important in the construction of urban spaces. In addition, providing visual guidance through maps can also help people explore romantic places and enrich dating ideas (ESRI UK, 2019; VISIT LONDON, 2021), which will encourage location-based scenario socialization. From this, the following research questions will be addressed in the paper: 1) How to express romance and love in cities? 2) What are the factors that affect the occurrence and distribution of urban romance? 3) In order to enhance the romance of cities through reasonable urban design, how to visualize and predict the distribution of urban romance?

### 1.2. PROBLEM STATEMENT AND PROJECT GOAL

Urban romantic relationship is still an understudied aspect of cities. Traditional methods such as questionnaires and interviews can help explore the public’s perception of romantic spaces, while they are time-consuming with only a small-scale sample investigated. Social media allows “individuals to communicate their impressions and interests to the general public” (Croitoru et al., 2013), which provides us with comprehensive insight into urban romance and love. Sina Weibo check-in data with user-generated content and geo-location coordinates enables the connection between user perception and city location, on this basis, urban romantic spots can be further screened by the matching of Weibo content and proposed romance dictionary. Combined with heat mapping, the distribution and aggregation of spatial romance can be visualized in a more intuitive way. However, the heatmap only serves as a visual analytic method based on historical or current data, which can neither predict the future romantic distribution, nor provide data supplement for suburbs or small cities with sparse check-in data. Therefore, the ability of visual analytics is too limited to give evidence-based recommendations for a more reasonable cross-regional urban planning or business location selection. The urban romance itself has dynamic attributes and is greatly affected by complex urban morphological features, which cannot be well

modeled through explicit programming and simple linear regression calculation. Such issues, can be addressed to a great extent, as the rise of artificial intelligence represented by deep learning techniques. Deep learning can cope with data-driven deep correlations modelling by multiple nonlinear transformations, which potentially provides a way for the efficient representation and accurate prediction of urban romance.

In this paper, as a typical image-to-image translation method, Generative Adversarial Networks (GAN) are adopted (Goodfellow et al., 2014) to automatically learn the correlations between romantic places and urban space, so as to achieve visual predictions. GAN, a popular deep learning model, has been gradually applied in urban planning and design realm due to its powerful spatial information processing and generation ability (Shen et al., 2020; He & Zheng, 2021). However, the training of GAN is often associated with convergence difficulties and high uncertainty. To increase prediction accuracy and reduce training time, we incorporate a priori knowledge into the input heatmap of romance distribution, i.e., the environments and various functional architectures highly relevant to urban romance are labelled with different colors and shades of grey. Our model has achieved ideal prediction results.

This paper proposes a GAN-based method to realize and verify the relationship between spatial romance and urban morphology, and provides a prediction for the future urban romance distribution across regions (Figure 1), which is of great value for inferring the implementation effect of various urban design plans on urban romance in advance, as well as the commercial site decision-making and space operations.

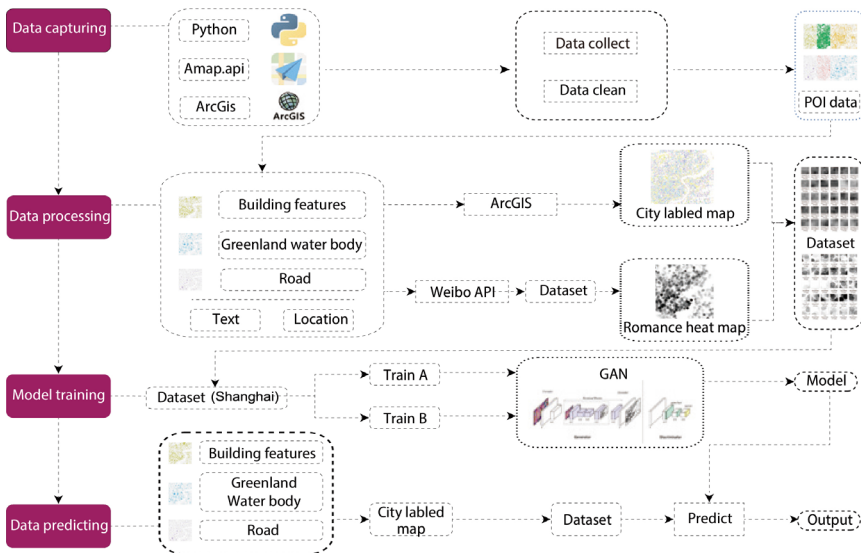


Figure 1. Overall Workflow

## 2. Methodology

### 2.1. GENERATIVE ADVERSARIAL NETWORKS

Deep learning is often a synonym for deep neural networks, which are the neural networks consisting of more than three layers (including inputs and outputs), implementing computational models by simulating the neural central system of an organism. In this paper, according to the experimental objectives, Generative Adversarial Networks (GAN) based on image processing are selected in the experiment. An adversarial neural network is a special adversarial process in which two neural networks compete with each other. The first network generates the data; the second network tries to distinguish the real data from the fake data created by the first network. The two play against each other to achieve a balance in game theory.

The first network is called the generator,  $G(z)$ , and the generator network takes random noise as input and attempts to generate sample data that is provided to the second network. The second network is called the discriminator,  $D(x)$ . The discriminator network takes real or generated data as input and tries to predict whether the current input is real or generated data. One of the inputs  $x$  is obtained from the real data distribution  $p_{data}(x)$ , which next solves a binary classification problem and produces a scalar ranging between 0 and 1. At the equilibrium point, which is also the optimal solution of the minimax Game, the discriminator considers the probability that the data which the generator output is the real data to be 0.5:

$$\min_G \max_D V(D, G) = E_{x \sim p_{data}(x)} [\log(D(x))] + E_{z \sim p_z(z)} [\log(1 - D(G(z)))]$$

Moreover, GANs are the neural networks based on input images and output images. The amount of love messages studied in this paper is extremely large, and each map may contain hundreds of data points of love messages. Also, to accommodate the visualization design habits of architects and to facilitate further regional optimization, we visualize the geo-location of romance messages and present them in a more intuitive heatmap rather than inputs the numerous data directly. Thus, we try to train the GAN to recognize the relationship between the two sets of images by taking the city map as input and the romance heatmap as output. The trained neural network model will be able to predict the future local love heat situation based on the existing city situation.

## 2.2. THE GENERATION OF ROMANCE HEATMAP

We select the cities in the Yangtze River Delta region as experimental subjects: Shanghai, Jiading, Hangzhou, and Ningbo as the training dataset cities, and Shanghai new towns, Nanjing, and Suzhou as the test dataset cities. The cultural characteristics and the economic development level of them are similar. The situations have certain regularities.

We first use Weibo data to create a romance heatmap. And use web crawlers to obtain web data. In this paper, we use Python+Selenium specifically for the Sina Weibo open platform API to receive data. Then we use a natural language processing model, WantWords Reverse Dictionary (Qi et al., 2020), to construct an emotion dictionary related to love messages. Love messages are applied to identify all synonyms related to romance, including 101 romance-related words such as dating, sweetness, and anniversary.

In the content filtering stage, taking Shanghai as an example, 8445 tweets and geo-located points are extracted from the original dataset whose dates are July-October

2021. We then use Rhino and Grasshopper to visualize the geospatial distribution of a large amount of data using colour shades(Figure 2). The visualization of filtered Weibo data can clearly show the distribution of romantic places.

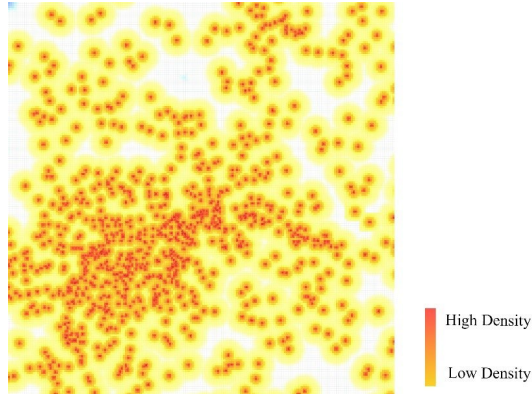


Figure 2. Romance heatmap

### 2.3. CORRELATION ANALYSIS AND THE COLLECTION OF CITIES LABELLED MAPS

We speculate that people's behavior of posting romantic messages is related to some urban morphological features. In this regard, we apply a correlation analysis between the geo-location of romantic messages and building features and selected Shanghai as the analysis object. We use the following three kinds of data, the romantic message points, the graphic of the street administrative divisions, and points of interest (POI) with function information collected on AutoNavi Maps. We count romantic message points and city POI of different functional types in each street division by ArcGIS.

Pearson correlation coefficient is used to conduct bivariate correlation analysis on them. The results show that the geo-location of romantic messages is strongly correlated with the features of 13 of the 16 categories (Table 1). Hotel and other accommodation services are most relevant to romantic messages, followed by shopping services and scenic spots. The commercial residences POI, which has the largest number, is extremely weakly related to romantic messages, which eliminates the interference of population concentration on the correlation results. The above results show that people's romantic-like social-emotional expression is related to urban design.

Poi function type	PCCs	Sig. (2-sided)	case number
Accommodation services	0.736**	0.0000000000000	87
Commercial	0.652**	0.0000000000044	89
Tourist attraction	0.644**	0.0000000002598	77
Financial and insurance	0.612**	0.0000000003974	86
Catering	0.611**	0.0000000003402	87
Communal facilities	0.572**	0.0000000104320	85
Company and Business	0.505**	0.0000004359302	89
Life Services	0.484**	0.0000015286857	89
Sports and leisure	0.480**	0.0000029637329	86
Governmental agencies	0.467**	0.0000073636040	84
Transport facilities	0.415**	0.0000584600176	88
Scientific,educational,cultural	0.401**	0.0001103573972	88
Health care	0.303**	0.0051347187389	84
Commercial residences	0.177	0.0996348878413	88
Automobile service	0.009	0.9450432391193	66
Motorcycle service	-0.130	0.3257917899378	59

\*\* . At 0.01 level (two-tailed), the correlation was significant

Table 1. Pearson correlation coefficient results.

After using statistical methods to preliminarily verify the correlation, we collect city maps as samples of the training dataset. The origin data is from AutoNavi Maps. We customize the color of green space, water body, and road in ArcGIS. Due to the lack of functional information in the building graphics, we connect the functional information of POI to the building graphics through the join tool to create building area of interest (AOI) graphics. Meanwhile, we sort the grey value of the building features graphic based on the correlation coefficient result. These labels serve as a supervisory signal to guide the network to process, extract, and transform visual information to achieve the maximum performance of the network on the required tasks (Figure 3).

The degree of population concentration may affect the intensity of romantic messages, and the romance heatmap would show a more scattered state of suburban areas. Thus, we further delete the outside area of the inner ring of the cities.

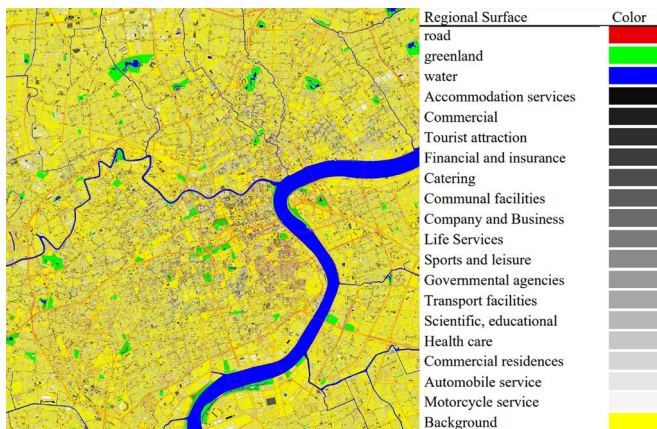


Figure 3. City labelled map.

After aligning the two images and removing the extra part, we divide the two large

images. So far, we have got 795 pairs of samples in the training dataset of city labelled maps and romance heatmaps that correspond to each other.

## 2.4. NEURAL NETWORKS TRAINING

Based on the data type of the experiment and the characteristics of the generative adversarial networks, high-resolution image-to-image translation with conditional adversarial networks (Wang et al., 2018) is used as the main operating algorithm. The training process was done on a computer with a GeForce RTX 2080 graphics card.

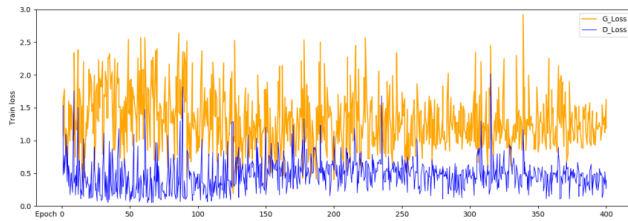


Figure 4. Generator loss ( $G\_LOSS$ ) and Discriminator loss ( $D\_LOSS$ ) during training

The loss values of the generator and discriminator are recorded during the training process. Figure 4 compares the loss values of the generator and discriminator. Training is a process in which the generator and the discriminator ‘compete’ with each other. Relatively, a higher discriminator loss value and a lower generator loss value represent a successful training process. We use a constant learning rate training gradient for the first 200 epochs and a decaying learning rate gradient for the next 200 epochs, which facilitates a more accurate fitting of the data to achieve better convergence. After 370 epochs, discriminator loss and generator loss tend to be stable, which means the training successes gradually. Meanwhile, according to the saved temporary results from the monitoring website (Figure 5), it can be found that at the initial epochs, the synthesized images are more blurred, while as the training proceeds, the synthesized images have stabilized after 370 epochs and can effectively respond to the city map, which means the accuracy of prediction increases. Hence, we decide to use the model trained after 400 epochs as the final prediction model.

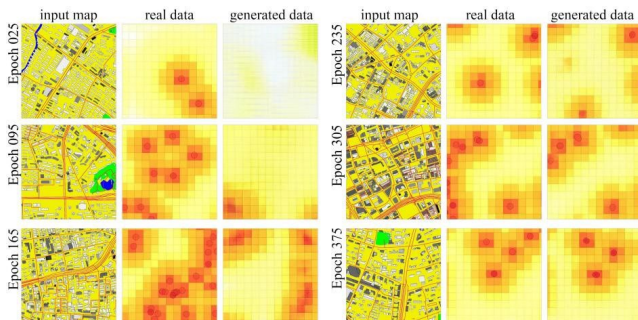


Figure 5. Predicted results for each training epochs.

### 3. Results

#### 3.1. TRAINING ACCURACY

To ensure that our prediction model can be accurate in practical applications, we decide to calculate the similarity of the trained maps in epoch 400 and the counterpart ground truth maps through the average hash algorithm.

We use the average hash algorithm to compare the difference between generated images and real heatmap images, and then calculate their hamming distance (Figure 6).

When the hamming distance is less than 5, the generated image is regarded to be particularly similar to the ground truth map; when the hamming distance is greater than or equal to 5 and less than or equal to 10, the generated image is considered to be quite similar to the ground truth map; and when the hamming distance is greater than 10, the two maps are considered to be unrelated. Thus, we divide the generated images into three categories: the particular like images, the very like images and the different images, and counted the number of these three categories of the generated images (Figure 6). z

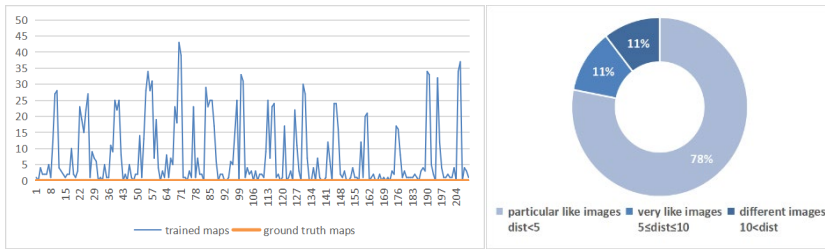


Figure 6. Hamming distances between trained maps and ground truth maps and Percentage of trained images with different levels of accuracy.

We find that 164 of the images are particular like images, 24 of them are very like images, and the remaining 22 images are different images. It means that the training result in epoch 400 has an accuracy rate of 78%, which shows that the result is reliable.

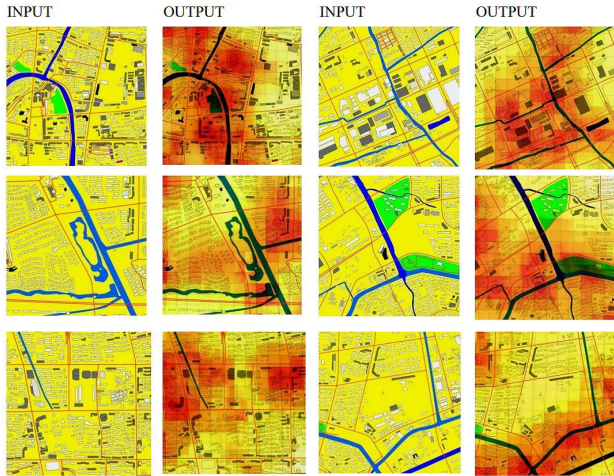
#### 3.2. ROMANCE HEATMAPS PREDICTION

We make city labelled maps of Shanghai's new towns, Nanjing, Suzhou as the input to the neural network and provide suggestions of the possible romance occurring place on the output romance heatmap (Figure 7). Through the direct observation and comparison, we find that in most cases, the density of distribution of predicted romantic places is related to three urban morphological features: green space, water bodies, and building features (with high Pearson correlation coefficient), and their influence on the density of romantic places is: water bodies > green space > building functions.

The result also shows a superimposed effect of these three elements. The high-mixed areas with water bodies, green space, and buildings with commercial, hotel, or catering functions attract more urban romance than individual elements do. This also illustrates the positive impact of high-mixed areas on urban vitality.



In addition, the commercial area along the street is relatively more likely to have romance heat in areas with similar grey values, which shows that romance contains a certain degree of commercial value. Block commercials may be more effective than box commercials.



*Figure 7. Generated results for predicting romance heatmap.*

However, we find that some areas where there is a high distribution of buildings with a high Pearson correlation show a low heat of romance. We speculate that the relationship between urban romance and urban morphology is still complicated and disturbed by many factors.

#### **4. Conclusion**

With the rapid development of urbanization, the relationship between urban design and citizen life has received more and more attention. Romantic places gradually become the infrastructure of cities. This research is based on the geographic location of romantic information on the common social media platform, and the text content is filtered by keywords related to romance, and the complex relationship between romantic places and urban morphology is predicted. We take the city map of the Yangtze River Delta region of China as the input of the GAN model, and present the output data in the form of a romance heatmap. The trained GAN model can also predict where romance will occur in other cities.

This research proves that deep learning can complete the analysis and prediction of urban romantic places, and gradually replace the old empirical methods. The next research project will improve the input data, add semantic information and train more accurate machine learning models. The successful prediction of love maps shows that embracing deep learning can draw knowledge in a wider range of urban perception and cognition scenarios, which empowers the construction of sustainable city and communities in the age of AI with data support.

## References

- Adams, D., & Tiesdell, S. (2012). *Shaping places: urban planning, design and development*. Routledge.
- Buchanan, P. (1988). What city? A plea for place in the public realm. *The architectural review*, 184(1101), 31-41.
- Croitoru, A., Crooks, A., Radzikowski, J., & Stefanidis, A. (2013). Geosocial gauge: a system prototype for knowledge discovery from social media. *International Journal of Geographical Information Science*, 27(12), 2483-2508.
- ESRI UK. (2019). *This gorgeous, interactive map shows Britain's 1,453 'most romantic streets'*, from <https://geoawesomeness.com/valentine-map-britain-street-name-uk/>
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2014). Generative adversarial nets, *Advances in neural information processing systems*, 2014, 2672–2680.
- He, J., & Zheng, H. (2021). Prediction of crime rate in urban neighborhoods based on machine learning. *Engineering Applications of Artificial Intelligence*, 106, 104460.
- Loring, M. L. (2012). Capturing the buzz: social media as a design informant for urban civic spaces.
- Norberg-Schulz, C. (1980). *Genius Loci: Towards a Phenomenology of Architecture*. New York: RIZZOLI INTERNATIONAL PUBLICATIONS, INC.
- Rahimi, S., Andris, C., & Liu, X. (2017). Using yelp to find romance in the city: A case of restaurants in four cities. In *Proceedings of the 3rd ACM SIGSPATIAL Workshop on Smart Cities and Urban Analytics* (pp. 1-8).
- Venkatesan, R., Koon, S., Jakubowski, M. H., Moulin, P., Inst, B. and Ave, N. M. . (2000) "Robust Image Hashing," In *IEEN International conference on image processing: ICIP 2000*, (pp. 1-3).
- Tissink, F. E. (2016). Narrative-driven design.
- VISIT LONDON. (2021). *Guide filled with Valentine's Day ideas.*, from <https://www.visitlondon.com/things-to-do/whats-on/valentines-day>.
- Wang, T.-C., Liu, M.-Y., Zhu, J.-Y., Tao, A., Kautz, J., Catanzaro, B. (2018). High-resolution image synthesis and semantic manipulation with conditional gans. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 8798–8807).
- Zukin, S. (1998). Urban lifestyles: diversity and standardisation in spaces of consumption. *Urban studies*, 35(5-6), 825-839.

# A MACHINE-LEARNING APPROACH TO URBAN DESIGN INTERVENTIONS IN NON-PLANNED SETTLEMENTS

ANNA BOIM<sup>1</sup>, JONATHAN DORTHEIMER<sup>2</sup> and AARON SPRECHER<sup>3</sup>

<sup>1,2,3</sup> MTRL Laboratory, Faculty of Architecture and Town planning,  
Technion- Israel Institute of Technology

<sup>1</sup> [anna.boim85@gmail.com](mailto:anna.boim85@gmail.com), 0000-0003-2272-3958

<sup>2</sup> [jonathan@dortheimer.com](mailto:jonathan@dortheimer.com), 0000-0002-7464-8526

<sup>3</sup> [asprecher@technion.ac.il](mailto:asprecher@technion.ac.il), 0000-0002-2621-7350

**Abstract.** This study presents generative adversarial networks (GANs), a machine-learning technique that can be used as an urban design tool capable of learning and reproducing complex patterns that express the unique spatial qualities of non-planned settlements. We report preliminary experimental results of training and testing GAN models on different datasets of urban patterns. The results reveal that machine learning models can generate development alternatives with high morphological resemblance to the original urban fabric based on the suggested training process. This study contributes a methodological framework that has the potential to generate development alternatives sensitive to the local practices, thereby promoting preservation of traditional knowledge and cultural sustainability.

**Keywords.** Non-planned Settlements; Cultural Sustainability; Machine Learning; Generative Adversarial Networks; SDG 11.

## 1. Introduction

Informal and non-planned settlements are highly complex and typically develop through small-scale, bottom-up actions. Accordingly, their morphology is a result of environmental factors and reflects their specific social structure, economy, traditions, and cultural values (Habraken, 1998). Often regarded as chaotic and non-functional, these settlements present a unique challenge to urban design practitioners and policymakers (Schaur, 1991). Adequate solutions to the substandard living conditions in non-planned settlements require a deep understanding of their initial formation and expansion processes and appropriate tools to evaluate potential policies or interventions (Patel et al., 2018).

To address this challenge, in this study, we present a machine-learning (ML) and computer vision approach that can be used as an urban design tool. We demonstrate a ML model trained to learn and reproduce the complex urban patterns of the town Jisr az-Zarqa in Israel, expressing the unique spatial qualities of this settlement. The lack of consistent planning policy in rural Arab settlements in Israel (Brawer, 1994) led to

unplanned development of Jisr az-Zarqa which, until 1988, had no official master plan (Israel Land Authority, 1992).

We argue that applying Artificial Intelligence (AI) to the design practice can be valuable due to its ability to learn urban patterns that resulted from dynamic and spontaneous self-organization processes. Therefore, AI tools can be meaningfully used by professionals, policymakers, and local communities to better visualize and reflect on the potential outcomes of different development scenarios. Moreover, AI tools can generate design solutions that are not merely imitating local forms but are also sensitive to the local culture and contributing to the promotion of more sustainable communities.

In this study, we contribute a generative machine-learning method to produce complex urban morphology and generate multiple alternatives as design recommendations, based on Pix2PixHD model (Wang et al., 2018).

The remainder of this paper is structured as follows. In Section 2, we briefly review relevant previous studies that used different approaches to investigate informal settlements. The methodology used in the present study is described in Section 3. In Sections 4-6, we present the process and the results of training our model on different data, testing it, and then using it to generate new alternatives for the urban fabric. The results are summarized and discussed in Section 7.

## 2. Related Work

Relevant approaches used in research on informal settlements include agent-based modelling (ABM), shape grammar, and, more recently, machine learning.

Agent-Based Modelling (ABM) is a method used to simulate bottom-up processes to predict urban development. In these simulations, decision-makers are represented as agents capable of responding to their environment and taking autonomous action. For instance, Patel et al. (2018) suggested a model that integrates agent-based modelling with Geographical Information System (GIS) to provide a platform for studying the emergence of slums. The authors argued that this model is helpful for testing slum policies before applying them in real-life settings. Likewise, Patt (2018) explored the applications of the ABM approach to urban redevelopment in informal contexts, with a particular focus on the public space network as the primary object of the multi-agent model.

The shape grammar approach defines initial shapes and a set of transformation rules applied to describe architectural evolution over time. For instance, using the shape grammar approach, Verniz and Duarte (2017) described the evolution of an informal settlement of Santa Marta in Rio de Janeiro and predicted future growth of this informal urban fabric. Similarly, Ena (2018) explored the potential use of shape grammar to regularize the favelas in Rio de Janeiro and better understand spatial ideas and politics behind their architecture.

However, a limitation of both ABM and shape grammar is that these two approaches depend on the assumptions resulting from an analysis of settlements or comprehension of individuals' dynamics and decision-making processes. Yet, the complexity of non-planned settlements makes this a challenging and time-consuming task.

In this context, considering that artificial neural networks can be trained to

recognize highly complex patterns, thus allowing design professionals to focus on the evaluation and decision-making phase, the machine learning (ML) approach becomes particularly relevant. Several previous studies proposed using a specific ML method called Generative Adversarial Networks (GANs) to design urban blocks. More specifically, Yao et al. (2021) used a GAN model pix2pix to generate plot layouts of varying building densities based on similar settings in Shanghai and further evaluated them using Octopus and Ladybug simulations. Furthermore, Fedorova (2021) suggested using GANs to design urban blocks of several European cities based on the surrounding morphology. These methods were applied to planned urban fabric dictated by local policies, masterplans, and regulations. In the next section, we explain how we applied a similar approach to generate a complex urban morphology of a non-planned settlement.

### **3. Methodology**

In this section, we describe the ML model used in the present study, the methods we used to generate training datasets, testing the trained model, using it to generate new urban morphology and evaluating the results.

#### **3.1. MACHINE LEARNING**

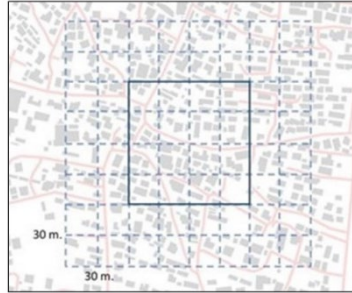
Artificial Neural Networks (ANNs) are computer software modelled after the metaphor of neurons in brains interconnected by synapses. Each neuron receives several inputs and computes an output. The connection between two neurons, called an edge, has a particular weight that influences the specific information transferred. ANNs made of several layers of aggregated neurons are called deep neural networks. When an ANN receives an input, it is processed by the neurons and the ANN generates an output. The main feature of ANNs is their capability to be ‘trained’ to adapt their output. For a more accurate output, the neurons’ computational functions and edges’ weights are modified during the training process.

In the present study, we used Pix2PixHD, a deep learning method that can synthesize photorealistic images from semantic label maps using conditional Generative Adversarial Network (cGAN) (Wang et al., 2018). A GAN is a system consisting of two deep networks—the generator, which generates an output, and the discriminator, which identifies whether or not the generated output is similar to the requested output. GANs are useful since they can be trained based on a dataset in an unsupervised way. In contrast to a regular GAN, a cGAN is a specific kind of GAN that receives an input. Several previous studies used similar methods to explore the utility of the ML method for urban and architectural plans generation (Fedorova, 2021; Ye et al., 2021; Chaillou, 2020).

#### **3.2. TRAINING DATA SET**

One of the challenges associated with creating an ML model is the large amount of data required for unsupervised training. Specifically, producing a large quantity of map images for each dataset can be highly time-consuming. Therefore, an efficient method to generate large amounts of data will increase accessibility of the suggested model to practitioners.

In the present study, we selected ArcGIS for data set extraction since it has a relatively accurate mapping of Jisr az-Zarqa. Furthermore, we leveraged ArcGIS's integrated Python scripting option and created a "map extraction tool" (MTRL



*Figure 1. Jisr Az-Zarqa – map of the town including its building masses (black) and roads (red). The map is divided into a grid of 30m offset. (Credit: MTRL 2021, Technion IIT)*

Technion IIT, 2021) that automatically exported maps in JPEG format (see Figure 2a, 2b). Using the Python script, we generated a total of 220 map tiles that were sufficient for training the GAN. The initial extent and scale of the first tile were defined, and the subsequent tiles were automatically exported with an offset of 30 meters on the X and Y coordinates until the entire area of the town was covered (see Figure 1).

### 3.3. EVALUATION

Upon completion of the training process, we tested the model on map images that were not included in the training set. To evaluate the success of the trained model, the test results were compared visually to ground-truth maps of the settlement in terms of urban structure and patterns. The complexity of non-planned urban fabric can be addressed through multiple aspects; however, the current study focuses solely on morphology. Since the data for the presented experiments is highly simplified, representing limited morphological information, the selected evaluation criteria were also limited to four basic characteristics- the model's ability to a) continue existing roads and suggest new ones, b) recognize topological features of the street network, c) reproduce building morphology with similar density and d) reproduce similar building typology. Topological feature of the street network refers to the linking of different elements, thus the number of streets that meet at each intersection. Density refers to the number of buildings per given area and building typology refers to the shape and area of the building masses.

### 3.4. SOFTWARE AND HARDWARE

The model was trained on a computer with a 12 core, 3.6GHz CPU with 32GB of memory and an Nvidia Quadro RTX 5000 graphic card. The computer's operating system was Ubuntu 20, and we installed the Python scripts using Anaconda. Pix2pixHD was retrieved from Wang et al. (2018).

In Sections 4-6, we report the results of three studies (Studies 1-3) where we trained

the Pix2PixHD model on different data, tested it, and used the results to generate new alternatives for the urban fabric of Jisr Az-Zarqa.

#### 4. Study 1- Pilot

In Study 1, we aimed to train the model to complete a partial map image with patterns that would resemble the existing morphology of the Jisr Az-Zarqa settlement.

##### 4.1. METHOD

A graphic layer of the Jisr az-Zarqa map was created based on ArcGIS's base map. The layer contained black outlined polygons as buildings and red polylines as roads and was exported as JPEG tile pairs using the map extraction tool. Each pair was made from a partially blank input map and a complete output map. To produce the input map, the output map was duplicated, and the right half was erased (see Figure 2b). Next, the model was trained for 200 epochs; the total duration of the training was ca. 36 hours. Finally, we tested the trained model using new input tiles and evaluated the generated outcome.

##### 4.2. RESULTS

Figure 2c shows the results of the trained model after 200 epochs given the input in Figure 2b. As can be seen in the figure, the model failed to identify the buildings polygons as patterns in the current settings or to complete the missing urban morphology. The generated forms were distorted polygons, and the generated pattern did not fill the entire frame. These results highlight that it is not straight-forward to train a ML model to generate such patterns. Accordingly, we concluded that a different image labelling, with adjusted structure and colors, would be necessary.

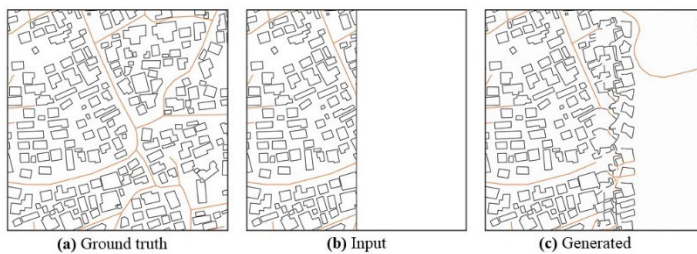


Figure 2. Study 1 - Test ground truth map tile of Jisr az-Zarqa (a), test input image (b), and generated output (c). In the process of training the model, tile (b) was used as input and (a) as output. (Credit: MTRL 2021, Technion IIT)

#### 5. Study 2- Nolli Maps

Based on the results of Study 1, in Study 2, we repeated the training process using image labelling by creating a different dataset format. The new format provided a 360-degree context for the blank area and used a simpler image labelling.

## 5.1. METHOD

We generated a new simplified dataset using the map extraction tool. Compared to the approach used in Study 1, in this study, there were two main differences. First, we changed map labelling by replacing the building polygons with solid black polygons (like Nolli maps). We hypothesized that this adjustment would mitigate the corrupt polygons. At this stage, the red polylines of roads were removed. The second difference was changing the input map format—specifically, we erased a rectangle in the center of the image (instead of the right half as in Study 1). We hypothesized that providing a surrounding urban context would be beneficial for the completion of patterns over the required area. The size of the new dataset remained the same (220 tiles), and the model was trained anew, this time for a total of 300 epochs, which took approximately 48 hours to complete.

## 5.2. RESULTS

The results of Study 2 showed a significant improvement. Specifically, the model successfully learnt the existing urban fabric and completed the blank part of the map accordingly. The polygons were complete, and the model covered the whole area (see Figure 3c). The generated map produced a new pattern with a high morphological resemblance to its surroundings and to the ground truth. The newly filled area contained 40 buildings of mainly rectangular shape with the total built area of 5702 sqm, while the ground truth map contained 42 buildings with the total built area of 5865 sqm. In addition, the model was able to fill in the continuous roads which were not explicitly marked but were visible due to the broader spaces between the buildings.



Figure 3. Study 2—Nolli maps dataset. Ground truth (a), the corresponding input map (b), and the map generated by the trained model (c). (Credit: MTRL 2021, Technion IIT)

## 6. Study 3- Generating New Urban Fabric

The aim of Study 3 was three-fold: 1) to increase the complexity and size of the dataset; 2) use the trained model to generate multiple alternatives to a new urban fabric as an infill in an empty area of Jisr az-Zarqa; and 3) test the model on different settlements.

### 6.1. METHOD



The training process was repeated, this time on a higher-complexity dataset. The Nolli maps from Study 2 were used, the size and the scale remained the same. However, the roads layer was added again as red polylines. To avoid overfit, the dataset was enlarged from 220 tiles to 1100 tiles by rotation and mirroring of the original dataset.

Thus far, the model was found to be able to generate one single output for each input, even when tested on the same map multiple times. However, for ML models to become a potential design tool, they should be able to produce a variety of alternatives. To address this challenge, we tested the model on several input maps with a small offset (1 m) between each. This time, the blank area of the map included an empty area potentially suitable for infill intervention.

The same model was tested on two different examples—namely, Fureidis, an Arab town near Jisr az-Zarqa, and New York. Despite being located on much steeper topography, Fureidis shares some cultural and morphological similarities to Jisr az-Zarqa; accordingly, this test was run to identify whether the trained model can be applicable to other similar settlements. New York—which is a planned, orthogonal grid—was tested as an opposite example.

## 6.2. RESULTS

The results revealed that, after 500 epochs, the model successfully generated an urban fabric that morphologically resembled the existing patterns of Jisr az-Zarqa (see Figure 4c). The continuous roads were consistent with the ground truth, and the intersections were of the same topology—namely, three-armed intersections and dead-end streets. The building density was similar- 47 buildings vs. 44 in the ground truth, however the scale of the building masses was smaller- with total built area of 5997 sqm vs. 8748 sqm in the ground truth. This finding can be explained by the fact that this particular area is the commercial and the historical center of the town; therefore, it is the first to undergo a densification process. Since the model completed the blank part with patterns learned from the context, it reproduced the same building typology as the surroundings which to date are not considerably densified yet.

However, as shown in Figure 4f, the results generated for Fureidis were not as accurate. Specifically, the building density was lower- only 61 buildings with a total built area of 6012 sqm vs. 78 buildings with total built area of 9428 sqm in the ground truth. The generated roads were less consistent with the ground truth in terms of continuity, likely due to the different structure of Fureidis where topography plays a more significant role, a dimension which is currently not addressed by the model. Notwithstanding, the topology of the road network was partially preserved. Moreover, the model adapted the existing network to self-generated morphology- replacing a planned round-about with a three-armed junction. This trend becomes even more evident in the test results for New York (see Figure 4i): since the model was not yet familiar with continuous, orthogonal grids, the road network was transformed to become dead-end streets and 3-armed junctions. The missing building masses were completed according to the Jisr az-Zarqa morphology, and the model attempted to not just fill in the blank part of the map, but also to modify the surrounding context.

As the testing results of study 3 for Jisr az-Zarqa revealed that the model can reproduce urban fabric with similar morphological features, Figure 5 demonstrates several alternatives of new infill within an existing urban. The different outputs generated by the model had only a slight offset in the input maps. Such small offset enables the user to remain in the same surrounding context yet creates different settings for the model with a high degree of variation between the outputs in terms of building layout and additional or modified roads.

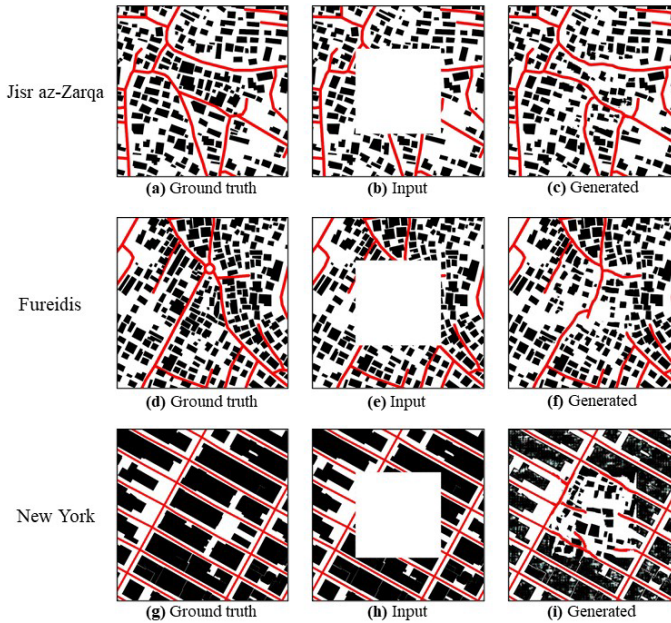


Figure 4. Study 3—Test results for the trained model. Upper row - Jisr az-Zarqa, middle row- Fureidis, bottom row- New York. Ground truth (left), input (center) and model output (right). (Credit: MTRL 2021, Technion IIT)

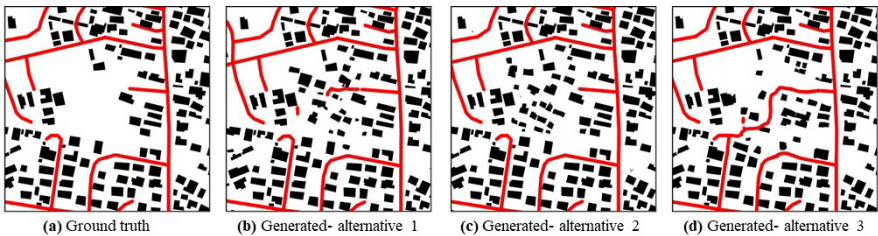


Figure 5. Study 3- An example of generating multiple alternatives (b-c-d) for a new infill fabric in Jisr az-Zarqa. (Credit: MTRL 2021, Technion IIT)

## 7. Discussion

### 7.1. MACHINE LEARNING FOR COMPLEX URBAN MORPHOLOGY GENERATION

The results of the present study revealed that machine-learning models hold great potential in the research and practice in informal and non-planned contexts. Unlike other popular approaches that require complex and time-consuming analysis, ML models are capable of learning local patterns and generating new outputs, as well as making adjustments in the existing fabric.

GAN models, specifically Pix2PixHD, are an effective method for urban design. An efficiently trained GAN model can reproduce a complex urban fabric of a non-planned settlement, with morphological resemblance of buildings and consistent topological qualities of the road network. The results of Studies 1-3 emphasize that the success of the ML approach largely depends on the quality of the training data. Specifically, we determined that good performance depends on distinct labelling, appropriate scale, and an appropriate balance between dataset size and training epochs in order to avoid overfit. In Studies 1-3, the best results were achieved at the scale of 1:2000, with dataset size of at least 1100 maps, and with 500 training epochs. In addition, we demonstrated that the trained model can be used in other settlements with similar morphology.

The ability to provide long-term sustainable solutions largely depends on the evaluation of the suggested policies and interventions. In this study, we presented a method to generate multiple output variations that allows professionals and the community to reflect on the potential outcomes of different scenarios as part of the design process.

### 7.2. LIMITATIONS AND FUTURE RESEARCH

While the presented method achieved the expected performance, it has several significant limitations in the broader scope of urban design. Since the model is two-dimensional, it cannot address topography, building heights, or three-dimensional visualization of the proposed layouts. Furthermore, as revealed by the results, it currently fails to generate new urban morphology without full context, such as a new extension to an existing settlement in an empty area. Furthermore, the proposed method cannot be adjusted to address design requirements, such as increased density, or minimal and maximal building sizes.

Accordingly, in order to improve the model's performance in practical applications, further research and development would be needed. Notwithstanding, the results of iterative series of experiments reported in this study contribute to establishing a methodological framework that can be used as a generative tool in informal contexts. Our contribution includes the production of datasets for training ML models, testing, and generating new urban layouts that can be further evaluated as urban design interventions.

To conclude, the proposed method can bridge the gap between top-down and bottom-up design practices and generate development alternatives that are sufficiently sensitive to the local communities, thus promoting preservation of traditional

knowledge and cultural sustainability.

## References

- Brawer, M. (1994). The internal structure of the traditional Arab village. In Grossman, D. and Meir, A. (Eds.), *The Arabs in Israel: Geographical Dynamics* (pp.99–113). Bar-Ilan University Press, Israel.
- Chaillou, S. (2020). ArchiGAN: Artificial Intelligence x Architecture. In: Yuan P., Xie M., Leach N., Yao J., Wang X. (Eds.) *Architectural Intelligence* (pp. 117-127). Springer, Singapore. [https://doi.org/10.1007/978-981-15-6568-7\\_8](https://doi.org/10.1007/978-981-15-6568-7_8)
- Ena, V. (2018). De-coding Rio de Janeiro's Favelas: Shape grammar application as a contribution to the debate over the regularisation of favelas. The case of Parque Royal. In *Proceedings of the 36th eCAADe Conference* (2) (pp. 429–438).
- Fedorova, S. (2021). Generative adversarial networks for urban block design. In *SimAUD 2021: A Symposium on Simulation for Architecture and Urban Design*, 2021.
- Habraken, N. J. (1998). *The structure of the ordinary: Form and control in the built environment*. The MIT Press, Cambridge, Massachusetts.
- Israel Land Authority (1992). *Jisr az-Zarqa Masterplan no. 356(870)*. Retrieved December 3, 2020, from: <https://apps.land.gov.il/TabaSearch/#/Plans/Plan/3004044>
- MTRL Technion IIT (2021). Map export script. [https://github.com/MTRL-lab/machine\\_learning\\_morphology](https://github.com/MTRL-lab/machine_learning_morphology)
- Patel A., Crooks A., Koizumi N. (2018) Spatial Agent-based Modeling to Explore Slum Formation Dynamics in Ahmedabad, India. In: Thill JC., Dragicevic S. (Eds.) *GeoComputational Analysis and Modeling of Regional Systems*. Advances in Geographic Information Science. (pp. 121–141). Springer, Cham. [https://doi.org/10.1007/978-3-319-59511-5\\_8](https://doi.org/10.1007/978-3-319-59511-5_8)
- Patt, T. R. (2018). Multiagent approach to temporal and punctual urban redevelopment in dynamic, informal contexts. *International Journal of Architectural Computing*. 16(3) (pp.199-211). <https://doi.org/10.1177/1478077118793127>
- Schaur, E. (1991). *IL39- Non-planned settlements*. Institute for Lightweight Structures, University of Stuttgart.
- Verniz, D., Duarte J.P. (2017). Santa Marta Urban Grammar: Towards an understanding of the genesis of form. In *Proceedings of the 35th eCAADe Conference* (2) (pp.477–484).
- Wang, T.-C., Liu, M.-Y., Zhu, J.-Y., Tao, A., Kautz, J., and Catanzaro, B. (2018). High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs. In *2018 Conference on Computer Vision and Pattern Recognition*, (pp. 8798–8807). IEEE.
- Yao, J., Huang, C., Peng, X. I. & Yuan, P. F. (2021). Generative design method of building group: Based on generative adversarial network and genetic algorithm. In *Proceedings of the 26th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA)* (1) (pp.61-70).
- Ye, X., Du, J., & Ye, Y. (2021). MasterplanGAN: Facilitating the smart rendering of urban master plans via generative adversarial networks. *Environment and Planning B: Urban Analytics and City Science*, 1-21. <https://doi.org/10.1177/23998083211023516>

# ENERGY-DRIVEN INTELLIGENT GENERATIVE URBAN DESIGN

*Based on deep reinforcement learning method with a nested Deep Q-R Network*

CHENYU HUANG<sup>1</sup>, GENGJIA ZHANG<sup>2</sup>, MINGGANG YIN<sup>3</sup> and  
JIAWEI YAO<sup>4</sup>

<sup>1</sup>*School of Architecture and Art, North China University of Technology*

<sup>2</sup>*Department of Architecture, Tamkang University*

<sup>3</sup>*East China Architectural Design and Research Institute Co., Ltd*

<sup>4</sup>*College of Architecture and Urban Planning, Tongji University*

<sup>1</sup> *huangchenyu303@163.com, 0000-0002-6360-638X*

<sup>2,3</sup> *{r2328643|minggangyin}@gmail.com* <sup>4</sup> *jiawei.yao@tongji.edu.cn, 0000-0001-7321-3128*

**Abstract.** To attain "carbon neutrality," lowering urban energy use and increasing the use of renewable resources have become critical concerns for urban planning and architectural design. Traditional energy consumption evaluation tools have a high operational threshold, requiring specific parameter settings and cross-disciplinary knowledge of building physics. As a result, it is difficult for architects to manage energy issues through 'trial and error' in the design process. The purpose of this study is to develop an automated workflow capable of providing urban configurations that minimizing the energy use while maximizing rooftop photovoltaic power potential. Based on shape grammar, parametric meta models of three different urban forms were developed and batch simulated for its energy performance. Deep reinforcement learning (DRL) is introduced to find the optimal solution of the urban geometry. A neural network was created to fit a real-time mapping of urban form indicators to energy performance and was utilized to predict reward for the DRL process, namely a Deep R-Network, while nested within a Deep Q-Network. The workflow proposed in this paper promotes efficiency in optimizing the energy performance of solutions in the early stages of design, as well as facilitating a collaborative design process with human-machine interaction.

**Keywords.** Energy-driven Urban Design; Intelligent Generative Design; Rooftop Photovoltaic Power; Deep Reinforcement Learning; SDG 11; SDG 12.

## 1. Introduction

Energy consumption in the building sector already accounts for more than 38% of greenhouse gas emissions and has risen to prominence as a significant contributor to air pollution and global warming. With various countries implementing carbon

neutrality targets, reducing urban energy consumption, increasing the use of renewable resources, and enhancing the sustainability of energy systems have become critical issues for urban planning and building design.

Accurate urban energy consumption prediction models are crucial for designers to predict the impact of planning and building design as well as decision makers to develop energy policies and pricing (Ahmad T et al., 2021). Based on the technical framework of "Performance-driven Generative Design", this study proposes a data-driven urban energy modelling approach and an energy-targeted generative design workflow. By introducing generative design and deep reinforcement learning, parametric meta models will undergo efficient design iteration in a controlled manner, ultimately converging to the optimal solution for energy performance.

## 2. Related works

### 2.1. ENERGY OPTIMIZATION IN THE EARLY STAGE OF DESIGN

Building Energy Simulation (BES) has developed mature tools and applications over the last three decades. BES are a complex task that require expert knowledge in a variety of disciplines, including building geometry, material thermal properties, and HVAC. Even the simplest simulations require a large amount of input data and a significant amount of time (Hong T et al., 2019). Due to these characteristics, BES is difficult to intervene during the early stage of design, when information is scarce. Pil Brix Purup reviewed the research framework for developing early BSE tools, concluding that current software development and information integration remain significant barriers (Purup, P. B. et al., 2020).

### 2.2. URBAN ENERGY PERFORMANCE MODELING

The majority of existing research has concentrated on evaluating and optimizing the energy consumption of single buildings (Lu S et al., 2021). However, the principles of building energy simulation differ at various scales, and the scale effect cannot be ignored as the scale of energy simulation increases. There is still a dearth of effective tools for simulating energy consumption at urban scales.

Abbasabadi's review discusses contemporary approaches to energy modeling, including data-driven approaches and physical-based simulations (Abbasabadi, N. et al., 2019). BES at the urban scale using a physical-based approach is a difficult task, and because no form has been developed during the early stages of urban planning, urban BES programs and optimization engines must be linked to parametric models. Additionally, the conventional approach is incapable of dealing with nonlinear coupling relationships that are complex (Shi, Z. et al., 2017). There is a need to develop a data-driven approach to district energy modeling in urban areas that is based on multi-factor coupling.

### 2.3. PERFORMANCE-DRIVEN BUILDING DESIGN

To improve the interaction between design and performance evaluation, it is common to combine generative design, simulation programs, and multi-objective optimization

to create performance-driven generative designs (Nguyen A T et al., 2014). As an important component of building performance, building energy optimization are combined with generative design to form energy-driven generative design. The results of the BES are established as the objective function for optimization. Energy-driven urban design will shape the urban form to accommodate energy infrastructure or technologies.

Numerous relevant investigations have been conducted utilizing multi-objective optimization techniques based on genetic algorithms. However, if the generative design has an excessive number of variable parameters, the search for optimal solutions can result in dimensional explosion. Increased dimensionality of the solution space in the urban context reduces the efficiency of the search and makes convergence more difficult. Cheng illustrates the obstacles faced by urban scale morphology optimization and proposes a reinforcement learning approach combining parametric geometry with BES engine to reveal the influence of urban design parameters (Chang S et al., 2019). In this paper, we used deep reinforcement learning to determine the optimal performance.

### 3. Method

#### 3.1. RESEARCH WORKFLOW

This study is oriented towards the early stages of design, aiming to provide urban configurations that reduce energy demand and simplify and accelerate the BES process. In particular, the influence of the surrounding environment is considered, while providing near-realistic application scenarios in combination with the urban rooftop PV potential power generation prediction.

Based on shape grammar, parametric meta models of three different urban forms were developed and batch simulated for its energy performance. A deep neural network named R-net are used to establish the mapping of key design parameters to energy consumption, which aim is to provide real-time energy performance predictions for the reinforcement learning process without going through the BES engine. Finally, the parametric model is optimized using deep reinforcement learning, and the optimization of the net urban energy consumption is performed using Deep Q network integrated with R-net to provide the high-performance urban model to be deepened. The research workflow is shown in the Fig 1.

#### 3.2. PARAMETRIC META MODELING

Based on Grasshopper shape grammar, the urban generation algorithm developed in this study can adapt different base red lines as input to automatically segment the internal street contours and generate building models. According to the use of the base set to change the type, the urban form is divided into three types: single-point, enclosed and mixed. The secondary morphology control indexes are established, including floor area ratio ( $FAR$ ), building density ( $\lambda_p$ ), average building height ( $H_{ave}$ ), standard deviation of building height ( $H_{std}$ ), average street height to width ratio ( $H/D$ ), shape factor ( $SF$ ) and orientation ( $Orien.$ ), where orientation refers to the direction of the parametric modelling process's land cut axis.

The site is divided into equal areas as much as possible during the automatic generation of the street profile in order to avoid producing deformed land and wasting land. The corresponding volume is controlled by the morphological indicators in order to ensure the urban meta model adapts to the real application scenario. The site's division mechanism is depicted in Fig 2, and the meta model production operations are depicted in Fig 3.

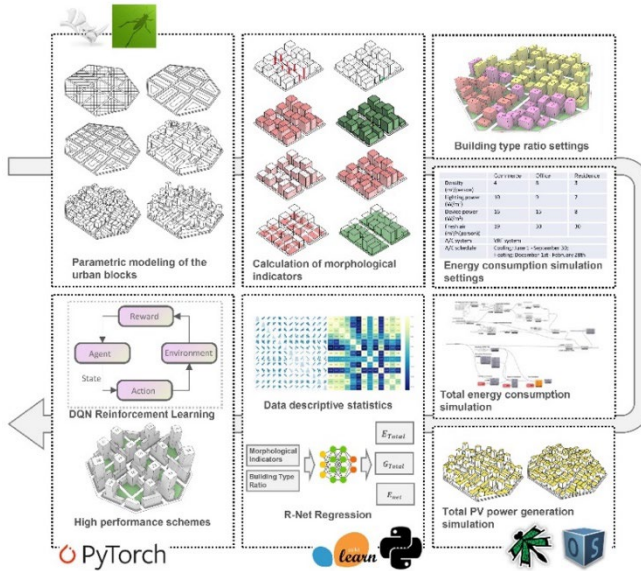


Figure 1. The research workflow

### 3.3. PARAMETRIC META MODELING

The Dragonfly plugin for Rhino & Grasshopper is used to calculate urban energy consumption. It utilizes OpenStudio as the calculation engine to run district-scale BES. For the purpose of predicting energy usage intensity during the early stages of design, default values for parameter settings such as the schedule and air conditioning system were applied. The fundamental parameters for the BES are listed in Table 1.

	Commerce	Office	Residence
Personnel density (m <sup>2</sup> /person)	4	8	3
Lighting power density (W/m <sup>2</sup> )	10	9	7
Device power density (W/m <sup>2</sup> )	16	15	8
Fresh air volume (m <sup>3</sup> /h(-person))	19	30	30
Air conditioning system	VRF system		
Air conditioning system turn-on schedule	Cooling period: June 1 - September 30 Heating period: December 1st - February 28th		

Table 1. The fundamental parameters for the BES

This calculation of energy consumption takes into account the HVAC system,



lighting, and electronic equipment. Simultaneously, the roof area of the building is extracted and the solar photovoltaic energy generation system is configured to operate at the fixed power level. Total energy consumption ( $E_{total}$ ) and total photovoltaic energy generation ( $G_{total}$ ) are included in the computation results. The term "net energy consumption ( $E_{net}$ )" refers to the  $E_{total} - G_{total}$ .

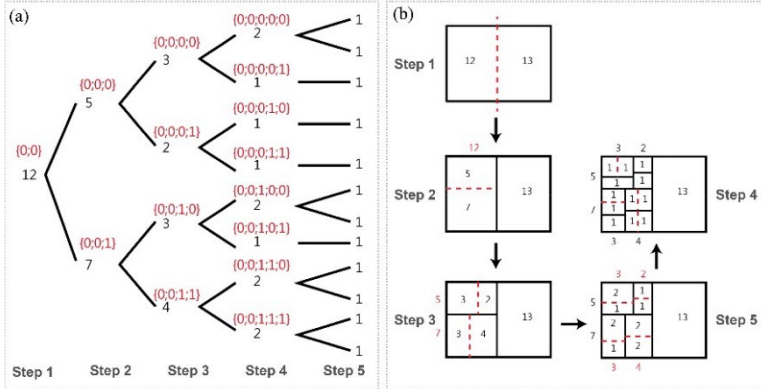


Figure 2. (a) Example of a data tree, (b) Site division based on data tree

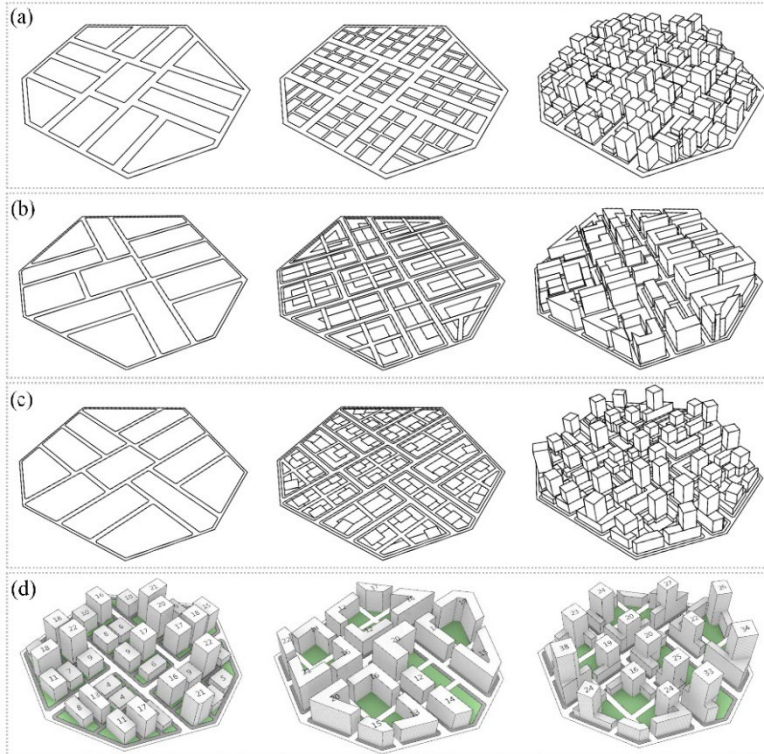


Figure 3. (a) Generative process of single-point type, (b) Generative process of enclosed type, (c) Generative process of mixed type, (d) Three examples of parametric meta models

Dragonfly was created in 3 stages. Step 1, a single building simulation was performed, taking into account the shading of nearby buildings within 50 meters. Step 2, the urban BES was converted into multiple single energy models. Because the functional ratios of the city have a significant impact on energy use patterns, the urban BES in this study was set up with three building types that were randomly assigned in the regional model: office building (*Office*) (15-35%), commercial building (*Comm.*) (15-35%), residential building (*Resid.*) (30-70%). Step 3, Large-scale single building energy consumption was calculated with parallel computing and aggregated as a training dataset for R-net.

### 3.4. INTERACTIVE PERFORMANCE OPTIMIZATION BASED ON DRL

This study uses deep reinforcement learning (DRL) to discover urban configurations with optimal energy performance. DRL is a branch of machine learning that has the potential to realise Artificial General Intelligence (AGI). The process of reinforcement learning is shown in Fig 4 (a). Reinforcement learning divides the world into two parts: the Environment and the Agent. The Agent interacts with the Environment by executing Actions and gets feedback from the Environment. In reinforcement learning, the feedback is represented in the form of Reward. The Agent learns how to interact with the Environment more effectively in the loop in order to maximize Reward.

Through the application of dynamic tactics, reinforcement learning enables the acquisition of extra information and more efficient search in complex environments. In comparison, as traditional optimization method, genetic algorithms take a large number of static strategies and choose the most profitable method and its variants to develop the next generation of strategies. Such a static method is difficult to adapt to the increasing search dimension, particularly in light of the urban generation challenge.

A typical DRL approach is Deep Q Network (Mnih, V. et al., 2015), which uses a neural network to construct a Q-Net mapping from states to values (the sum of all possible future Rewards), and then executes the action with the highest value in the action space at each state, as illustrated in Fig 4 (b).

The first step to optimization is to convert the urban generation process into a deep reinforcement learning process and to configure the environment. The urban parametric meta model is defined as the agent in our task, and the states and actions are defined in Fig 4 (c). Each moment can be represented by a collection of normalized feature parameters  $\mathbf{s}_t$  based on the prior specification of the extraction of urban morphological indicators and the scaling of building types.

$$\mathbf{s}_t \in \mathbb{R}^N, N = 10$$

Where  $N$  denotes the number of feature values, in this case,  $N = 10$ . Normalization is performed to eliminate the magnitude effect. By setting each feature value of the initial State  $\mathbf{s}_0$  to 0.500, the action of the urban meta model at each moment is the combination of each feature value moving one Step distance to the left or right on the number axis since 0.500 which is shown in Fig 4 (c). Step is specified as the hyperparameter of DQN in this study. Given that some feature values remain constant throughout an action, the full action space  $\mathbf{A}_t$  is as follows:

$$\mathbf{A}_t \in \mathbb{R}^M, M = \left( \sum_{i=1}^N C_N^i \right)^2$$

where  $M$  is the number of possible actions of the urban meta model.

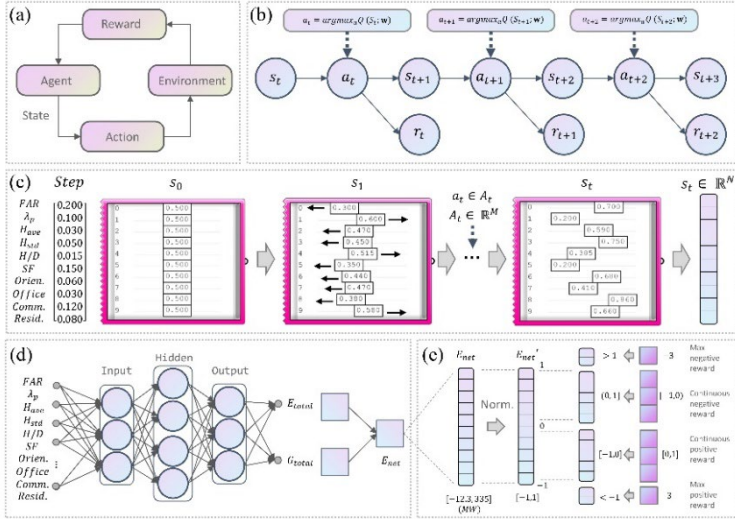


Figure 4. (a) Reinforcement learning principle (b) Markov decision process (c) Environment configuration: State and Action definitions (d) R-Net deep network architecture (e) Environment configuration: Reward definitions

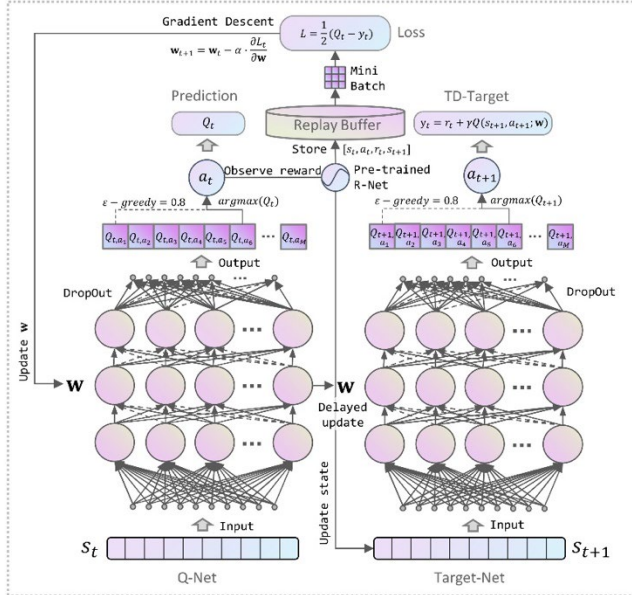


Figure 5. A nested Deep Q-R Network for DRL training

The reward function still needs to be defined in the DRL environment configuration, i.e., to establish a mapping between the urban meta model and its energy performance. We compared several machine learning methods and selected the best

performing neural network to build the regression model, as illustrated in Fig 4 (d). The net energy consumption was calculated from the total energy consumption and the total PV generation for defining reward. The interval [-1,1] was used to map the value domain of net energy consumption. The corresponding reward is given according to the value of the normalized net energy consumption. We encourage the agent to perform morphological exploration beyond the batch simulation results, by offering the maximum positive reward. The reward function is defined in Fig 4 (e). 4 (d) and 4 (e) together form the R-Net, which is used for later DQN training.

DQN is a Temporal Difference (TD) algorithm using an Off-line policy. The Off-line policy refers to the replay buffer, which may store experience and learn in small batches, enabling not only distributed training, but also avoiding the correlation between successive experiences. The TD algorithm is implemented by creating two deep neural networks, Q-Net and Target-Net. They are structurally identical except for the fact that the weights  $w$  of Q-Net are updated in real time while the weights  $w$  of Target-Net are updated in a delayed way to maximize training efficiency. In this study, a pre-trained R-Net is nested within the DQN network to ensure that the Q-Net can instantaneously observe the present moment of  $r_t$  following the morphological change of the urban meta model. The improved DQN is illustrated in Fig 5.

### 4. Results

#### 4.1. DESCRIPTIVE STATISTICS ON DATASET

The batch urban BES engine was set up to randomly sample within the same site red line by connecting with a parametric meta model, and finally 444 total samples were collected as the dataset for descriptive statistics and machine learning. Correlation analysis was performed to establish a link between morphological control indicators, functional configurations, and energy usage. Pearson correlation coefficient results indicated that the outcomes were compatible with current morphological investigations and justified the dataset. The bivariate scatter matrix and correlation analysis of dataset are shown in Fig 6.

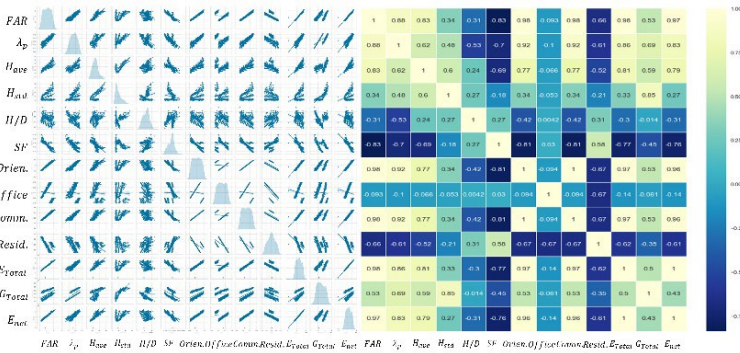


Figure 6. Bivariate scatter matrix and correlation analysis

#### 4.2. COMPARISON OF MACHINE LEARNING MODELS

The purpose of using machine learning is to provide DQN with rapid reward feedback without the use of urban BES. This section compares numerous machine learning models using Scikit-learn, an open-source machine learning tool for Python, in order to avoid the effects of different machine learning algorithms' inductive preferences. Using the hold-out method, 80% of the 444 data were used as the training set and 20% as the test set to train machine learning regression models. Table 2 compares the performance of several machine learning models. The neural network outperformed all other networks in terms of RMSE and R2. Following that, a deep neural network was developed using PyTorch with hyperparameter optimization, and eventually, an R-Net was constructed to predict the reward of the DRL process.

Table 2. Comparison of machine learning models

Machine learning method	Response: $E_{total}$ (MW)		Response: $G_{total}$ (MW)	
	RMSE	R <sup>2</sup>	RMSE	R <sup>2</sup>
Multiple linear regression	11.605	0.96	1.4832	0.92
Regression tree	14.429	0.94	1.5967	0.91
Support vector regression	11.662	0.96	9.4151	0.97
Gaussian process regression	11.535	0.96	1.0358	0.96
Ensemble tree	7.560	0.98	0.7665	0.97
Neural Network	5.982	0.98	0.6572	0.98

### 4.3. OPTIMAL URBAN FORMS

The DQ-RN environment was constructed in this study using the PyTorch deep learning framework, and the training process was accelerated using CUDA. By adjusting  $\epsilon - greedy$ , several exploration and exploitation strategies were evaluated, and  $\epsilon - greedy = 0.8$  was chosen for training. After up to 20 hours of training, the total reward curve exhibited a trend toward convergence. The end states of the extracted single-point type 11000 epoch, enclosed type 18300 epoch, and mixed type 35300 epoch urban meta-models are shown in Fig. 7, and their net energy consumption has been drastically reduced relative to the beginning state. The mixed model's net energy consumption is already less than the dataset's minimum net energy consumption. Meanwhile, the trained models are adaptable to a variety of site boundary inputs, and the morphological optimal solution can be obtained in less than 100 steps.

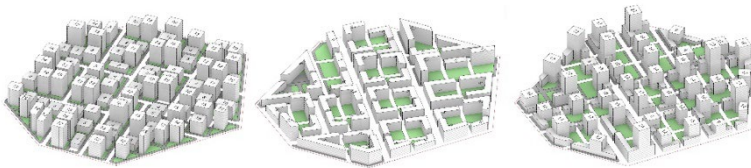


Figure 7. Optimal urban forms of 3 urban types

## 5. Results

The purpose of this study is to develop an automated workflow capable of providing urban configurations that minimizing the energy use while maximizing rooftop photovoltaic power potential. Based on shape grammar, parametric meta models of

three different urban forms were developed and batch simulated for its energy performance. Deep reinforcement learning is introduced to find the optimal solution of the urban geometry. In this study, we propose an improved Deep Q-R Network, by nesting an R-Net within the DQN. This allows for real-time mapping of urban morphological factors to energy performance, guaranteeing that the deep reinforcement learning process is efficient. The trained model can be used for urban optimization search in similar scenarios and will provide several possible solutions in a short period of time, resulting in a net energy consumption reduction.

Current research has limitations, such as DQN can only handle discrete behaviors of agent. In urban generation tasks, the design parameters vary as continuous values, so a hyperparameter Step had to be added in the paper to bridging this gap. However, such an operation also reduces the diversity of urban metamodel generation. In the future, combining adversarial generative networks (GANs) with parametric urban meta models will have the ability to deal with other design requirements like different distributions of volumes and aesthetic considerations.

### Acknowledgements

This research was financially supported by the National Natural Science Foundation of China (grant number 51908410) and Shanghai Sail Program (grant number 19YF1451000).

### References

- Abbasabadi, N., & Ashayeri, M. (2019). Urban energy use modeling methods and tools: A review and an outlook. *Building and Environment*, 161, 106270.
- Ahmad, T., Zhang, D., & Huang, C. (2021). Methodological framework for short-and medium-term energy, solar and wind power forecasting with stochastic-based machine learning approach to monetary and energy policy applications. *Energy*, 231, 120911.
- Chang, S., Saha, N., Castro-Lacouture, D., & Yang, P. P. J. (2019). Multivariate relationships between campus design parameters and energy performance using reinforcement learning and parametric modeling. *Applied energy*, 249, 253-264.
- Hong, T., Kim, J., & Lee, M. (2019). A multi-objective optimization model for determining the building design and occupant behaviors based on energy, economic, and environmental performance. *Energy*, 174, 823-834.
- Lu, S., Wang, C., Fan, Y., & Lin, B. (2021). Robustness of building energy optimization with uncertainties using deterministic and stochastic methods: Analysis of two forms. *Building and Environment*, 205, 108185.
- Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529-533.
- Nguyen, A. T., Reiter, S., & Rigo, P. (2014). A review on simulation-based optimization methods applied to building performance analysis. *Applied Energy*, 113, 1043-1058
- Purup, P. B., & Petersen, S. (2020). Research framework for development of building performance simulation tools for early design stages. *Automation in Construction*, 109, 102966.
- Shi, Z., Fonseca, J. A., & Schlueter, A. (2017). A review of simulation-based urban form generation and optimization for energy-driven urban design. *Building and Environment*, 121, 119-129.

# DATA-DRIVEN EVALUATION OF STREETS TO PLAN FOR BICYCLE FRIENDLY ENVIRONMENTS: A CASE STUDY OF BRISBANE SUBURBS

GABRIELLE TOOHEY<sup>1</sup>, TOMMY BAO-NGHI NGUYEN<sup>2</sup>, RITVA VILPPOLA<sup>3</sup>, WAISHAN QIU<sup>4</sup>, WENJING LI<sup>5</sup> and DAN LUO<sup>6</sup>

<sup>1,2,3,6</sup>*School of Architecture, The University of Queensland.*

<sup>4</sup>*Department of City and Regional Planning, Cornell University.*

<sup>5</sup>*Center for Spatial Information Science, The University of Tokyo.*

<sup>1</sup>*g.toohey@uq.net.au, 0000-0002-2730-1378*

<sup>2</sup>*nguyen026@outlook.com, 0000-0002-2167-7497*

<sup>3</sup>*r.m.vilppola@outlook.com, 0000-0002-0029-5692*

<sup>4</sup>*wq43@cornell.edu, 0000-0001-6461-7243*

<sup>5</sup>*liwenjing@csis.u-tokyo.ac.jp, 0000-0003-2590-5837*

<sup>6</sup>*d.luo@uq.edu.au, 0000-0003-1760-0451*

**Abstract.** Empirical cycling data from across the world illustrates the many barriers that car-dependent cities face when implementing cycling programs and infrastructure. Most studies focus on physical criteria, while perception criteria are less addressed. The correlations between the two are still largely unknown. This paper introduces a methodology that utilises computer vision analysis techniques to evaluate 15,383 Google Street View Images (SVI) of Brisbane City against both physical and perception cycling criteria. The study seeks to better understand correlations between the quality of a street environment and an urban area's 'bicycle-friendliness'. PSPNet Image Segmentation is utilised against SVIs to determine the percentage of an image corresponding with objects and the environment related to specific cycling factors. For physical criteria, these images are then further analysed by Masked RCNN processes. For perception criteria, subjective ranking of the images is undertaken using Machine Learning (ML) techniques to score images based on survey data. The methodology effectively allows for current findings in cycling research to be further utilised in combination via computer visioning (CV) and ML applications to measure different physical elements and urban design qualities that correspond with bicycle-friendliness. Such findings can assist targeted design strategies for cities to encourage the use of safer and more sustainable modes of transport.

**Keywords.** Bicycle-friendly; Quality Streetscapes; Active Living; Visual Assessment; Computer Visioning; Machine Learning; SDG 3; SDG 11.

## 1. Introduction

Australia is one of the highest carbon consumers in the world (Global Change Data Lab, 2020) where transport emissions comprise the third-largest contributor of greenhouse gas nationally (Department of Industry, Science, Energy and Resources, 2021). Private vehicles contribute to approximately 80% of transit activity across major cities such as Brisbane City (Australian Bureau of Statistics, 2018). Making the need for more sustainable transport in Australia vital to achieve key targets of the United Nations Sustainable Development Goals 11.2 and 3.6 (United Nations, 2021).

It is widely documented that cycling as part of a person's daily routine can reduce greenhouse gases associated with car usage, whilst also significantly improving physical and mental wellbeing (TNMT, 2021, Qja et al., 2011). Despite this, the uptake of cycling has been slow in Australian cities such as Brisbane.

Empirical cycling data from across the world illustrates the many barriers that car-dependent cities face when implementing cycling programs and infrastructure. This paper highlights how such findings can be further utilised and analysed at macro-scales, to identify and evaluate key factors that contribute to the overall bicycle-friendliness of cities such as Brisbane City. Through this methodology, targeted design strategies could be better informed to enable and encourage more cycling activity.

## 2. Literature Review

### 2.1. OBJECTIVE AND SUBJECTIVE MEASURES

Objective methods of analysis provide important findings into the general physical features required to enhance cycling environments and encourage more activity. It is becoming more prevalent, however, that physical elements alone fail to consider the subtle role that overall and underlying perceptions of an environment can have on levels of activity (Forbes-Mitchell and Mateo-Babiano, 2015, Osborne and Grant-Smith, 2017, Rossette et al., 2017, Winters et al., 2012).

Perceptions can be difficult to measure and will vary between demographics, however, studies such as Ewing and Handy (2009) have begun to successfully quantify subjective urban design elements and by doing so, effectively measure built-environment perceptions based on human behaviours.

### 2.2. COMPUTER VISION AND MACHINE LEARNING IN STREET MEASURES

A number of recent studies have demonstrated how the use of computer vision and machine learning (ML) in quantifying social science fundamentals can effectively predict, undertake and illustrate micro-urban analysis of environments at a macro scale (Naik et al. 2014, Qiu et al. 2021, Yin and Wang, 2016). Naik et al. (2014), for example, successfully measured perceived safety, utilising and converting survey data on urban perceptions to predict the perceived safety scores of streets across 21 cities worldwide (Qiu et al. 2021). Qiu et al. (2021) further measured four subjective, key urban design qualities and validated their associated scoring through objective points of interest data.



### 2.3. BIKEABILITY MEASURES

Much of existing literature on cycling either explores the quality and availability of physical infrastructure, perceptions of safety and risk, or local urban form, however, it remains unclear how these factors correlate at a macro-scale (Forbes-Mitchell and Mateo-Babiano, 2015, Osborne and Grant-Smith, 2017, Rossette et al., 2017, Winters et al., 2012).

Despite current data and research highlighting the common factors that encourage or deter cycling activity, further analysis is required to understand their correlations and affect in determining the bicycle-friendliness of an environment. Drawing from these methodologies and findings, our study utilises CV and ML technology to effectively evaluate Brisbane's cycling environment.

## 3. Methodology

### 3.1. CONCEPTUAL AND ANALYTICAL FRAMEWORK

Figure 1 illustrates the framework established to evaluate the bicycle-friendliness of a selected area, of which this study comprises 22 suburbs of inner Brisbane City.

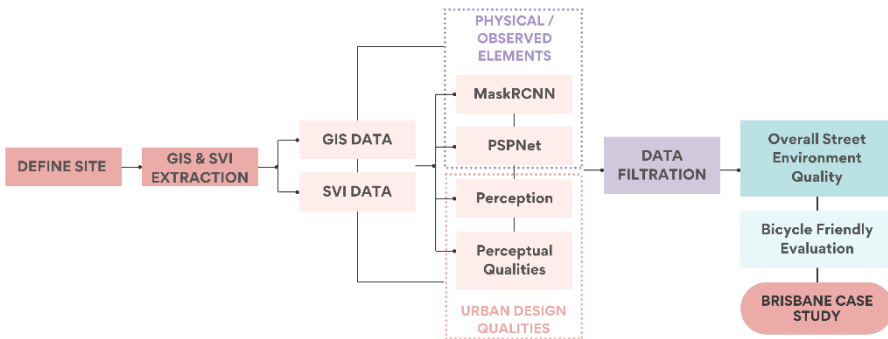


Figure 1. Conceptual Analytical Framework

### 3.2. OVERARCHING DEFINITIONS

#### 3.2.1. Physical and Observed Elements

Most studies utilise objective quantities to measure varying cycling environments and draw conclusions about their qualities (Osborne and Grant-Smith, 2017). This study also identifies observed elements which may not be identified as easily through visual observation but can be evaluated through data collection and extrapolation.

Brisbane specific survey data and international research publications have identified several physical and observed elements as being influential in modal choice. Drawing from these findings, the below elements were selected for inclusion in the study's methodology.

- Collision Numbers: The number of road collisions reports involving bicycles
- Speed: The maximum speed that motorised vehicles can travel on the road
- Bikelane: The existence of bicycle lanes
- Sidewalk: The existence of footpaths
- Streetlight: The street lighting percentages
- Traffic Volume: The number of total vehicles including cars, trucks, motorcycles.

### 3.2.2. *Urban Design Qualities*

Personal perceptions of physical elements inform subjective individual measures relating to senses of safety, comfort, and pleasure (Ewing et al., 2006, van Hagen, 2019). A person's experience of a cycling environment can therefore be greatly influenced by senses of safety, where people 'dare to cycle' and senses of comfort and attractiveness where people are 'invited to cycle' (van Hagen, 2019). Research in the Netherlands has shown that attractiveness appears to be more important than speed, easiness, and comfort to the majority of bicycle riders (Van Hagen, 2019). Therefore, attractiveness is an important consideration which can lead to more subtle qualities of the built environment and urban design integrating with physical elements to achieve better bikeability. Existing guidance on designing bicycle infrastructure in accordance with principles of comfort and attractiveness include the consideration of perceptual qualities such as greenery, openness and aesthetics (Ewing et al., 2006, Ewing and Handy, 2009, Groot, 2016, Van Hagen, 2019). Drawing from such findings, the below urban design elements were selected for inclusion in the study's methodology.

- Visual Order: Based on the consistency of physical elements including arrangement of buildings, paving materials, broken glass, character and scale (Ewing et al., 2006, Ewing & Handy, 2009, Griew et al., 2013, Rundle et al., 2011, Qiu et al., 2022)
- Aesthetic (Imageability): Based on the distinctness in arrangement of physical elements and whether it captures emotions, impressions and/or attention (Ewing et al., 2006; Ewing & Handy, 2009; Ma et al. 2021; Qiu et al., 2021)
- Ecology (Greenness): Based on the proportionality between physical elements including vegetation and building facade (Ewing & Handy, 2009; Ma et al., 2021)
- Enclosure: Based on the proportionality between vertical height of physical elements and horizontal width of the space (Ewing et al., 2006; Ewing & Handy, 2009; Salesses et al., 2013; Dubey et al., 2016; Ma et al. 2021; Qiu et al., 2021)
- Complexity: Based on the diversity of physical elements including user numbers, architectural and landscape variety (Ewing et al., 2006; Ewing & Handy, 2009; Salesses et al., 2013; Dubey et al., 2016; Qiu et al., 2021)
- Human Scale: Based on the proportionality of between physical elements and humans and the speed that humans move (Ewing et al., 2006; Ewing & Handy, 2009; Salesses et al., 2013; Dubey et al., 2016; Qiu et al., 2021).

### 3.3. WORKFLOW

#### 3.3.1. *Obtaining Data*

Quantum Geographic Information System (QGIS), Google Street View Imagery (SVI) and basic information was extracted from the local road data of Brisbane City, providing a series of points which each include a SVI, coordinates, street name, street type and the maximum speed limit. Additionally, two other datasets were used in the calculation of the street scores, that being the locations of existing cycle paths and historic locations of collisions between cyclists and other vehicles. Each SVI is assigned a score based on its locality to either one of these data sets.

#### 3.3.2. *Image Processing*

Once the SVIs were extracted, a series of image processing was applied to further quantitate the images for analysis. Image segmentation was used to extract pixel ratios from SVIs of individual qualities, and Mask Region-Based Convolutional Neural Networks (Mask R-CNN) applied to count the number of specific objects. Subjective ranking of the SVIs could then be achieved according to various qualities through the use of ML.

The first process uses the pyramid scene parsing network (PSPNet) identifying the pixel ratios between each element in the image with a high accuracy, capable of distinguishing between streetscape elements such as trees, sky, roads, and buildings (Zhao et al., 2017). For this research, measuring of street lighting and sidewalks were later used in the evaluation as part of the objective analysis.

The second process used the Mask R-CNN, a powerful deep-learning framework with the task of instance segmentation (He et al., 2017, Qiu et al., 2021). This provides the objective analysis with a dataset of the number of vehicles and people in each SVI.

The final image processing was the application of ML in determining if an SVI portrays certain abstract qualities such as order, aesthetic, ecology, enclosure, complexity and human scale (Ewing et al., 2006, Ewing and Handy, 2009, Griew et al., 2013, Qiu et al., 2021, Qiu et al., 2022, Rundle et al., 2011). The ML process of comparing the elements and converting the preferences to street scores used Microsoft's skill based training system, Truskill algorithm, based on a series of surveys of 300 images from a number of university students, details of which can be found in Qiu et al., 2021. Once the images process was complete, the values assigned to each of the SVI were compiled for data manipulation. Further details of the scoring process can also be found in Qiu et al., 2021.

#### 3.3.3. *Compilation and Ranking*

Each step of the image processing produces a set of data points linked to the SVI and its coordinates, these variables are filtered down to the following:

- Basic information: Coordinates, Road typology, Surface material, Maximum speed
- PSPNet: Streetlight, Sidewalk
- MaskRCNN: Volume of cars, trucks, busses, motorcycles and cyclists

- Subjective Scoring: Order, Aesthetic, Ecology, Enclosure, Complexity, Scale

For each of the variables, the scores are rated on a scale from Poor (0) to Good (5) through a process of scaling the current domain ( $Low_1, High_1$ ) to the target domain ( $Low_2, High_2$ ), otherwise known as reparameterization. These scores enable evaluation of overall street environment quality.

Achieving the reparameterized value ( $y$ ), the following formula was used on the set of data,  $X$ :

$$\begin{aligned} \text{For } High_1, f(x) &= \max \{x \in X\} \\ Low_1, f(x) &= \min\{x \in X\} \end{aligned}$$

From the reparameterized values, each variable is totalled and further reparameterized to a domain between 0 and 1, with 1 being the highest value in order to determine the SVI's 'cumulative street score' and effectively measure bicycle friendliness.

$$y = Low_2 + (x - Low_1) \times \frac{(High_2 - Low_2)}{(High_1 - Low_1)}$$

Finally, the model is further validated by comparing the 'cumulative street score' results with cycling activity data from the subject area (Brisbane City). Due to the vast amounts of data, a computational workflow is needed to manage each of the suburbs individually before assembling the data together again. The initial analysis of the site was computed in QGIS, which extracted the SVIs for image processing done in Python and produced numerous datasets which could be curated and cleaned in Rhino / Grasshopper for efficiency.

#### 4. Results

Importantly, due to timing constraints attributed to the large amount of data requiring further validation with raw cycling activity data, the below results are preliminary. Nonetheless, figures 3 and 4 provide valuable insight into this model's ability to identify key correlations influencing an urban areas' bicycle-friendliness.

Figures 3 and 4 provide example street results, illustrating the differences between a high street score, in terms of its physical and urban design elements, versus a low street score. Figure 2 below reiterates the model's bicycle-friendly evaluation method:

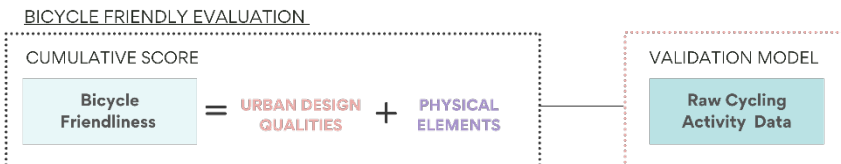


Figure 2. Bicycle-friendly Evaluation - Framework



Figure 3. Bicycle-friendly Cumulative Street Score Results - Example Low Street Score



Figure 4. Bicycle-friendly Cumulative Street Score Results - Example High Street Score

Despite the above results being preliminary, we are already beginning to see trends hinting at a general lack of physical elements that could improve overall bikeability such as separated cycling lanes, sidewalks and lighting. Therefore, an immediate planning priority can be made in relation to implementing more cycling infrastructure.

Additionally, main roads which are identifiable by the higher speed limits, are alluding to a strong relationship with other physical and observed elements of higher

collision numbers and vehicle activity. They also seemingly show a negative relationship with key urban design qualities such as scale, aesthetic (imageability) and ecology (greenness). Such an evaluation could assist engineers identify roads with high collision rates and implement varying safety measures to appropriately address these issues.

This model demonstrates how a number of key trends for Brisbane's overall bikeability at a macro scale can quickly be determined. Based on our preliminary street view analysis alone, results suggest SVIs resulting in a high (cumulative) street score typically score well among most urban design elements, despite not necessarily scoring well in physical elements.

## 5. Conclusion

From our preliminary findings, key factors that seem to be having the most influence in bicycle-friendliness in Brisbane City appear to be individual environment's relationship with traffic, where safety plays a key role (whether through separated bikelanes or calmer streetscapes), in hand with urban design qualities such as ecology (greenness) and human scale. A significant hurdle for Brisbane City's bicycle-friendliness is therefore not likely limited by just its lack of physical infrastructure, but also its lack of key urban design elements that ensure a space exudes a combined sense of safety, comfort and attractiveness.

### 5.1. EFFECTIVENESS OF PROPOSED MICRO-MACRO ANALYSIS FRAMEWORK

It is acknowledged that long-existing techniques in micro-urban analysis cannot be completely replicated at a macro-scale, however, whilst this method may not immediately replace existing techniques, it offers many benefits to planners and policymakers. Utilising and analysing SVI datasets, for example, allows for an analysis reflective of the pedestrian perspective, at varying urban scales and at a significantly lower cost than conventional methods.

### 5.2. LIMITATIONS AND NEXT STEPS

Limitations of this study are primarily associated with time and resource constraints, in addition to general limitations of computer-vision technologies, as detailed below.

#### 5.2.1. Data Limitations

Firstly, elements informing the data collection were limited by time and resourcing, due mainly to limitations in developing necessary computerised scripts for undergoing and validating the analysis. Further studies could therefore benefit from incorporating additional cycling elements into the analysis, as well as validating the cumulative scores with cycling activity data.

The data analysis is also limited to visual senses that can be identified by imagery and GIS mapping data, of which perceptions may differ in reality. For example, the road may appear smooth in imagery but could be uneven in reality, which can influence perceptions of physical elements. Traffic volume data is also likely skewed by counts

being derived from images of varying times of the day. Future studies may benefit from advances in technology associated with street view imagery.

Furthermore, limited by time and resourcing, Brisbane specific survey data was not collected or incorporated into the study and as such, subjective measures relied on existing research relating to perceptions and behaviours from around the world. Perception data from across the world may not accurately reflect perceptions of the local demographics of Brisbane City. Future studies could therefore benefit from undertaking additional surveys in their respective subject areas, to better align perception influences with local demographics.

### 5.2.2. Further Analysis

Using computer vision and machine learning techniques, a micro-level urban analysis of cycling environments can be undertaken at a macro-scale by combining objective and subjective measures of bikeability. By doing so, correlations between objective and subjective influences can be evaluated, and an insight into the overall bicycle-friendliness of urban areas demonstrated.

Future studies could establish additional scripts to incorporate missing bikeability elements and further analysis could be undertaken through additional surveying of human behaviours specific to the subject area and / or targeted design strategies to refine and validate the model's findings.

## References

- Australian Bureau of Statistics. (2018). *Census of Population and Housing: Commuting to Work - More Stories from the Census, 2016*. Commonwealth of Australia. Retrieved May 20, 2021, from <https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/2071.0.55.001~2016~Main%20Features~Feature%20Article:%20Journey%20to%20Work%20in%20Australia~40>.
- Department of Industry, Science, Energy and Resources. (2021, May 25). *National Greenhouse Gas Inventory Quarterly Update: December 2020*. Commonwealth of Australia. Retrieved May 31, 2021, from <https://www.industry.gov.au/data-and-publications/national-greenhouse-gas-inventory-quarterly-update-december-2020#annual-emissions-data>.
- Dubey, A., Naik, N., Parikh, D., Raskar, R., & Hidalgo, C. A. (2016). *Deep Learning the City: Quantifying Urban Perception at a Global Scale*. Computer Vision – ECCV 2016 Lecture Notes in Computer Science, 196-212. doi:10.1007/978-3-319-46448-0\_12.
- Ewing, R., Handy, S. (2009). Measuring the Unmeasurable; Urban Design Qualities Related to Walkability. *Journal of Urban Design*, 14(1), 65-84. doi:10.1080/13574800802451155.
- Ewing, R., Handy, S., Brownson, R. C., Clemente, O., & Winston, E. (2006). Identifying and Measuring Urban Design Qualities Related to Walkability. *Journal of Physical Activity and Health*, 3(S1). doi:10.1123/jpah.3.s1.s223.
- Forbes-Mitchell, J. & Mateo-Babiano, I. (2015). Cycling to Work and the Gender Gap in Brisbane: a study of the environmental, sociocultural and individual determinants of gender disparity in commuter cycling in inner-Brisbane. In *7th State of Australian Cities National Conference 2015*. State of Australian Cities Conference (SOAC).
- Global Designing Cities Initiative. (2021). *Traffic Calming Strategies*. Retrieved November 24, 2021, from <https://globaldesigningcities.org/publication/global-street-design-guide/designing-streets-people/designing-for-motorists/traffic-calming-strategies/>.

- Groot, R. D. (2016). *Design manual for bicycle traffic*. Ede, The Netherlands: CROW.
- He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 2961-2969). IEEE international conference on computer vision (ICCV).
- Hull, A. & O'Holleran, C. (2014). Bicycle infrastructure: can good design encourage cycling? *Urban, Planning and Transport Research: An Open Access Journal*, 2(1), 369-406. <http://dx.doi.org/10.1080/21650020.2014.955210>.
- Ma, X., Ma, C., Wu, C., Xi, Y., Yang, R., Peng, N., & Ren, F. (2021). Measuring human perceptions of streetscapes to better inform urban renewal: A perspective of scene semantic parsing. *Cities*, 110, 103086. doi:10.1016/j.cities.2020.103086.
- Naik, N., Philipoom, J., Raskar, R., & Hidalgo, C. (2014). Streetscore-Predicting the perceived safety of one million streetscapes. In *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 793-799). IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- Oja, P., Titze, S., Bauman, A., Geus, B. D., Krenn, P., Reger-Nash, B., & Kohlberger, T. (2011). Health benefits of cycling: A systematic review. *Scandinavian Journal of Medicine & Science in Sports*, 21(4), 496-509. doi:10.1111/j.1600-0838.2011.01299.x.
- Osborne, N. & Grant-Smith, D. (2017). Constructing the cycling citizen: A critical analysis of policy imagery in Brisbane, Australia. *Journal of Transport Geography*, 64(1), 44-53. <http://dx.doi.org/10.1016/j.jtrangeo.2017.08.015>.
- Qiu, W., Li, W., Liu, X. & Huang, X. (2021). Subjectively Measured Streetscape Perceptions to Inform Urban Design Strategies for Shanghai. *International Journal of Geo-Information*, 10, 493. doi:10.3390.
- Qiu, W., Zhang, Z., Liu, X., Li, W., Li, X., Xiang, X., Xiaokai, H. (2022). Subjective or objective measures of street environment, which are more effective in explaining housing prices? *Landscape and Urban Planning*, 221, 104358. doi:10.1016.
- Rossetti, T., Saud, V. & Hurtubia, R. (2019). I want to ride it where I like: measuring design preferences in cycling infrastructure. *Transportation*, 46(1), 697-718. doi:10.1007/s11116-017-9830-y.
- Salesses, P., Schechtner, K., & Hidalgo, C. A. (2013). The Collaborative Image of The City: Mapping the Inequality of Urban Perception. *PLoS ONE*, 8(7). doi:10.1371/journal.pone.0068400.
- TNMT. (2021, May 11). *The environmental impact of todays transport types*. TNMT. Retrieved November 24, 2021, from <https://tnmt.com/infographics/carbon-emissions-by-transport-type/>.
- United Nations. (2021). *Sustainable Development Goals*. Department of Economic and Social Affairs. Retrieved October 20, 2021, from <https://sdgs.un.org/goals>.
- Van Hagen, M., Govers, B., & Coffeng, G. (2019). Dare, able and invited to cycle! The pyramid of train customer needs applied to cycling policy. In *Association for European Transport 2019 European Transport Conference*. European Transport Conference (ETC).
- Winters, M., Babul, S., Becker, H.J.E.H., Brubacher, J.R., Chipman, M., Crompton, P., Cusimano, M.D., Friedman, M., Harris, A., Hunte, G., Monro, M., Reynolds, C.C.O., Shen, H., & Teschke, K. (2012). Safe Cycling: How Do Risk Perceptions Compare With Observed Risk? In *Canadian Journal of Public Health*, 103(S3), S42-S47. doi:10.2307.
- Yin, L.; Wang, Z. (2016). Measuring Visual Enclosure for Street Walkability: Using Machine Learning Algorithms and Google Street View Imagery. *Appl. Geogr.* 2016, 76, 147–153.
- Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). Pyramid scene parsing network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2881-2890). IEEE Conference on Computer Vision and Pattern Recognition (CVPR).



# TRANSIT-ORIENTED DEVELOPMENT ASSISTIVE INTERFACE (TODAI)

*A Machine Learning-Powered Computational Urban Design Tool to Enhance TOD Planning Processes*

GARRY ZHANG<sup>1</sup>, LEO LIN MENG<sup>2</sup>, NICOLE GARDNER<sup>3</sup>, DANIEL YU<sup>4</sup> and MATTHIAS HAEUSLER<sup>5</sup>

<sup>1,2,3,4,5</sup> UNSW / Computational Design

<sup>2</sup> HDR Inc./Data-Driven Design

<sup>1</sup>z5241635@ad.unsw.edu.au, 0000-0002-9855-6845

<sup>2</sup>lin.meng@student.unsw.edu.au, 0000-0002-6279-9052

<sup>3</sup>m.haeusler@unsw.edu.au, 0000-0002-8405-0819

<sup>4</sup>n.gardner@unsw.edu.au, 0000-0001-6126-6716

<sup>5</sup>daniel.yu@unsw.edu.au, 0000-0002-7788-548X

**Abstract.** Transit-oriented Development(TOD) is widely regarded as a sustainable development paradigm for its sensible space planning and promotion of public transit access. Research in providing decision support tools of TOD may contribute to the Sustainable Development Goals, especially towards sustainable cities and communities (SDG goal 11).While the existing Geographic Information System(GIS) approach may well inform TOD planning, computational design, simulation, and visualisation techniques can further enhance this process. The research aims to provide a data-driven, computational-aided planning support system (PSS) to enhance the TOD decision-making process. The research adopts an action research methodology, which iteratively designs experiments and inquires through situating the research question in real-world practice. A work-in-progress prototype is provided - Transit-Oriented Development Assistive Interface (TODAI), along with an experiment in a newly proposed metro station in Sydney, Australia. TODAI provides real-time visualisation of urban forms and analytical data indicators reflecting key considerations relevant to TOD performance. A regressive machine learning model (XGBoost) is used to make predictions of analytical indicators, promptly producing outcomes that may otherwise require a costly computational operation.

**Keywords.** Urban Planning; Transit-Oriented Development; Planning Support System; Machine Learning; SDG 11.

## 1. Background

Against the backdrop of burgeoning urbanisation, cities are becoming packed with people and vehicles. With overpopulation and traffic congestion becoming problems in modern cities, it is increasingly vital to make urban development sustainable. According to the United Nations New Urban Agenda, by 2030, 60% of the world's population will live in cities, which will be doubled by 2050 (New Urban Agenda, 2016). As is highlighted by the United Nations in its 2030 Agenda for Sustainable Development, it is one of our top priorities to make cities and human settlements inclusive, safe, resilient, and sustainable (goal 11).

Transit-Oriented Development (TOD), a term first coined by Peter Calthorpe (Calthorpe, 1993), is one of the planning paradigms that has gained increasing popularity among many cities worldwide. TOD describes developments near transit hubs, often characterised by compactness, walkability, and mixed usage. Many claims that TOD reduces car reliance by building walkable communities, improving public transit access, and encouraging well-thought-out spaces. There are many convergences between the targets of SDG Goal 11 and aims of TOD, such as providing access to public transits (target 11.2), providing public spaces (target 11.7), providing spaces that support positive social-economical outcomes (target 11. a).

### 1.1. TRANSIT-ORIENTED DEVELOPMENT (TOD)

Transit-Oriented Development (TOD) has attracted growing attention in academia and practice in the past decades. The development aims to reduce automobile usage by providing higher density in areas with good access to public transport and locating public amenities within walkable distance (Lima et al., 2016, Ibraeva et al., 2020). Implementing TOD to a targeted site often requires thorough site research and analysis and introducing multi-aspect attributes as indexes for factor analysis, and evaluating the potential impacts on local ridership and urban characteristics (Cervero and Kockelman, 1997). Cervero et al. (1997) defined 3Ds in evaluating local travel demand: Density, Diversity and Design. Furthermore, the study established 13 variables related to the 3Ds, analysed them with coefficient values, and concluded that better transportation accessibility, squared urban layout, and concentrated mix-used facilities positively promoted less mobile vehicle usage. Several other researchers have introduced similar indexes (Sung and Oh, 2011, Zhou et al., 2020, Ewing et al., 2017). The evaluation process across these multidimensional aspects will influence the precision and comprehension for the planning, design, and implementation process for a successful application of Transit-Oriented Development (Thomas and Bertolini, 2017).

### 1.2. PLANNING SUPPORT SYSTEM (PSS) FOR TOD

PSS describes the integration of Geographical Information System (GIS) with the specific use in modelling for different planning schemes (Harris and Batty, 1993). Researchers had started implementing such a system in TOD and generated concrete presentations in visualising and optimising urban forms and layouts. Lima et al. (2016) explored the TOD urban generation process with Rhino and Grasshopper, employed the TOD attributed into the optimising algorithm and provided a clear visualisation for

the audience to understand urban optimisation outcome based on the theory of TOD. This interface transformed what Cervero had discussed in the 1990s, which only provided a complex and extensive number of figures into a state-of-the-art planning interface while also can be easily altered and reconstructed based on different inputs. Similarly, Pettit (Pettit, 2005) had also implemented PSS in a case study scenario to link and predict land-use relationships with the algorithms in predicting potential local growth with GIS data. In such a way, PSS enables planners to not only integrate data but also work collaboratively with the stakeholders.

### 1.3. MACHINE-LEARNING-POWERED TOD

Machine learning techniques are applied in various TOD and urban planning researches. Machine learning improves the ability to optimise and predict potential input attributes and outcomes. In the project named CityMatrix, Zhang et al. (2018) produced an interactive block model for an indicative urban layout, which adapts the user input with a slider that controls the overall height of the building and population density simultaneous visual feedback for building ratings and city performance. Machine learning was used to make city performance predictions based on previous agent-based simulation data. However, CityMatrix was based on an orthogonal grid urban layout, which may not be sufficient for representing cities where existing urban fabrics are organic. A similar approach has also been used by Lima et al. (2016), who used a genetic algorithm to relocate amenities, resulting in better walkability performance. However, the research focused on factors related to accessibility. Other factors related to TOD, such as development area, population, the urban form, can also be included in such PSS.

In summary, there was abundant research in evaluating and optimising for part of the process of TOD. However, there is not yet a dedicated computational design tool to reflect on several features of the TOD, especially urban density, mixed-use conditions, and improved walkability. Our research will test the possibility of using machine learning algorithms to provide an advanced TOD assistive interface.

## 2. Methodology

In this research, the action research methodology is applied along with the design research method. A thorough iterative design process is framed to experiment with urban geographical information data. Such methodology aims to generate TODs outcomes with the appropriate datasets to encompass the research aim as mentioned and develop a planning support system that reflects key TOD features.

The 'key infrastructure' within the action research process is that the 'research scientist' work collaboratively with the 'practitioner that oriented to the practical problems as the 'organisational scientists' (Baskerville and Wood-Harper, 1996). Therefore, this research is partnered with industry practice HDR inc. to investigate real-world problems when developing the TOD interface.

The plan to be conducted as 'action taking' is within the cycle of the action research process (Susman, 1983). Moreover, while designing out the features in Grasshopper, the industry partner HDR inc. also proves both technical and industry realised issues during the development process during the meeting where the early prototypes are presented. The prototype is tested with a real-world metro location located in Five Dock, Sydney, as one of the newly proposed Sydney metro stations.

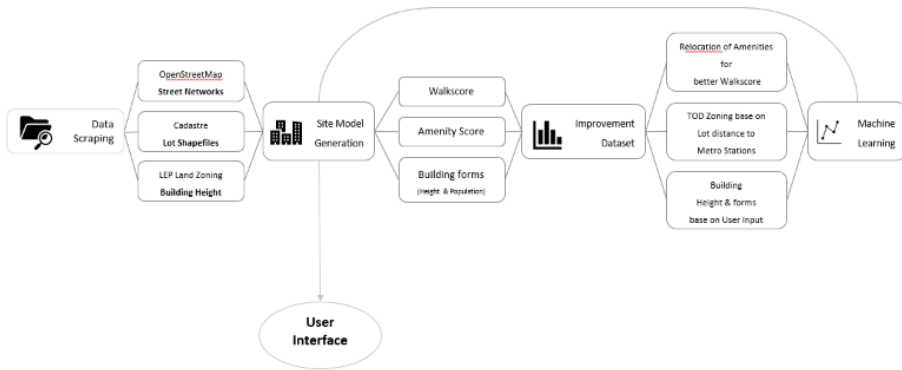


Figure 1. TODAI Development Framework

After defining the research aim, we seek to plan the actions for developing the back-end design and development workflow, which includes the study of the previous TOD principles as identified in other research papers; the features of TODs to develop on; and the primary Grasshopper processes for realising such features. The plug-in Urbano is also used for importing GIS data from OpenStreetMap as well as the tool to analyse some features of TODs such as walkability (walkability score) and amenity popularity (amenity score), Galapagos is used for improving the urban forms and amenity locations for better walkability and all to be visualised by using Human UI as the plug-in for developing the TOD assistive interface.

### 3. Transit-Oriented Development Assistive Interface

#### 3.1. DATA COLLECTION

The foremost step in developing the TOD model is to collect useful geographical data with embedded metadata. The geographical data is used as we are only redeveloping the urban forms in this research rather than taking away the street layouts.

As shown in Figure 1, the three primary geographical datasets are acquired. Street networks, existing amenity locations, and their metadata are acquired from OpenStreetMap. The cadastral boundaries are stored as Esri shapefiles, made available from the government's open data platform. Planning constraints, such as zoning and building height limits, are processed and made publicly available from the AURIN platform.

#### 3.2. TOD SITE ANALYSIS

Two TOD-specific analyses of the existing urban fabric are conducted when developing the TOD urban forms. In such a manner, the improvements of the urban forms would be generated based on the identified existing urban issues. With the reallocation of amenities and transit-centred development, TOD seeks to promote better walkability to transit stations and local amenities. Hence, less auto-mobile usage is needed, and the primary transportation method will be transformed to walking. This section elaborates the analysis based on urban walkability, urban density, and urban design.

### 3.2.1. Walkability Analysis

The process of generating the walkability calculation is by calculating the distance from the locations of the local amenities to the transit centre / local amenities; this character has also been discussed in the past decade (Lamour et al., 2019, Singh et al., 2017, Sung and Oh, 2011). It is generally encouraged to locate residential buildings near where the walkability score is high (Hall and Ram, 2018, Dogan et al., 2018). Hence, the initial step in the analysis process is to calculate the walkability in the case study site. The calculation of walkability can be seen as calculating the average of the networked distance between origins (residence) and destinations (amenities), weighted by the possibilities of visiting each amenity during the week (amenity index).

To further realise the result from the walkability score, the overall visualised result is completed, the Grasshopper process is shown in Figure 2.

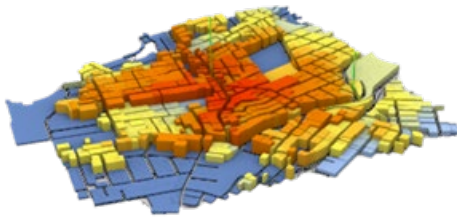


Figure 2. Left: Existing Site Walkability Score Visualisation

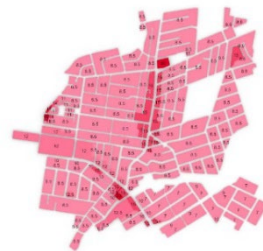


Figure 3. Right: Existing Building Height Limit

### 3.2.2. Building Height Limit Analysis

Zoning requirements under the New South Wales legislation of Local Environmental Plans (LEP) directly relate to the building height limits and building types to build on different land lots. The zoning revitalisation under the principle of TODs is also a crucial element in providing various housing types for different income groups, and essentially, is linked to the local real estate value (Ibraeva et al., 2020). The regeneration of the land zoning could enhance the densification and mix-used zone around the transit centre while providing affordable housing around the rear area to the transit centre (Ibraeva et al., 2020). This definition of zoning has also been discussed in the previous sections.

To compare the existing zoning and the TOD featured redevelopment on the case study site. An analysis of the existing zoning condition has been visualised using the dataset imported from the LEP data source. Furthermore, the height limits on the current conditions of building lots have been visualised. This building height limit data is also visualised and compared in the interface. The height limit visualisation is shown in Figure 3.

### 3.2.3. TOD Based Urban Redevelopments

Based on the two analyses completed, a TOD based generation process is followed. In this section, two improvements will be discussed: relocation of the amenities and the extra amenity types added for a better walkability score; redevelopment on local land zoning and building forms generation. The systematic improvements can build an overall TOD improved local blueprint and operated as part of the final assistive interface.

### 3.2.4. TOD height zoning and Building forms generation

The existing urban fabric at the study site exhibits small and segregated building lots due to historic settlement. Building sites are first amalgamated algorithmically before further process.

After the amalgamated land parcels are generated, an optimisation process is developed to minimise the overall shadow casting by the generated built forms using a genetic algorithm through Galapagos. This optimisation reduced the shadow area in the study area, especially around the high-density development areas, so that residents and workers may have adequate access to sunlight for wellbeing. (Figure 4).

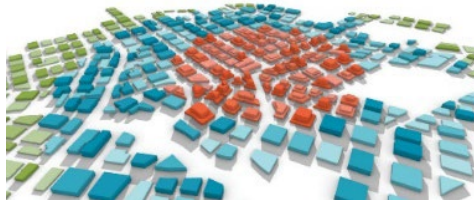


Figure 4. TOD Zoning Allocation

### 3.2.5. Relocation of amenities for better walkability score

Further optimisation of the walkability is conducted by relocating the amenities using genetic algorithms inspired by (Lima et al., 2016). The objective function minimises the networked distance between residential buildings and public amenities. After the walkability score analysis, an overview of the existing situation is visualised.

There were 39 different amenities on the original site. In this step, some more amenities are added to improve the walkability performance. Moreover, in the end, 55 amenities are located on the TOD enhances the site. The final 'walkability score' has also been increased due to such an event. (See Figure 6)



Figure 5. Left: Improved Walkability Score



Figure 6. Right: Galapagos Optimization for Amenities Allocation

### 3.2.6. Assistive Interface

Finally, an interface with user input and a data dashboard is developed to reflect on the TOD planning and development indicators discussed in Section 3.2, including zoning, density, and walkability score. The interface is developed using HUMAN UI in Grasshopper. All passive factors for changing the building forms and amenity locations are integrated into the interface to manipulate the TOD redevelopment process as desired. At the same time, the geometries will also be visualised in the Rhino preview window (Figure 7 and Figure 8).



Figure 7. TOD Urban Feedback Interface

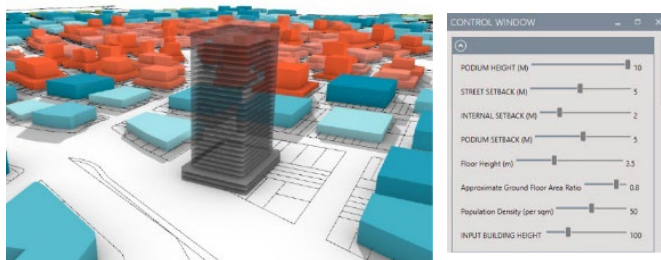


Figure 8. User-directed building input

### 3.2.7. XGBoost - Regressive machine learning

Regressive machine learning models learn the numeric representation of the given dataset and make predictions based on input values by mapping the data domain.

Compared to complex simulations like Walkscore, making predictions through a trained machine learning model is often more streamlined and takes significantly less time to compute. XGBoost (short for eXtreme Gradient Boosting) is an ensemble learning algorithm and is a tree-boosting technique (Chen & Guestrin, 2016). XGBoost ensembles many smaller learning models with parallelisation, gradient boosting, sparse awareness. These features make XGBoost one of the more popular machine learning algorithms for regression with less requirement in data quality. A low-code XGBoost implementation named H2O is selected for our machine learning training.

Firstly, a training dataset is collected. By enumerating all possible user inputs in Grasshopper, a training dataset is automatically captured in a CSV file, including its resultant indicators corresponding to each possible input combination. Secondly, the CSV file is consumed by using H2O. Each feature column is labelled with input (input features) or output (targets) for training. A trained XGBoost model is produced after the dataset is fitted. The trained model is loaded in H2O and ready to predict by giving new input.

In the end, a socket script is set up for bi-directional communication with the H2O platform. Grasshopper sends the input data to the H2O platform for prediction upon the new user input. The H2O predicts based on the given values and returns a numeric result. This result is sent back to TODAI and is presented to the user in near real-time. This process has significantly reduced the time required to show the result for TODAI. However, it is noted that further work is required to produce a robust prediction model.

#### **4. Discussion**

This research proposes a framework to provide a data-driven, computational planning support system for Transit-oriented Development. This framework contains the following steps: data collection through publicly available sources, rule-based urban model generation, simulation and analysis, scenario testing and optimisation, and visualisation through a user interface. The simulation and optimisation are further enhanced using the machine learning technique in order to provide prompt user feedback.

The research showcased an early prototype based on existing urban conditions in the Five Dock area in Sydney to test hypothetical TOD and its potential outcome by the proposed metro station. The urban form with optimised amenity allocation and data dashboard has been reflected in TODAI, giving insights on the TOD planning and development process by providing real-time feedback from given planning input.

However, there are also limitations to this research. First, the walkability score analysis's accuracy relies on a clean street network dataset, and a 5 to 10 % margin of error is assumed with current crowd-sourced OSM data quality. This issue can be overcome by using a manually cleaned street network dataset suitable for walkability analysis.

Moreover, more features can be included to build a comprehensive TOD model, such as the street widths, local transportation big data, intercity public transport connectivity, and vegetation, instead of solely street network, amenity and planning constraint.

The interface is still under development. Future studies can engage with various



computational algorithms within this framework to expand the TODAI interface. Such as optimising the TOD automatic form generation with a convolutional neural network for translating the data to a systematic dataset and reducing the computation power needed for the later form optimisation process than the current genetic algorithm. Cellular automata can generate optimised zoning by integrating the TOD feature and indexes for amenities allocations and building types of automation. Further studies can also cover a variety of data sources that may be relevant to TOD development.

## 5. Conclusion

The exploration of TODAI has indicated the potential of studying TOD's features and characteristics using computational design tools. By using Rhino + Grasshopper, TOD redevelopment algorithms are capable of processing a large dataset and urban geometries and producing an ideal TOD urban optimisation and the outcome.

The research also contributes to planning support systems through computational tool development and machine learning techniques. The interface provides visualised outcomes of the TO development and encourages public participation by enabling non-expert users to explore different input combinations and interactively navigate different scenarios. This tool can also aid application for sustainable future developments in the Architecture, Engineer, and Construction (AEC) professionals to work cooperatively in the feasibility stage of a single project and a masterplan overview.

Two types of machine learning algorithms enhance the process: 1. evolutionary solvers optimise local decisions, where amenity location affects walkability, and built form affects solar access; 2. The regressive ensemble learning model provides an "instant" prediction of the outcome, making real-time feedback in TODAI possible.

By exploring the possibilities of applying the TOD paradigm into a real-world scenario, not only can we promote such tools to more fields and markets, but we can also integrate more urban planning concepts into the same project and produce urban development outcomes on the fly.

## References

- Baskerville, R. L., & Wood-Harper, A. T. (1996, 1996/09/01). A critical perspective on action research as a method for information systems research. *Journal of Information Technology*, 11(3), 235-246.  
<https://doi.org/10.1080/026839696345289>
- Calthorpe, P. (1993). *The next American metropolis: ecology, community, and the American dream*. New York : Princeton Architectural Press.
- Cervero, R., & Kockelman, K. (1997, 1997/09/01/). Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment*, 2(3), 199-219.  
[https://doi.org/https://doi.org/10.1016/S1361-9209\(97\)00009-6](https://doi.org/https://doi.org/10.1016/S1361-9209(97)00009-6)
- Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System.  
<https://doi.org/10.1145/2939672.2939785>
- Dogan, Timur & Samaranyake, Samitha & Saraf, Nikhil. (2018). Urbano - A New Tool to Promote Mobility-Aware Urban Design, Active Transportation Modeling and Access Analysis for Amenities and Public Transport. *simaud.2018*.

- Ewing, R., Tian, G., Lyons, T., & Terzano, K. (2017, 2017/04/01/). Trip and parking generation at transit-oriented developments: Five US case studies. *Landscape and Urban Planning*, 160, 69-78.  
<https://doi.org/https://doi.org/10.1016/j.landurbplan.2016.12.002>
- Hall, C. M., & Ram, Y. (2018, 2018/06/01/). Walk score® and its potential contribution to the study of active transport and walkability: A critical and systematic review. *Transportation Research Part D: Transport and Environment*, 61, 310-324.  
<https://doi.org/https://doi.org/10.1016/j.trd.2017.12.018>
- Harris, B., & Batty, M. (1993). Locational Models, Geographic Information and Planning Support Systems. *Journal of planning education and research*, 12(3), 184-198.  
<https://doi.org/10.1177/0739456X9301200302>
- Ibraeva, A., Correia, G. H. d. A., Silva, C., & Antunes, A. P. (2020, 2020/02/01/). Transit-oriented development: A review of research achievements and challenges. *Transportation Research Part A: Policy and Practice*, 132, 110-130.  
<https://doi.org/https://doi.org/10.1016/j.tra.2019.10.018>
- Lamour, Q., Morelli, A. M., & Marins, K. R. d. C. (2019, 2019/06/01/). Improving walkability in a TOD context: Spatial strategies that enhance walking in the Belém neighbourhood, in São Paulo, Brazil. *Case Studies on Transport Policy*, 7(2), 280-292.  
<https://doi.org/https://doi.org/10.1016/j.cstp.2019.03.005>
- Lima, F. T., Kos, J. R., & Paraizo, R. C. (2016). Algorithmic approach toward Transit-Oriented Development neighborhoods: (Para)metric tools for evaluating and proposing rapid transit-based districts. *International journal of architectural computing*, 14(2), 131-146.  
<https://doi.org/10.1177/1478077116638925>
- NCHRP. (2005). *Transit-Oriented Development: Developing a Strategy to Measure Success*.  
<https://doi.org/10.17226/23319>
- Susman, G. (1983) *Action Research: A Sociotechnical Systems Perspective*. In: Morgan, G., Ed., *Beyond Method: Strategies for Social Research*, Sage, Newbury Park, 95-113.
- Pettit, C. J. (2005). Use of a Collaborative GIS-Based Planning-Support System to Assist in Formulating a Sustainable-Development Scenario for Hervey Bay, Australia. *Environment and planning. B, Planning & design.*, 32(4), 523-545.  
<https://doi.org/10.1068/b31109>
- Sung, H., & Oh, J.-T. (2011, 2011/02/01/). Transit-oriented development in a high-density city: Identifying its association with transit ridership in Seoul, Korea. *Cities*, 28(1), 70-82.  
<https://doi.org/https://doi.org/10.1016/j.cities.2010.09.004>
- Thomas, R., & Bertolini, L. (2017). Defining critical success factors in TOD implementation using rough set analysis. *Journal of transport and land use*, 10(1), 139-154.  
<https://doi.org/10.5198/jtlu.2015.513>
- Habitat III. 2022. *The New Urban Agenda - Habitat III*. [online] Available at: <<https://habitat3.org/the-new-urban-agenda/>> [Accessed 13 February 2022].
- Zhang, Y., Grignard, A., Aubuchon, A., Lyons, K., & Larson, K. (2018). *Machine Learning for Real-time Urban Metrics and Design Recommendations*.
- Zhou, J., Yang, Y., & Webster, C. (2020). Using Big and Open Data to Analyse Transit-Oriented Development: New Outcomes and Improved Attributes. *Journal of the American Planning Association*, 86(3), 364-376.  
<https://doi.org/10.1080/01944363.2020.1737182>

# PERCEIVING FABRIC IMMERSED IN TIME

## *An Exploration of Urban Cognitive Capabilities of Neural Networks*

ZHIYONG DONG<sup>1</sup>

<sup>1</sup>*South China University of Technology, State Key Laboratory of Subtropical Building Science*

<sup>1</sup>*ar\_dongzhiyong@mail.scut.edu.com, 0000-0003-0958-2986*

**Abstract.** City develops gradually with the lapse of time. Cities, as a ‘container’, are injected new urban elements along the trajectory of the times and the progress of human civilization, constructing the historical structures involved past, present and future. Thus, the cultural information of each era is preserved in the urban fabric together and urban fabric features are complex and rich, which are difficult to capture in traditional design methods. In this paper, we try to use Generative Adversarial Networks (GAN), one of the neural network algorithms, to explore the inner rules of complex urban morphological features and realize the perception of the urban fabric. Neural networks are innovatively applied to the larger and more complex city generation in this experiment. First, we collect European urban fabric as the dataset, then label data to facilitate machine training, use GAN to learn the feature of the dataset by adjusting parameters, and analyze the effect of the generated results. The automatic feature learning capability of the neural networks is used to summarize the inherent patterns and rules in urban development which is difficult for human to discover.

**Keywords.** Deep Learning; Generative Adversarial Networks; Generative Design; Morphology Cognition; Urban Fabric; SDG 11.

## 1. Introduction

Cities are human settlements that develop gradually over time. Maurice Willmore Barley (1977) has mentioned that any social value and behavior, no matter how abstract they are, can generally be reflected in material forms. And the city is the concentrated expression of geographical environment factors, social value, economy and politics. Colin Rowe (1983) has expressed his opposition to the almost utopian urban design, which is completely detached from the social and temporal elements.

Architects are used to chasing a set of rules as guidelines for design, while the real city is a complex aggregation under the influence of time, which is difficult to sort out and cannot be simply overlapped. Descending to the level of city plan, urban fabric that is immersed in time is also not directly designed by a designer.

Urban fabric is a collage of various periods, part of which are determined by local regulations, part by the monarch's will, and part by aesthetic consciousness. With the continuous development of computer-aided technology, rule-based generative design methods are applied to architectural morphology design at present. Although the applications of logical algorithms take architects out of their inherent thinking, the collection of complex rules in urban design is hard to express by algorithms. Urban designs done under the traditional and rule-based design methods inevitably become the dilemma of Rowe's writing.

The great breakthroughs of artificial intelligence focus on the direction of neural networks and deep learning in recent years. Rather than directly processing the input data, neural networks simulate the neural thinking of the human brain, allowing the machine to automatically analyze and summarize the data feature. The application of neural networks to architecture also provides a new method to architecture generative design. The advantage of this data-driven method to generative design is that it does not require any parameters of the desired design solution to be determined in advance, use the machine to learn the visual characteristics of the existing urban fabric automatically, and then generate novel projects that mimic data feature.

## 2. The complexity of urban fabric and data-driven generation methods

### 2.1. THE COMPLEXITY OF URBAN FABRIC

*Urban Fabric* is the term that describes the physical characteristics of urban areas, which includes the streetscapes, buildings, hard landscape, signage, roads and other infrastructure. According to the Conzen school, what historians and architects roughly call urban fabric is composed of three interrelated elements. First, the city plan, which refers to the street system; second, the plot pattern, the division of land; and third, the layout of buildings under the plot pattern (Jeremy, 2001). City plan referred to here is what M.P. Conzen calls "a system of land division and registration consisting of legally protected land ownership". The land use pattern represents the various uses of ground and space. Finally, urban fabric refers to the actual three-dimensional physical structure based on the subdivision of land parcels.

Kevin Lynch (1960) argues that the city's morphology, actual function, the ideas and values that people give to the city combine to create a miracle. It is the reason that urban fabric built over hundreds of years has shown complicated that is not easy to capture. Lewis Mumford (1961) refers to such urban spaces as "containers" from their origins, cities have been special structures designed to store and transmit the fruits of human civilization. In addition to the complex social factors, urban fabric also contains the architecture and construction ideas of each period. Colin Rowe (1983) attributes the irregularity of the urban fabric to time, and his *collage city* theory is to consider the existing structure of the city as a skeleton, which continuously is injected new urban elements with the spirit of the times. The city involves the historical structure in the present and future renewal construction. So, it is possible to keep the cultural information of each era side by side in the urban

carrier. The early skeleton would “decay and die” over time, and time is the element that designers most easily overlook but are also incapable of capturing.

In addition to time, the order of different periods also determines the urban fabric. Urban historian F. Castagnoli (1971) says, an irregular city is a form that develops when the people who live on the land take full control of the city. If a ruling body had pre-divided the land before handing it over to the users, then another uniform pattern of cities would appear accordingly. Spiro Kostof (1991) attributes this irregular urban fabric to order, whether the controller of the order is the individual people or the city governor. The order also changes constantly over time, the complexity and changes of urban fabric are presented with the advance of time.

## 2.2. DATA-DRIVEN GENERATIVE METHODS

Neural network algorithms bring new ways of thinking for architectural design. In the era of artificial intelligence and data science, data and design mapping mechanism based on the neural nets-oriented and data-driven generative method can not only avoid accidental and one-sided problems due to individual experience but also use neural networks' ability to automatically learn features to find solutions for complex problems (Hinton & Salakhutdinov, 2006). The neural network algorithm automatically summarizes the inherent rules from the input dataset and describes the common characteristics of the data in the form of coding (Bengio et al., 2013), which saves architects the time of data analysis and programming, and can also realize more experience-based architectural projects. In this case, data is the main driving element, and data filtering and processing are the main works that architects need to complete (Hansen et al., 2009). The type and quantity of data will directly determine the result and accuracy of the generated project.

When facing complex urban fabric, it is possible to use the ability of neural networks to automatically learn morphological features to perceive and learn the complex time-immersed cities. Machine learns from the dataset of the existing urban fabric. Driven by a large amount of data, machine summarizes by itself the universal patterns and rules in the complex fabric left by the development of the city.

## 2.3. OBJECTIVE

The interdisciplinary application of computer technology and urban design has a long history. In previous studies, the application of deep learning in the field of architectural design can only address layout problems in which the logic is relatively clear, such as, residential floor plan zoning (Weixin et al., 2018) and buildings layout in small plots (Runjia Tian, 2021). With the continuous development of neural network algorithms, more complex urban generations with unclear rules will also become the focus of discussion, which will provide more comprehensive design guidance.

On this basis, the research purpose of this paper is to explore the strategy of urban fabric perception based on neural networks and to realize the automatic

urban morphology generative design using Generative Adversarial Networks (GAN) model in deep learning technology. After effective training, the generative adversarial network model can quickly generate urban layout solutions that meet the requirements according to the site conditions in a short time, providing candidate solutions while inspiring architects to think about diverse designs in a short time.

In this paper, we take the recent European urban fabric as the experimental sample. Firstly, we collect data on the layout of European cities; then we process the data to facilitate machine learning, keep the urban layout and the function and height data of buildings, and label them. And then we use GAN to train them by adjusting the parameters. Finally, the trained model is used to automatically generate the city morphology by inputting the initial data of the base such as geographical information.

### 3. Methodology

#### 3.1. GENERATING ADVERSARIAL NETWORKS

Based on the data type and the characteristics of neural networks, we use Image-to-Image Translation with conditional Generative Adversarial Networks (Isola, Zhu et al., 2017) as the main algorithm, which is one of neural networks algorithms.

Generative Adversarial Networks (GAN) consist of a Generator and a Discriminator. In the training process, the generator first generates an alternative image from the latent space and passes it to the discriminator. The discriminator takes either the real image or the alternative image as input and tries to distinguish whether the current input is the real data or the alternative data. During the training period, the whole process makes the generator and the discriminator play with each other, and they promote each other until the discriminator considers the result generated by the generator as real. Because of this, the advantages between the two combine to complement each other and thus improve the learning ability of the machine. The application of Generative Adversarial Networks is further enhanced by the proposal of conditional Generative Adversarial Networks (cGAN) (Mirza & Osindero, 2014). Compared with GAN, which are generative models that learn the mapping of random noise  $z$  to output image  $y$ , cGAN learns the mapping of random noise  $z$  and conditional variable  $x$  to output image  $y$ . In this experiment, the conditional variable  $x$  is chosen as the input of image information.













#### 3.2. DATA COLLECTION AND PROCESSING

According to the experimental objectives, urban fabric data that meet the requirements are collected and processed to construct the experiment dataset. Data processing includes data sorting, cleaning, analysis and labeling.

To get the original dataset, we use the open data from OpenStreetMap and batch its map information by using QGIS and python. We select 560 pairs of data from 69 cities, which can represent the typical urban layout of different scales in Europe and have the comparative learning value of Generative Adversarial Networks in

machine learning. We then classify the dataset type of training dataset in the experiment: training dataset A is the image which means basic geographical environment, including the main roads, natural green spaces and water system of the site, and training dataset B is the image which means urban layout within the existing site, including building function, building height, and public space.

In the data processing step, the original city images are simplified in the form of color labeling, with different colors representing various elements in the map (Figure 1). The color labeling acts as a supervised signal that guides the network to process, extract and transform the visual information to maximize the performance of the network on the desired task. In this experiment, the three channels of RGB colors are used to represent building information. R and G channels represent building functions and the B channel represents building height. With labeling of natural elements, roads and rail, these labels can represent most of the factors in urban morphology.

Regional Surface	RGB values	Hex	Opacity	Color
Commercial	255, 80, 0	#FF5000	100%	
Residential	255, 150, 0	#FF9600	100%	
Industrial	255, 255, 0	#FFFF00	100%	
Public	67, 75, 0	#434B00	100%	
height_1-5F	0, 0, 70	#000046	100%	
height_6-10F	0, 0, 120	#000078	100%	
height_10-20F	0, 0, 180	#0000B4	100%	
height_20F+	0, 0, 255	#0000FF	100%	
railway	0, 0, 0	#000000	100%	
road	255, 0, 0	#FF0000	100%	
greenland	160, 255, 54	#A0FF36	100%	
water	183, 255, 255	#B7FFFF	100%	

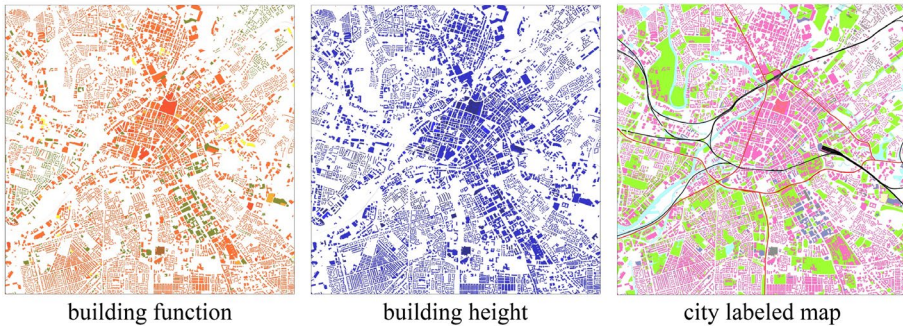


Figure 1: Sample labeling rule

Then, the labeled OSM data are separated according to attribute and layer. As in Figure 2, training dataset is made based on the classification. The range represented by each image is 6.25km<sup>2</sup>.

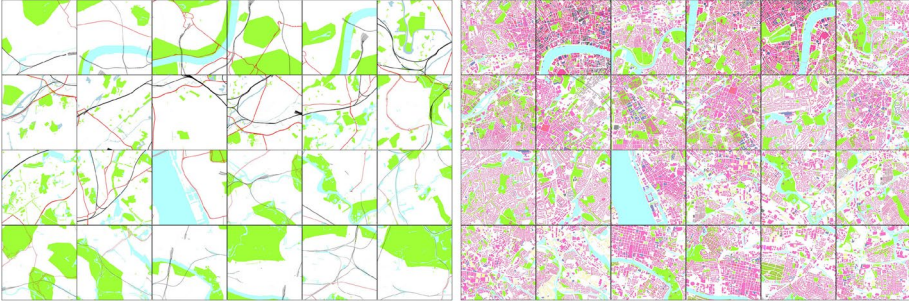


Figure 2: Part of training dataset

### 3.3. TRAINING AND TESTING

Generative Adversarial Networks are learned by using the collected data. As mentioned before, the training sets are divided into two parts: Training dataset A contains the base conditions and initial limitations, (i.e., elements such as roads, railways, green spaces, and waters), and training dataset B contains the design content of the project, (i.e., building functions, building heights and artificial green space).

The images of the training dataset are fed into the GAN model separately, and in order to achieve the learning of large-scale urban fabric, the images are adjusted into  $1024 \times 1024$  pixels to match the data structure of the adversarial network. High-resolution images put new demands on the machine's arithmetic power and the original algorithm, and we achieve effective learning of the machine by adjusting the parameters. In addition, we adjust the number of net layers and the number of neurons in each layer of the neural network, to avoid the frequent overfitting phenomenon in deep learning training and break the limits of machine arithmetic power.

The loss values of the generator and discriminator are recorded in Figure 3 during the training process. The training process is the process of generators and discriminators playing against each other. At the equilibrium point, which is also the optimal solution of the minimax game the discriminator considers the output of the generator as the real data, with a probability of 0.5. Meanwhile, a higher discriminator loss value and a lower generator loss value means that the training process tends to succeed. The training gradient is set to a constant learning rate before 200 epochs, and a decaying learning rate is used after 200 epochs, which facilitates better fitting of the data. From the loss images, it can be found that the losses of the generator and discriminator stabilize after 330 epochs.



As in Figure 4, the generated images which are displayed on the monitor web during the training, show that the generated images are blurred at the 10th epoch, the generated images at the 50th to 160th epoch recorded are gradually clearer and the city streets in the images tend to be more obvious. In the 380th epoch, the machine has learned the urban fabric effectively and grasped the more obvious urban morphological features. Finally, we choose the 400th epoch trained model as the urban generated experimental model. After the model training is completed, inputting the testing data including roads, railways, green spaces, and rivers, the new urban fabric responding to the base conditions can be quickly output.

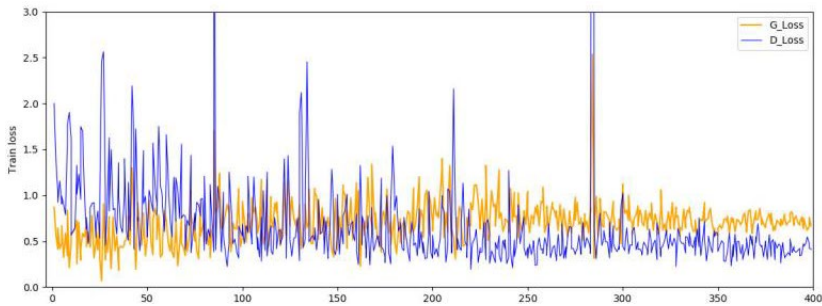


Figure 3: Generator loss (G\_LOSS) and Discriminator loss (D\_LOSS) during training.

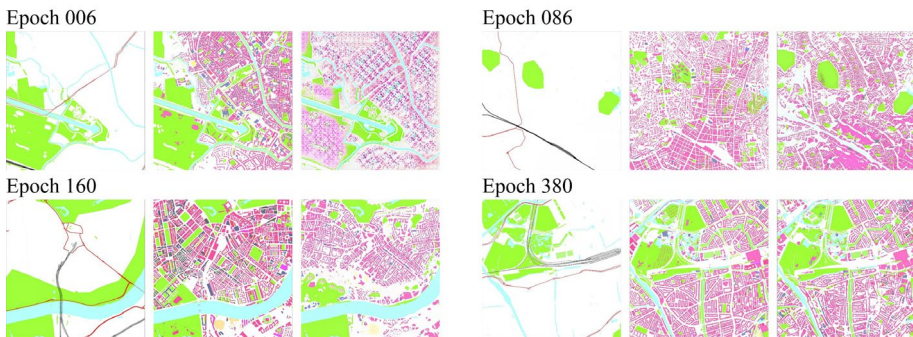


Figure 4: Output results for each training epoch

#### 4. Analysis of results

In this experiment, the existing base information is used to generate urban fabric using neural networks, and the generated image is compared with the existing city image for similarity to verify the generation effect of neural networks. After comparing with Existing city areas, the final model is tried to generate a new urban fabric for undeveloped city areas.

##### 4.1. COMPARING WITH EXISTING CITY AREAS

First, we use the geographic information in the existing cities as the test dataset input. In general, the generated results show different degrees of similarity and difference with the existing urban fabric areas of Liege, Lille, and Vienna (Figure 5) which are selected as the input for comparison.

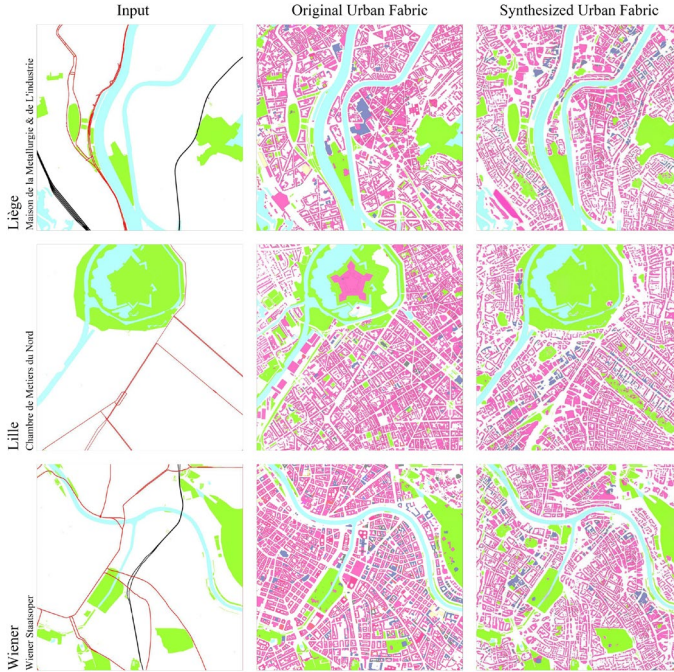


Figure 5: Compare between real images and generated images

From the comparison between Liege's synthetic fabric and the existing fabric, we can find that the original city's baroque road networks with Renaissance characteristics are replaced by an irregular road network; the block scale shows a certain degree of unity; large plots are dismantled; the original high-rise buildings are replaced by low-rise buildings, while high-rise buildings appear along the main roads and the river. In addition, more artificial green spaces appear scattered throughout the city in the synthetic fabric. The radioactive axis network of Lille has also been replaced by a large number of blocks and a regional unification of block scales. The exhibition center, which is more visible in the original city, has been replaced by green space. A continuous green space has appeared along the main road on the east side. Compared to Liege and Lille, the overall layout of the generated fabric of Vienna remains consistent with the existing fabric and shows its consistent diversity, but the number of simulated artificial green spaces increases, the location of green spaces is slightly different from those in the existing fabric, and the functional height of a small number of buildings changes.

#### 4.2. GENERATING FOR UNDEVELOPED CITY AREAS

We briefly summarize the features of existing fabric in the dataset in advance and verify the generated projects by (1) the intrinsic characteristics of the city (e.g., block scale, street layout, building distribution), and (2) the response to the input information (e.g., railroad, highway, water system, public green space).

From the synthesized urban fabric images (Figure 6), they are easy to find that the urban layout maintains the main characteristics of the existing European cities: the block scale is relatively uniform, its side length is controlled at 50-100m; the street layout is mostly a natural form of streets based on the appropriate modulus of housing buildings, which is locally expressed as an axial network layout; the public green space is evenly distributed between the gaps in the blocks; most of the high-rise buildings are distributed along the main roads and water. Apart from that, the generated fabric images also show that buildings setback along the road roads, railroads, and water systems. Railroad stations are generated in the dense areas of railroads. At the same time, both effectively deal with the relationship between synthetic fabric and the base conditions: the reasonable avoidance of roads, the simulation of public green areas. We can also find that the edges of the input natural green areas are partially encroached by buildings and other man-made facilities.

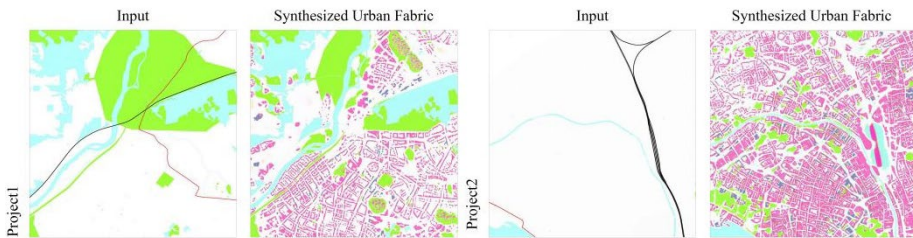


Figure 6: Generated Urban Design Projects

### 4.3. APPLICATION

The above analysis shows that using deep learning, the machine can accomplish the learning of the complex morphology of the city at a large scale. The complex relationships among building forms, building functions, roads, railroads, green spaces, and water systems in the urban fabric are systematically summarized.

In the process of future applications, on the one hand, the trained GAN model will provide architects and designers with rapid generation of urban morphological intentions, and the data-driven design method can avoid the chance and one-sidedness arising from individual experience; on the other hand, the automatic feature learning capability of deep learning perceives the hidden inner laws in urban data, and the urban fabric generated by machine can inspire designers and scholars to think further about the city.

## 5. Conclusion

This experiment has completed the generation of urban morphology based on GAN. It demonstrates that deep learning can be an effective tool in perceiving urban fabric. By collecting and processing the data, the machine summarizes the

general patterns of the traditional European cities, which have lasted for hundreds of years and are immersed in time. Time is reflected in the generated urban fabric, shapes of streets, the scale of blocks and the picturesque traditional streetscapes that are difficult for humans to perceive but relatively easy for machines.

The data-driven perceptive and generative method opens up a new way of thinking for computational architectural design. Although the application of this method in real projects has not been proven, architects can use this method to broaden their horizons with the power of machines. The urban fabric generation tool based on GAN can acquire and process more complex urban elements, quickly providing a preview of new design solutions for urban tasks. From the perspective of design practice, data processing, neural networks adjusting and generated results evaluating is the next step for this research.

## References

- Barley, M. W. (1977). *European towns: their archaeology and early history*. Academic Press.
- Bengio, Y., Courville, A. & Vincent, P. (2013). Representation Learning: A Review and New Perspectives, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, 1798-1828.
- Castagnoli, F. (1971). *Orthogonal Town-Planning in Antiquity*. MIT Press.
- Hansen, C., Johnson, C., Pascucci, V. & Silva, C. (2009). The Fourth Paradigm: Data-Intensive Scientific Discovery. Microsoft Research.
- Hinton, G. E., Salakhutdinov, R. R. (2006). Reducing the Dimensionality of Data with Neural Networks. *Science*. 313 (5786): 504–507.
- Huang, W., Zheng, H. (2018). Architectural Drawings Recognition and Generation through Machine Learning. In *Proceedings of the 38th Annual Conference of the Association for Computer Aided Design in Architecture, ACADIA 2018* (pp. 156–165). The Association for Computer Aided Design in Architecture (ACADIA)
- Isola, P., Zhu, J. Y., Zhou, T. & Efros, A.A. (2017). Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, (pp. 1125-1134).
- Kostof, S. (1991). *Formation of the city: The historical process of urban patterns and urban significance*. China Building Industry Press
- Lynch, K. (1960). *The image of the city*. MIT Press.
- Mirza, M. & Osindero, S. (2014). Conditional generative adversarial nets. *ArXiv*, vol. abs/1411.1784.
- Mumford, L. (1961) *The City in History*. Harcourt Inc.
- Rowe, C., Koetter, F. (1983). *Collage city*. MIT Press.
- Tian, R. (2021). Suggestive Site Planning with Conditional GAN and Urban GIS Data. In *Proceedings of the 2020 DigitalFUTURES* (pp. 103-113). [https://doi.org/10.1007/978-981-33-4400-6\\_10](https://doi.org/10.1007/978-981-33-4400-6_10)
- Whitehand, J.W.R. (2001). British Urban Morphology: the Conzenian tradition, in *Urban Morphology*, 5(2),103-109.

# AN ASSESSMENT OF TOOL INTEROPERABILITY AND ITS EFFECT ON TECHNOLOGICAL UPTAKE FOR URBAN MICROCLIMATE PREDICTION WITH DEEP LEARNING MODELS

NARIDDH KHEAN<sup>1</sup>, SERJOSCHA DÜRING<sup>2</sup>, ANGELOS CHRONIS<sup>3</sup>, REINHARD KÖNIG<sup>4</sup> and MATTHIAS HANK HAEUSLER<sup>5</sup>

<sup>1,2,3,4</sup>*City Intelligence Lab, Austrian Institute of Technology, Austria.*

<sup>1,2,4</sup>*Bauhaus-Universität Weimar, Germany.*

<sup>5</sup>*Computational Design, UNSW, Australia.*

<sup>1</sup>*nariddh.khean@ait.ac.at, 0000-0003-0549-7784*

<sup>2</sup>*serjoscha.duering@ait.ac.at, 0000-0002-1003-1236*

<sup>3</sup>*angelos.chronis@ait.ac.at, 0000-0002-6961-2975*

<sup>4</sup>*reinhard.koenig@uni-weimar.de, 0000-0002-3579-8855*

<sup>5</sup>*m.haeusler@unsw.edu.au, 0000-0002-8405-0819*

**Abstract.** The benefits of deep learning (DL) models often overshadow the high costs associated with training them. Especially when the intention of the resultant model is a more climate resilient built environment, overlooking these costs are borderline hypocritical. However, the DL models that model natural phenomena—conventionally simulated through predictable mathematical modelling—don't succumb to the costly pitfalls of retraining when a model's predictions diverge from reality over time. Thus, the focus of this research will be on the application of DL models in urban microclimate simulations based on computational fluid dynamics. When applied, predicting wind factors through DL, rather than arduously simulating, can offer orders of magnitude of improved computational speed and costs. However, despite the plethora of research conducted on the training of such models, there is comparatively little work done on deploying them. This research posits: to truly use DL for climate resilience, it is not enough to simply train models, but also to deploy them in an environment conducive of rapid uptake with minimal barrier to entry. Thus, this research develops a Grasshopper plugin that offers planners and architects the benefits gained from DL. The outcomes of this research will be a tangible tool that practitioners can immediately use, toward making effectual change.

**Keywords.** Deep Learning; Technological Adoption; Fluid Dynamics; Urban Microclimate Simulation; Grasshopper; SDG 11.

## 1. Introduction

Deep learning (DL)—a sub-field of machine learning predicated on a family of algorithms known as neural networks—has been toyed with for decades. However, DL

has only gained its recently burst in popularity in 2012. Professor and computer scientist, Olga Russakowsky, notes '2012 was really the year there was a massive breakthrough in accuracy, but it was also a proof of concept for deep learning models, which had been around for decades' (Agrawal, 2018, p. 28-29). Most attribute this revitalisation to a met threshold in 'huge, flexible stores of data' combined with increased 'access to [more] powerful computers' (Browne, 2019, p. 64). Since then, where DL took a landmark leap in accuracy, it has been applied in both academia and industry, in almost all fields ranging from medical diagnosis (Hafiz and Bhat, 2020) to agriculture (Kamilaris and Prenafeta-Boldú, 2018). Even in the urban planning industry, DL has made its mark. However, despite its wide breath of application, DL is not a 'silver bullet'. While 2012 was the year that we saw deep learning's rise, 2018 was the year it became cool to criticise deep learning (Broussard).

The efficacy of deep learning algorithms is fickle. 'Neural networks can model very complex patterns [...] and, as such, are very powerful. In fact, they are so powerful that they can even model the noise in the training data' (Baesens, 2014, p. 51), an insidious phenomenon known as 'overfitting'. Furthermore, DL algorithms are 'black boxes', non-transparent and opaque modelling methods which leaves any attempt to understand why the model predicted what it did a fruitless endeavour. Finally, DL models are only as good as the data used to train it. 'We tend to think of data as the immutable truth, but we forget that data and data-collection systems are create by people' (Broussard, 2018, p. 57), thus is susceptible to error. These factors, and other mathematical and computational restrictions (Baesens, 2014, p. 58), make DL suitable for only certain types of problems, and when misapplied, can lead to horrific ramifications (Angwin et al., 2016).

This research will focus on the application of DL within urban microclimate analysis prediction for climate resilience. (Its suitability will be discussed in the literature review). These models, when applied to a computationally intensive simulation such as wind factor analysis, provides orders of magnitude improvement through prediction rather than simulation. However, discounting the costs associated with training are incredibly myopic and selective. Training DL models—a costly process involving data acquisition and engineering, consuming computation time and resources, as well as the need for continual model assessment and retraining upon conceptual drift—all contribute to a horrific carbon footprint. A study from the University of Massachusetts, Amherst, concluded that the process of training large AI models emits roughly five times the lifetime emissions of the average American car (Strubell et al., 2019). Despite the righteous proclamation for sustainability and environmental resilience through deep learning (which could arguably be seen as an oxymoron in itself), DL models need to go beyond a costly and environmentally damaging research toy and move toward a commercially viable product to be adopted by those in industry who would use it to make tangible positive environmental impact.

## 2. Literature Review

'All models are wrong, but some are useful' (Domingos, 2015, p. 149). Each modelling method has a variety of strengths and weakness. And deep learning is no exception. An understanding and careful consideration of these attributes will be the difference between success and gross misapplication. Thus, what problems within urban planning

can be most suitably addressed with deep learning?

One such set of problems focus on predicting the outcomes of a computationally intensive simulation process based on natural phenomena. Issues of data quantity and quality—'multiple data sources, subjective judgement, limited computing facilities, and size of data' (Baesens, 2014, p. 152)—are ever-present when it comes to training deep learning models. However, by creating a training dataset through simulation over a variety of real-world and generated urban forms, the engineer has complete control and can drastically minimise these issues. Furthermore, because the DL model is trained upon data generated in simulation, there is minimal concern to do with human-and equipment-error. And the fact that the data is generated means that the size of the dataset is never an issue, because if one needs a larger dataset, they could simply run more simulations; time is the only limiting factor. 'Data makes prediction better' (Agrawal, 2018, p. 174). Through these sets of problems, the issues that surround both data quality and quantity are drastically minimised. Conversely, the benefits gained from using DL methods over typical simulation approaches come down to the staggering gains in computational speed and efficiency. Because, despite the large amounts of time and computational power needed to train a DL model, once trained, the resources needed to make a prediction is orders of magnitude less than that of running a simulation. Once trained, the simulation (now prediction) process is made much more efficient.

This research focuses on the application of deep learning on urban microclimate analyses, which in itself includes a breath of studies, such as urban heat island, rainfall, and pedestrian comfort simulations. Typically, these studies are conducted through arduous simulation approaches. The promise of DL for microclimate prediction is to model the patterns between urban form and the input parameters with the simulation results; to predict, rather than simulate. A 2019 meta-analysis concludes that the then current majority of DL application in urban planning focuses on 'urban growth or urban land use analysis while there [was] less interest in environmental studies' (Grekousis, 2019, p. 253). For research contribution toward that of lesser studied, this research will further narrow down its scope by assessing the application of deep learning on wind factor analysis based on computational fluid dynamics (CFD).

There has been a plethora of research done toward the development of deep learning models for CFD prediction (Zhu et al., 2018; Kulkarni et al., 2019; Erdemir et al., 2020), however, there is a lesser subset dedicated for their application in urban planning (Ding and Lam, 2019; Kaseb et al., 2020; Mokhtar et al., 2021; Sun et al., 2021). A 2021 review provides a comprehensive overview of the published research on the application of DL in CFD analysis, for both the indoor and outdoor environment. Within their review, they identify that the 'vast majority of applications in the built environment are limited to substituting [the conventional approach] [...] to achieve faster predictions' and that their main objectives 'are to reduce computational cost while keeping the same order of accuracy' (Calzolari and Liu, 2021). However, the study also concedes that 'most often, fast predictions come at a cost of degraded accuracy'.

While the 2021 review compares research outcomes on algorithm choice and model speed and accuracy, what is not taken into account is their accessibility. 'Technology is a tool. That is true whether it's a hammer or a deep neural network' (McAfee, 2017, p. 330). Stretching that further, unless that tool is designed and offered

in a manner that is highly accessible and available with minimal barriers to entry, such a tool may lay unused, affecting no one. Yes, although it is true that conventional 'CFD simulation requires engineering knowledge and expensive computational resources, [creating] a barrier for people to use [it]' (Ding and Lam, 2019), DL models also require a highly specific set of engineering skills and knowledge to use; that of MLOps. MLOps ('machine learning operations', borrowing from the word, 'DevOps') is the task of deploying or embedding a machine learning model into a program, application, or API that allows access to said model. Without sufficient MLOps, the complexity barrier still exists, but has just shifted. Of the aforementioned research published on DL CFD prediction, none make any mention of its deployment or use (albeit such focus is arguably outside the scope of the reviewed publications).

### 3. Research Objective

There has been recent innovation toward improving or replacing conventional computational fluid dynamics simulations with deep learning models for the sheer efficiency gains offered by such methods. The common objective of these efforts is to increase the speed and decrease the computational costs of analyses, while minimising accuracy loss. However, there is significantly less research emphasis placed on their usage and access, namely through MLOps. Despite the evidence suggesting that the models improve upon the resource costs of conventional methods by literal orders of magnitude, this research suggests that to make a tangible and measurable impact with these models, they need to be positioned with a low barrier to entry so that they can be used by other researchers and, more importantly, those in industry.

The objective of this research is to contribute to the broader uptake of DL models for rapid CFD studies, through the development and implementation of an API wrapper within a visual programming environment more familiar to urban planners (Grasshopper).

### 4. Methodology

This research builds upon the work and research of the City Intelligence Lab, namely, upon the 'Intelligent Framework for Resilient Design' (InFraReD) (Chronis et al., 2020). InFraReD is a web-based platform for intelligent and resilient urban design that leverages deep learning to accelerate a variety of urban microclimate analyses. As of the time of writing, these models are accessible via two methods: the InFraReD web application as well as through a GraphQL application programming interface (API). It is the latter method by which this research leverages to create a Grasshopper plugin that takes a step toward lowering the barrier of usage.

The Grasshopper plugin developed during this research was done so following an applied, action research methodology. Widely attributed to social psychologist, Kurt Lewin, action research started as a 'methodology for intervening in and researching the major social problems of the day' (Hopkins, 2008, p. 48). According to Lewin, action research 'consisted in analysis, fact-finding, conceptualisation, planning execution, more fact-finding or evaluation; and then a repetition of this whole circle of activities' (Kemmis, 1988, p. 13). This research largely follows the original incarnation of the action research methodology for the case study.



The software development for this research was conducted within the span of a week and was done so with the intention to use the outcomes for the 2021 Digital Futures workshop.

## 5. Case Study

Contributing toward the broader uptake of DL models for urban microclimate studies, this research intends to develop a wrapper around an existing platform of DL models in a software environment more familiar to urban planners: the 3D modelling software, Rhino 6, and its parametric scripting counterpart, Grasshopper. The existing platform that provides access to DL microclimate models is the Intelligent Framework for Resilient Design (InFraReD), a web application and API made available by the City Intelligence Lab (Chronis et al., 2020). However, these two methods of use still prove some difficulty, and thus, resistance. The web app asks the user to shift their software ecosystem and the data import/export function is not yet available (forcing the user to recreate all their geometry in the app), and conversely the API still requires some understand of programming which has proven to be too high a barrier of entry. Consequently, a third, more accessible method, is the projected research outcome.

At the time of writing, InFraReD has four microclimate studies available to users: solar radiation, sunlight hours, wind speed, and wind comfort. To perform a prediction using these models, there are two required inputs: urban geometry, and the input parameters that effect the post-processing for each model (i.e., wind speed, wind direction, etc.). However, before even making a prediction, one needs to overcome the challenges of formulating an authenticated request, of synchronising data between Grasshopper and InFraReD's servers, and then visualising the resultant predictions. By far, the greatest challenge was the shift of mental model between that of a web application and a Grasshopper plugin, not helped with the time restraints of the task. Each of these challenges and their designed solutions are explained below.

There are several programming languages by which a Grasshopper plugin can be written, however the choice of Python was one made considering the expertise of the research group and the time allotted for the plugin's development.

### 5.1. AUTHENTICATION

To access InFraReD's services, one first needs to authenticate. As the DL models are resource demanding for the hardware it is currently deployed on, we are monitoring and limiting its use by keeping track of who has access.

Authentication is handled quite simply through the use of cookies. Using either the web application or the API, all a user would need to do is send a POST request to InFraReD's authentication route with their log in credentials. Upon valid credentials, a response storing their authentication cookie would be returned. Within the cookie, is a JSON web token (JWT), 'a compact, URL-safe means of representing claims to be transferred between two parties' (Jones et al., 2015). This claim can be used to statelessly verify a user on InFraReD's backend, and any subsequent request they may send, so long as the request contains that cookie. Thus, the task of leveraging this authentication technique within a Grasshopper plugin is an exercise of extracting this cookie and injecting it in all subsequent requests from the other plugin components.

First, an authentication component was developed that required the username and password of InFraReD credentials. The component would send a request to InFraReD, and upon successful authentication, the component would receive a response with the cookie containing the JWT. To preserve this cookie and reuse it in the other component's requests, the cookie was extracted and stored in a 'sticky' variable, made available in the 'scriptcontext' module, making the variable accessible in other components in the Grasshopper script.

```
...
class Authenticate(component):
    ...
    def RunScript(self, U, P):
        ...
        # Extract cookies
        for cookie in response.Cookies.GetEnumerator():
            if cookie.Name == 'InFraReD':
                auth_token = cookie.Value
        ...
        # Store stickies
        sc.sticky['InFraReDAuthToken'] = auth_token
    ...
```

Each InFraReD Grasshopper component checks for this sticky variable for the authentication cookie before it computes anything further.

```
...
def RunScript(self, ...):
    ...
    # Validate authentication
    auth_token = sc.sticky.get('InFraReDAuthToken')
    if not auth_token:
        component.AddRuntimeMessage(
            self,
            Grasshopper.Kernel.GH_RuntimeMessageLevel.Error,
            "[Message]"
        )
    ...
```

## 5.2. DATA SYNCHONISATION

InFraReD is a 'stateful' application. As opposed to its counterpart known as stateless applications, stateful applications store data and events that may have occurred with a user's interactions, which then allows future experiences to be adjusted. Common examples of stateful web applications include online banking and email. For the case of InFraReD, the stored data includes project metadata, 'snapshots' (which is InFraReD's term for design variants), urban geometries, and a variety of analysis configuration variables (including which analyses are set to run project-wide, what input parameters for each analysis, etc.). The fact that InFraReD's database stores urban geometry, brings with it a host of benefits, significantly, removing the need to always send geometry data within a request each time a prediction needs to be run. Conversely, if an external application, such as Rhino/Grasshopper, were to create geometry outside of InFraReD, this geometry would need to be mirrored, and the issue of data synchronisation becomes a potential challenge (addressed in the Evaluation section).

The components designed to synchronise Rhino/Grasshopper geometry with the InFraReD server must parse that data into a format that InFraReD can comprehend. InFraReD represents building geometry as a building footprint and a height. And so, any geometry type that may be considered a building (i.e., boundary representations, meshes, etc.) need to be simplified and reduced to a curve and floating-point number.

```
import rhinoscriptsyntax as rs
import ghpythonlib.components as ghcomp
...
def extract_footprint_and_height(building):
    mesh = rs.coercemesh(building)

    # Footprint
    crv = ghcomp.MeshShadow(
        mesh,
        rs.CreateVector([0, 0, 1]),
        rs.CreatePlane([0, 0, 0])
    )
    footprint = ghcomp.SimplifyCurve(crv).curve

    # Height
    z = [v.Z for v in ghcomp.DeconstructMesh(mesh).vertices]
    height = max(z) - min(z)
    ...
```

Note, there are some complexities removed in favour for brevity and concision (such as a recursive function that called 'ghcomp.SmoothMesh()' if the 'ghcomp.MeshShadow()' method failed to return a valid curve).

Having extracted the footprint and height, to formulate a request to the InFraReD server, the building data is then converted to a GeoJSON string and converted to a GraphQL mutation request.

### 5.3. PREDICTION VISUALISATION

The InFraReD web application was designed to be asynchronous. Despite being orders of magnitude faster than a CFD simulation, making a prediction could take up to 3 seconds, not to mention the pre- and post-processing involved (depending on hardware). At the time of writing, InFraReD is deployed on a single node (however the research team has been working on using Kubernetes to scale to a multi-node cluster). Thus, if there were to be several prediction requests sent to InFraReD in a very short span of time, these jobs would be added to a queue (handled by Apache Kafka) and be resolved sequentially.

When InFraReD makes a prediction, the resultant inference is postprocessed and stored on the server as a normalised matrix. This was an intentional design decision, as, although this offsets the responsibility of visualising the results up to the technology providing access to InFraReD (i.e. Unity in the web browser, or whichever technology one chooses if they access InFraReD's API), this maintains a high level of interoperability, and assures that no technology cannot use and manipulate the visualisation in their own fashion. Thus, the developed Grasshopper plugin must also include a method to visualise the matrix of results. Due to the time restrictions of the development period, this process was achieved with Aviary.

#### 5.4. OTHER CHALLENGES

This research paper could not hope to include all the challenges that were encountered and resolved (and not resolved) during the development of the Grasshopper plugin. The three that were discussed here were intentionally selected as they posed the most development difficulty, in conjunction with providing the most fuel for discussion in the evaluation section. For the documentation and source code, please visit [redacted].

### 6. Evaluation

Due to the brevity of development time, there are some software design decisions made to favour the rapid completion of the project over a robust solution. As economist Thomas Sowell so famously positions, 'there are no solutions, only trade-offs' (Sowell, 2007). However, if the objective of this research is for the general uptake and use of deep learning models, this work needs some further action research iterations to reach 'production-ready'. Below are what this research has identified as the outcome's major pain-points, which coincide with the team's next steps.

The first major inconvenience is the compute graph. Despite early efforts in the project, the developed Grasshopper plugin was almost a one-to-one conversion of the GraphQL API into a series of Python requests wrapped in Grasshopper components. There is a general failure to consider the difference in the mental model between that of a web application API and a Grasshopper script. As such, using the Grasshopper plugin, there is a rigorously strict method by which the plugin should be used, so not as to disrupt the compute graph. If so much as one connection triggers a component that should not run, an entire InFraReD project could be ruined (this would not affect the geometry and such on the Grasshopper side, only on InFraReD's servers). Coupled with the fact that InFraReD is stateful, and how data on Grasshopper must be mirrored with that on InFraReD's servers, if one ill-timed API request was triggered, the local and server-side projects could lose synchronisation and the disparities would continue to diverge on subsequent requests. The way to resolve this is to rethink how the deep learning models would best be used from the perspective of a Grasshopper user, rather than that of a web application. Then, collating these chunks of behaviour, recreating the plugins to include the necessary GraphQL API requests to perform such actions.

Despite the computational speed and efficiency gained over conventional simulation approaches, deep learning predictions still take some time to compute, not to mention the overhead of sending that data over the web. A single prediction made with the Grasshopper plugin takes roughly 3 seconds. However, because geometry is an input for the 'predict' component, every geometric manipulation would add at least those 3 seconds for computation. This is unnecessary for each incremental adjustment. One simple method to prevent such issues, offered in the workshop and in the developed tutorial video, is to have every InFraReD Grasshopper component disabled unless intending to send a request. If a user wants to send a request, they would re-enable the required components, and then re-disable them once complete. This is arguably a very 'un-Grasshopper-like' solution. Thus, simple buttons on the components, used to trigger the request to the API, could be the better solution.

Finally, Grasshopper and InFraReD represents building geometry differently, and in fact, in Grasshopper, this is entirely up to the user. Because of the 'one-size-fits-all'

solution of converting geometry to meshes, there is potential for loss of detail. However, this problem is drastically exacerbated when considering how InFraReD represents building geometry: footprint and height. Thus, if a user has geometry more complex than a vertical extrusion, the building would be simplified. The reason for this is primarily due to how the deep learning models were trained. There is no solution that could have been implemented within the development of the plugin, and requires the recreation of a training dataset, retraining the DL models, and rebuilding the data input pipelines on InFraReD's servers, well beyond the scope of this research. This is a vital stipulation that all users of InFraReD should be aware of.

Some smaller issues include: the method of credential input is not secure and can easily be stolen if a user saves the Grasshopper script and disseminates it (a pop-up box would be the solution here); the process of result visualisation would be more reliable if included in the plugin rather than relying on an external plugin (which would allow the implementation of domain-specific colours gradients for the different analyses), amongst others. The outcomes of this research are far from ready for wide-spread adoption—the feedback from the participants of the 2021 Digital Futures workshop indicated as much. However, this research succeeds in illuminating a side to microclimate analysis through deep learning that has not received enough attention and provides a method by which one can tangibly contribute. This research hopes to act as a steppingstone for further developments in this field.

## 7. Conclusion

While there has been a general trend toward embedding deep learning processes for improving or replacing conventional approaches in almost all areas of research since 2012, there has been substantially less emphasis placed on the use of such technology. This imbalance is especially problematic, and even hypocritical, as one considers how incredibly resource-intensive it is to train a DL model, upon the application of deep learning for climate resilience. The outcomes of this research draw attention to the necessity of developing methods to access DL-based microclimate predictions, and further develops one such method. Leveraging the existing InFraReD API, which offers rapid DL-based prediction on solar radiation, sunlight hours, and wind factor analysis, this research implemented a wrapper around the API within a software environment more familiar to urban planners. This developed Grasshopper plugin was then used in the 2021 Digital Futures conference workshops.

## References

- Agrawal, A. (2018). *Prediction Machines: The Simple Economics of Artificial Intelligence*. Boston, Massachusetts: Harvard Business Review Press. ISBN: 978-1-6336-9567-2.
- Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). *Machine Bias*. Propublica. Retrieved November 27, 2021, from <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
- Baesens, B. (2014). *Analytics in a Big Data World: The Essential Guide to Data Science and its Applications*. Hoboken, New Jersey: Wiley. ISBN: 978-1-1188-9270-1.
- Broussard, M. (2019). *Artificial Unintelligence: How Computers Misunderstand the World*. Cambridge, Massachusetts: The MIT Press. ISBN: 978-0-2620-3800-3.

- Browne, J. (2019). *Think, Make, Imagine: A Brief History of the Future*. London: Bloomsbury Publishing. ISBN: 978-1-5266-0572-6.
- Calzolari, G., & Lui, W. (2021). Deep learning to replace, improve, or aid CFD analysis in build environment applications: A review. *Building and Environment*, 206, 108315. <https://doi.org/10.1016/j.buildenv.2021.108315>.
- Chronis, A., Aichinger, A., Duering, S., Galanos, T., Fink, T., Vesely, O., & Koenig, R. (2020). InFraReD: An Intelligence Framework for Resilient Design. In *25th International Conference on Computer-Aided Architectural Design Research in Asia*. The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA).
- Ding, C. & Lam, K. P. (2019). Data-driven model for cross ventilation potential in high-density cities based on coupled CFD simulation and machine learning. *Building and Environment*, 165, 106394. <https://doi.org/10.1016/j.buildenv.2019.106394>.
- Domingos, P. (2015). *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. New York: Basic Books. ISBN: 978-0-4650-6570-7.
- Erdemir, G., Zengin, A. T., & Akinci, T. C. (2020). Short-term wind speed forecasting system using deep learning for wind turbine applications. *International Journal of Electrical and Computer Engineering*, 10(6). <https://doi.org/10.11591/ijece.v10i6.pp5779-5784>.
- Grekousis, G. (2019). Artificial neural networks and deep learning in urban geography: A systematic review and meta-analysis. *Computers, Environment and Urban Systems*, 74, 244-256. <https://doi.org/10.1016/j.compenvurbsys.2018.10.008>.
- Hafiz, A. M., & Bhat, G. M. (2020). A survey of deep learning techniques for medical diagnosis. *Information and communication technology for sustainable development*, 161-170. [https://doi.org/10.1007/978-981-13-7166-0\\_16](https://doi.org/10.1007/978-981-13-7166-0_16).
- Hopkins, D. (2008). *A teacher's guide to classroom research*. Maidenhead: Open University Press. ISBN: 978-0-3352-2175-2.
- Jones, M., Bradley, J., & Sakimura, N. (2015, May). *JSON Web Token (JWT)*. RFC 7519. Retrieved November 27, 2021, from <http://www.rfc-editor.org/rfc/rfc7519.txt>.
- Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture. *Computers and Electronics in Agriculture*, 147, 70-90. <https://doi.org/10.1016/j.compag.2018.02.016>.
- Kaseb, Z., Hafezi, M., Tahbaz, M. & Defani, S. (2020). A framework for pedestrian-level wind conditions improvement in urban areas: CFD simulation and optimization. *Building and Environment*, 184, 107191. <https://doi.org/10.1016/j.buildenv.2020.107191>.
- Kemmis, S. (1988). *The Action research planner*. Warrn Ponds, Victoria: Deakin University Press. ISBN: 978-0-7300-0521-6.
- Kulkarni, P. A., Dhoble, A. S., & Padole, P. M. (2019). Deep neural network-based wind speed forecasting and fatigue analysis of a large composite wind turbine blade. *Journal of Mechanical Engineering Science*, 233(8), 2794-2812. <https://doi.org/10.1177/0954406218797972>.
- McAfee, A. (2017). *Machine, Platform, Crowd: Harnessing our Digital Future*. New York: W.W. Norton & Company. ISBN: 978-0-3932-5429-7.
- Mokhtar, S., Beveridge, M., Cao, Y., & Drori, I. (2021). Pedestrian Wind Factor Estimation in Complex Urban Environments. In *13th Asian Conference on Machine Learning*.
- Sowell, T. (2007). *A conflict of visions: ideological origins of political struggles*. New York: Basic Books. ISBN: 978-0-4650-0205-4.
- Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and Policy Considerations for Deep Learning in NLP. In *57th Annual Meeting of the Association for Computational Linguistics, ACL 2019*, The Association for Computational Linguistics (ACL).
- Sun, C., Zhang, F., Zhao, P., Zhao, X., Huang, Y., & Lu, X. (2021). Automated Simulation Framework for Urban Wind Environments Based on Aerial Point Clouds and Deep Learning. *Applications of AI and Remote Sensing in Urban Systems*, 13(12), 2383. <https://doi.org/10.3390/rs13122383>.

# MACHINE LEARNING TOOL FOR SUSTAINABILITY EVALUATION: THE CASE OF NEIGHBOURHOODS' DESIGN

NOAM M. RAANAN<sup>1</sup>, HATZAV YOFFE<sup>2</sup>, GAL ZEEV<sup>3</sup> and  
YASHA J. GROBMAN<sup>4</sup>

<sup>1,2,3,4</sup>*Technion - Israel Institute of technology.*

<sup>1</sup>*noammraanan@gmail.com, 0000-0002-5499-6275*

<sup>2</sup>*hatzav.y@campus.technion.ac.il, 0000-0003-0262-3052*

<sup>3</sup>*gal.zeev@campus.technion.ac.il,*

<sup>4</sup>*Yasha@technion.ac.il, 0000-0003-4683-4601*

**Abstract.** This paper proposes a framework for machine learning to evaluate landscape design. In this study, we measured key performance indicators of landscape-development plans using a convolutional neural network (CNN) approach to predict the performance level of the design. The model used 3749 performance evaluations from 36 professionals, covering six sustainability criteria in 32 neighbourhoods' designs. Results show a high agreement level between experts on the performance level of the designs. The study contributes to computational sustainability by showing the potential in evaluation-automation of urban resiliency, ecological enhancement, and design for wellbeing, using expert knowledge and machine learning.

**Keywords.** Urban Design, Landscape Architecture, Computational Sustainability; Machine Learning; Convolutional Neural Network; Landscape Sustainability; SDG 9; SDG 11; SDG 13; SDG 15.

## 1. Introduction

In the Anthropocene, the urban population grows on account of natural habitats (Almond, Grooten, and Peterson 2020). Urban sprawl, densification, and development on open land contribute to the ongoing degradation of existing ecosystem services in cities, which due to climate changes, results in severe urban ailments such as increased vulnerability to hurricanes, flooding, wildfires, and heatwaves. (SCBD, 2012). Sustainable urban landscape development that enhances ecologies and strengthens communities' resiliency is essential for adapting to climate change and increasing the well-being of urban dwellers (United Nations 2021). Although Sustainable design of buildings is considered a standard practice in the architectural engineering and construction (AEC) industry, sustainable design of neighbourhoods is still considered a labour-intensive, expert dependent, and time-demanding process with low implementation rates (Pedro et al. 2019). Recent technological advancements in Building information modelling (BIM) and Geo-design (Ayman, Alwan, and McIntyre 2020; Wang, Pan, and Luo 2019) promote automation, analysis, and visualization of sustainability performance in design, making computational sustainability an emerging

field of research (Chong, Lee, and Wang 2017). Integration of automation using machine learning (ML) methods that provide access to what were once tedious tasks, helps to reduce calculation and evaluation time. It also facilitates finding new relations between elements, for tasks such as energy simulations (Hendrycks, Lee, and Mazeika 2019; Sebestyen, and Tyc 2020), and classification of architectural building types using images (Date and Allweil 2021).

In urban landscape design, ML methods were used for tasks such as classification of buildings, land use and roof types, quality evaluation of building facades, visual quality of streets, and citizens' perception of the built environment such as safety aspects, in growing numbers each year (Tebyanian 2020).

Although significant work has been undertaken in this area, there seems to be less focus on evaluating urban landscapes at the neighborhood level; specifically regarding sustainability aspects: design's social, ecological, and resiliency. (Yoffe, Plaut, and Grobman 2021). Furthermore, to our knowledge, none of the automation processes found in the existing literature uses expert-based evaluation datasets.

This study aims to advance computational sustainability by introducing a framework for sustainability evaluation using expert-based knowledge and machine learning. It also suggests a preliminary method that uses a convolutional neural network algorithm trained to evaluate the sustainability properties of plans based on expert-evaluation data collection.

This paper presents the proposed method and a demonstration experiment, covering 32 neighborhood development plans and 36 landscape sustainability expert evaluations, followed by a discussion of the experiment results.

## **2. Methodology**

The proposed framework extracts domain experts' knowledge using a survey and translates it to sustainability KPIs. A machine learning-based approach implements the acquired knowledge to automate sustainability evaluation, supporting expert feedback input for the next design iteration (figure 1). The method was divided into three steps: (1) data collection via survey, (2) data analysis and preparation, and (3) computation method validation (Figure 2). The final training dataset included 3749 performance evaluations from 36 professionals, covering six sustainability criteria of 32 neighborhood designs.



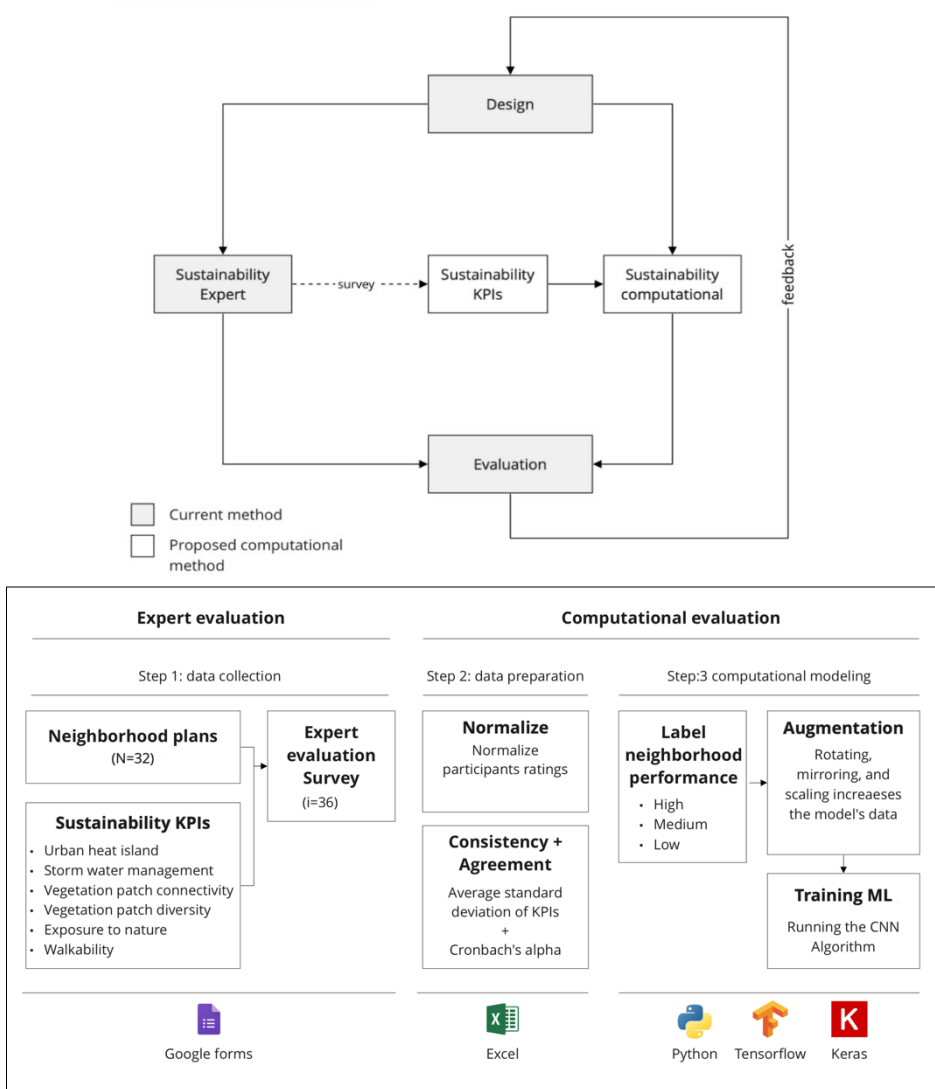


Figure 1: From expert evaluation to computational evaluation. A tree step method.

## 2.1. STEP 1: DATA COLLECTION

Expert data was collected via a survey among landscape professionals. Experts rated six sustainability criteria in 32 rendered neighborhood plans. Neighborhoods plans, randomly selected from a web image platform (see Figure 3). Word combination used in image search included: 'urban development', 'neighbourhood', 'landscape design',

'neighbourhood plan', 'landscape plan'. We choose plans similar in development level and rendering style. Also, plans had similar architectural elements, such as buildings, streets, and vegetation. Chosen plans had no text labelling, an indication of their location or origin, or any other form of information besides graphics.

Three sustainability aspects were evaluated: ecology, resilience, and society, each containing two key performance indicator (KPI) criteria. The achievement level for each criterion was described on a scale of 1 to 7 (Table 1). The choice of KPIs was based on local AEC industry needs for evaluating ecological, social, and resilience design aspects in landscape development (Yoffe, Plaut, and Grobman 2021).

Table 1: Sustainability aspects and evaluated criteria.

	<b>Sustainability criteria (KPIs)</b>	<b>Evaluation description</b>
<b>Ecology</b>	<b>Vegetation patch connectivity</b>	Does the design and dispersion of vegetation encourage fragmentation or connectivity (based on trees and streams)? (1) fragmented, (7) connected
	<b>Vegetation patch diversity</b>	Does the design create similar (1) or diverse (7) spaces and outdoor experiences?
	<b>Urban heat island (UHI)</b>	How much, to your opinion, does the landscape mitigate the urban heat island effect and encourage moderated micro-climate? (the evaluation of UHI is determent by the ratio between hard surfaces and soft surfaces, SRI of surfaces, and the dispersion of landscaping and vegetation). High mitigation (7), Low mitigation (1)
<b>Resiliency</b>	<b>Stormwater management</b>	How much, to your opinion, does the design (dispersion of open spaces) enable implementation of stormwater management methods (retention areas, blue roofs, etc.) and decrease of flooding damages? Very poor (1), Very well (7)
	<b>Exposure to nature</b>	How much, to your assumption, does the design expose the residents to the environment and the natural components in the design? Low exposure to none (1), High exposure(7)
<b>Social</b>	<b>Walkability</b>	Does the design encourage a safe and comfortable walk between the neighborhood points of interest (enables continuous walk, shade, interest, multiple walking possibilities)? Doesn't encourage walkability (1), Highly walkable and connected (7)



Figure 2: Examples of rendered plans from the survey

Each neighborhood was represented by a plan image (Figure 2), followed by the sustainability evaluation questions. The survey was distributed to the professional landscape community during the 18th annual Israeli Association of Landscape Architects (ISALA) meeting and conference.

## 2.2. STEP 2: DATA PREPARATION

Data preparation included determining the agreement level between experts and cataloguing the neighborhood plans for computational ML training. Data preparation had two goals: to determine consistency and to determine agreement. After normalization of the evaluation results, consistency was calculated using Cronbach's alpha equation for each criterion on all 32 neighborhoods. We defined consistency as coefficient  $\alpha > 0.70$  as it is an accepted benchmark by many studies in the social studies field (Cortina, 1993). Agreement among experts was found by calculating the standard deviation of each criterion.

## 2.3. STEP 3: COMPUTATIONAL MODELLING

The final step included exploring sustainability evaluation using machine prediction. This study scope examined the KPI with the highest expert agreement rating, meaning the lowest standard deviation (figure 3). We chose the 'Vegetation patch diversity' criterion, aiming to test all six KPIs in the future.

The neighborhoods were divided into three achievement levels: high, medium, and low based on the average normalized score in the specific criteria where low  $< 0.33$ , medium  $< 0.66$ , and high  $> 0.66$ . The ML model architecture is a convolutional neural network [CNN] used in small data learning experiments like Chollet (2016). The script was written using the Keras and Tensorflow libraries. The algorithm included an augmentation stage where each neighborhood image was randomly rotated, flipped, scaled, and shifted to create more training data. Overall, the model created an additional 20 training images from each image. Following augmentation, data was split into train and validation sets, creating two directories, a test directory containing 625 images and a validation directory with 155 images, a customary ratio of 80-20. Finally, each of the folders was split into three (3) subfolders based on the overall score received by the experts.

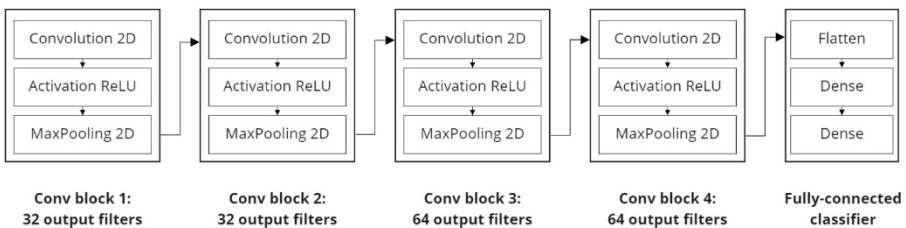


Figure 3: ML model included five convolutional blocks, divided into 16 batches and 30 epochs.

The model architecture consists of five convolutional blocks (figure 4). The data was divided into 16 batches for a training run with 30 epochs. Adjustments were made

in the last fully connected block after each run to try and achieve better accuracy until reaching the best performance with the available toolset.

The study scope tested the prediction accuracy of computer analysis for one sustainability KPI as a proof of concept.

### 3. Results

The professional survey showed promising results regarding consistency and agreement between experts, from similar backgrounds, for visual evaluation of sustainability aspects of plans (Table 2).

The survey results showed high consistency in participant evaluations (coefficient  $\alpha > 0.80$ ) in all evaluated participants. Participant evaluation numbers varied for each criterion due to the incompleteness of the survey, resulting in a range between 9-11 participants used to measure Cronbach's alpha for each KPI.

Like consistency, agreement levels – average standard deviation, were in an accepted range between 0.19-0.23, meaning experts agreed on the performance level of the neighborhood by visual analysis only (Table 2). In addition, the coefficient of variant across all KPIs was less than one, meaning a strong similarity in answers between different experts (Table 2).

The results achieved by the ML algorithm for measuring vegetation patch density were low with an overall average accuracy of 13.5% and could not be considered a successful prediction. This might be explained either by insufficient modeling data amounts, by the architecture of the CNN model in interpretation of the neighborhoods or a combination of both.

Table 2: Experts' evaluation results analysis

	Criterion	n	$\alpha$	ASD	CoV
Resiliency	Urban heat island	n = 11	0.80	0.19	0.59
	Stormwater management	n = 9	0.87	0.22	0.71
Ecology	Vegetation patch connectivity	n = 9	0.92	0.22	0.78
	Vegetation patch diversity	n = 11	0.86	0.19	0.53
Social	Exposure to nature	n = 10	0.86	0.21	0.62
	walkability	n = 9	0.89	0.23	0.57

Legend: n - number of experts used for  $\alpha$ ,  $\alpha$ -Cronbach's  $\alpha$ , ASD - average standard deviation, CoV-coefficient of variation

#### **4. Discussion and conclusions**

This study introduced a framework and a method aimed at connecting expert knowledge on sustainability with computational evaluation using ML abilities. The suggested method uses a CNN algorithm trained to evaluate the sustainability properties of plans, based on an evaluation dataset, achieved using a survey among landscape sustainability experts.

The study demonstrated positive results in the agreement level between expert evaluations of neighborhood plans. It strengthened the idea that domain experts agree on the sustainability performance level when evaluating projects. This ability, which captures the qualitative evaluations and transforms them into quantitative data, is the cornerstone for useable datasets in ML-based tools. However, the low prediction results of the applied CNN tool highlight the importance of increasing the sustainability evaluation datasets prior to the successful application of this method by practitioners.

Machine learning in urban and landscape design is an emerging and rapidly growing field (Yoffe, Plaut, and Grobman 2021). In a literature review done by Tebyanian (2020), out of 71 papers reviewed, 34 were published in 2019 and 15 in 2018, portraying this phenomenon. Not only the academy is looking towards ML methods but practitioners as well. The American Society of Landscape Architecture (ASLA) (2019) shows that more than 25% of landscape architecture firms intend to adopt AI/ML as part of their arsenal of tools, motivating researchers and reassuring that studies like this have a place in the frontlines of landscape architecture computation.

The proposed method has successfully demonstrated a workflow for utilizing domain-expert knowledge to advance the design process by disseminating sustainability evaluation abilities to non-professionals. However, the scope of the data gathered was not sufficient to achieve high computation prediction rates.

Future studies which would extend professional datasets and the measured sustainability criteria can significantly contribute to the implementation of sustainable design in ongoing and planned developments. Introducing ML automation to sustainability evaluations could help a labor-intensive become broadly accessible to the AEC industry and support global efforts in creating ecologically rich, resilient environments and more liveable cities.

#### **Acknowledgements**

We would like to thank the Israeli Association of Landscape Architects ISALA, especially to Michal Bitton, for helping disseminate the survey during the 18th ISALA Expo and Meeting. Rendered plan images courtesy of the Commons Landscape Design Studio, Lerman Architects, Lavi Natif, Engineering and Consultants Ltd., Arim Urban Development Ltd., iStern Project Management, Shikun & Binui, Moshe Zur Architects & Town Planners Ltd.

## References

- Almond, R., Grooten, M., and T Peterson, T. (2020). *Living Planet Report 2020-Bending the Curve of Biodiversity Loss*. Gland, Switzerland: World Wildlife Fund.  
<https://f.hubspotusercontent20.net/hubfs/4783129/LPR/PDFs/ENGLISH-FULL.pdf>.
- ASLA. (2019). *Design Software Survey Results - The Field*. Retrieved from  
<https://thefield.asla.org/2019/09/26/design-software-survey-results/>.
- Ayman, R., Zaid, A., and Lesley, M. (2020). BIM for Sustainable Project Delivery: Review Paper and Future Development Areas. *Architectural Science Review*, 63 (1), 15–33.  
<https://doi.org/10.1080/00038628.2019.1669525>.
- Chollet, F. (2016). *Building Powerful Image Classification Models Using Very Little Data*. Retrieved June 5, 2016, from, <https://blog.keras.io/building-powerful-image-classification-models-using-very-little-data.html>.
- Chong, H. Y., Cen Y. L., and Xiangyu W. (2017) A Mixed Review of the Adoption of Building Information Modelling (BIM) for Sustainability. *Journal of Cleaner Production*. Elsevier Ltd. <https://doi.org/10.1016/j.jclepro.2016.09.222>.
- Cortina, J. M. (1993). What Is Coefficient Alpha? An Examination of Theory and Applications. *Journal of Applied Psychology*, 78 (1), 98–104.
- Date, K., and Allweil, Y. (2021). Towards a New Image Archive for the Built Environment. *Environment and Planning B: Urban Analytics and City Science*.  
<https://doi.org/10.1177/23998083211011474>.
- Hendrycks, D., Kimin, L., and Mantas M. (2019). Using Pre-Training Can Improve Model Robustness and Uncertainty. In *36th International Conference on Machine Learning, ICML 2019*, 4815–26.
- Pedro, J., A. Reis, M. D. P., and Silva, C. (2019). A Systematic Review of the International Assessment Systems for Urban Sustainability. In *IOP Conference Series: Earth and Environmental Science*. 323. Graz, Austria: IOP Publishing Ltd.  
<https://doi.org/10.1088/1755-1315/323/1/012076>.
- Sebestyen, A., Tyc, J. (2020). Machine Learning Methods in Energy Simulations For. Edited by L Werner and D Koering. *Anthropologic: Architecture and Fabrication in the Cognitive Age - Proceedings of the 38th ECAADe Conference*, 11 (M1): 613–22.
- Secretariat of the Convention on Biological Diversity. (2012). *Cities and Biodiversity Outlook - Action and Policy*. Montreal, Quebec: Secretariat of the Convention on Biological Diversity. <http://www.cbd.int/en/subnational/partners-and-initiatives/cbo>. Website: [www.cbd.int](http://www.cbd.int).
- Tebyanian, N. (2020). Application of Machine Learning for Urban Landscape Design: A Primer for Landscape Architects. *Journal of Digital Landscape Architecture*. 2020 (5), 217–26. <https://doi.org/10.14627/537690023>.
- United Nations. (2021). *The Sustainable Development Goals Report 2021*. New York, NY: Department of Economic and Social Affairs United Nations, Population Division.  
<https://unstats.un.org/sdgs/report/2021/The-Sustainable-Development-Goals-Report-2021.pdf>.
- Wang, H., Yisha, P., and Xiaochun L. (2019). Integration of BIM and GIS in Sustainable Built Environment: A Review and Bibliometric Analysis. *Automation in Construction*. Elsevier B.V. <https://doi.org/10.1016/j.autcon.2019.03.005>.
- Yoffe, H., Plaut, P., and Grobman, Y. J. (2021). Towards Sustainability Evaluation of Urban Landscapes Using Big Data: A Case Study of Israel's Architecture, Engineering and Construction Industry. *Landscape Research*. 00 (00), 1–19.  
<https://doi.org/10.1080/01426397.2021.1970123>.





# UAV-BASED PEOPLE LOCATION TRACKING AND ANALYSIS FOR THE DATA-DRIVEN ASSESSMENT OF SOCIAL ACTIVITIES IN PUBLIC SPACES

JEROEN VAN AMEIJDE<sup>1</sup> and CARSON KA SHUT LEUNG<sup>2</sup>

<sup>1,2</sup>*The Chinese University of Hong Kong*

<sup>1</sup>*jeroen.vanameijde@cuhk.edu.hk, 0000-0002-3635-3305*

<sup>2</sup>*kashutleung@cuhk.edu.hk, 0000-0003-2936-0344*

**Abstract.** In sustainable high-density cities, public spaces play an important role in supporting social and community health and well-being. Amidst ongoing urbanisation, it is of increasing importance to study public space interaction patterns and placemaking processes that contribute to the quality of life of urban residents. This paper reports on the development of a new methodology for the computational tracking and analysis of social activities in urban spaces, using Computer Vision Object Detection (CVOD) techniques to create digitalised pedestrian trajectory data. Referring to concepts from humanistic geography and time geography, our method offers a new platform for data-driven urban place studies, detecting co-presence and social interaction in relation to urban morphology. This paper focuses on the development of Machine Learning protocols, algorithms for tracing and mapping pedestrian trajectories in a georeferenced photogrammetry model, and computational analysis of co-presence. The resulting workflow forms a foundation for future research around detecting, analysing and quantifying behavioural parameters, to evaluate the ability of public spaces to support social interaction and placemaking.

**Keywords.** Public Space Analysis; Pedestrian Location Tracking; Computer Vision Object Detection; Machine Learning; SDG 11.

## 1. Introduction

### 1.1. PUBLIC SPACE, SOCIAL ACTIVITIES AND PLACEMAKING

Amongst the growing awareness of the need to reduce the environmental impact of urbanization, the Compact City concept provides appealing benefits such as enhanced land-use and transportation efficiency, walkability, and other conveniences (Westerink et al., 2013). In high-rise cities such as Hong Kong, the design of public spaces is a crucially important challenge, as they facilitate important everyday activities and socialising, and serve as an extension of people's limited domestic space (Gou et al., 2018). The notion of 'placemaking' emphasizes the role of "context, local conditions, and place-specific culture and experience" in shaping neighbourhood communities and

well-being (Williams, 2014, p. 75). Scholars have asserted that place is produced by people's relationships with their environment, geographical behaviour, and the social structures and identities of space and place (Tuan, 1976). 'Secondary interactions' (Jacobs, 1961) can stimulate casual neighbouring, which can lead to social connections, integration, attachment to place (Talen, 1999). Placemaking can contribute to improved social capital, the collective ability to secure resources and opportunities (Friedmann, 2010). Hence, it is of increasing importance to study how sociable public spaces shape interaction patterns and placemaking processes that contribute to the quality of life of urban residents. An increasing number of data-driven studies in recent years has confirmed the association between co-presence and social behaviour (De Stefani and Mondada, 2018; Zakariya et al., 2014). Integrating overlapping insights from the fields of time geography and urban place studies, our research aims to detect behavioural parameters that can be observed, analysed and quantified to evaluate the ability of public spaces to support social interaction and placemaking.

## 1.2. SPATIAL-TEMPORAL PUBLIC SPACE ANALYSIS

Previous studies around documenting activities in public spaces have mixed qualitative and quantitative approaches, following methodologies for the ethnographic study of space outlined by Whyte (1980), Gehl and Svarre (2013) and Low (2016; 2019). These methods combine several observational techniques, including population counts, movement maps and behavioural maps (Low, 2019), and are typically limited in scale across time and geographical space. With the rapid development of information and communication technologies, analysis of increasingly accessible geo-referenced urban data offers a quantitative human-centred approach to capturing the social dynamics of public space. Recent research focuses on the detection of user behaviours and routes based on different types of activity information and Big Data analysis (Biljecki & Ito, 2021; Chen et al., 2016). One category of research, which involves computer vision-based systems to detect and track pedestrians, has been fast developing due to the number of possible applications such as crowd size measurement, transport security, pedestrian traffic management, etc. (Sidla et al., 2006). The research presented in this paper applies a methodology related to this field, but expands it towards the social processes of placemaking in public spaces.

## 2. Methodology

The research presented in this paper employed Unmanned Aerial Vehicle (UAV) mounted cameras, to create a digital capture of an urban space morphology through photogrammetry modelling, and UAV and building-mounted cameras in connection to a newly created pedestrian Real-Time Locating System (RTLs). To obtain RTLs data, multiple cameras were used to obtain fixed position videos of the site from surrounding vantage points. The videos were then analysed with Computer Vision Object Detection (CVOD) techniques to create digitalised pedestrian trajectory data. This data was integrated within the digital models of the urban space using perspective transform algorithms. Quantitative relationships between social activities and public space layout design were then extracted, spatialised and analysed using a customised information processing workflow.

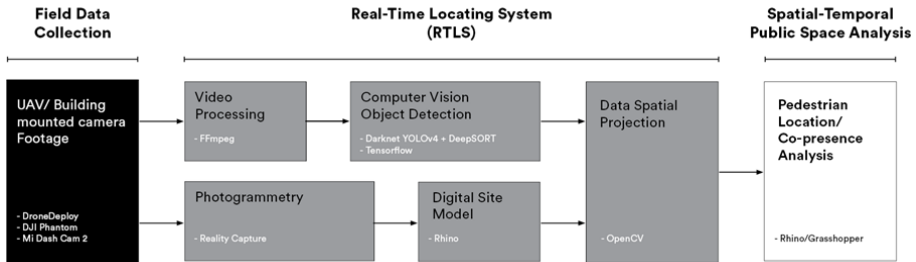


Figure 1. Workflow for a RTLS based Spatial Temporal Analysis

Figure 2 shows the site selected for experimentation, the University Mall at the central campus of The Chinese University of Hong Kong. This area is part of the main access route across the central campus, which connects the main library, administrative buildings, various faculties, and adjacent bus stops.

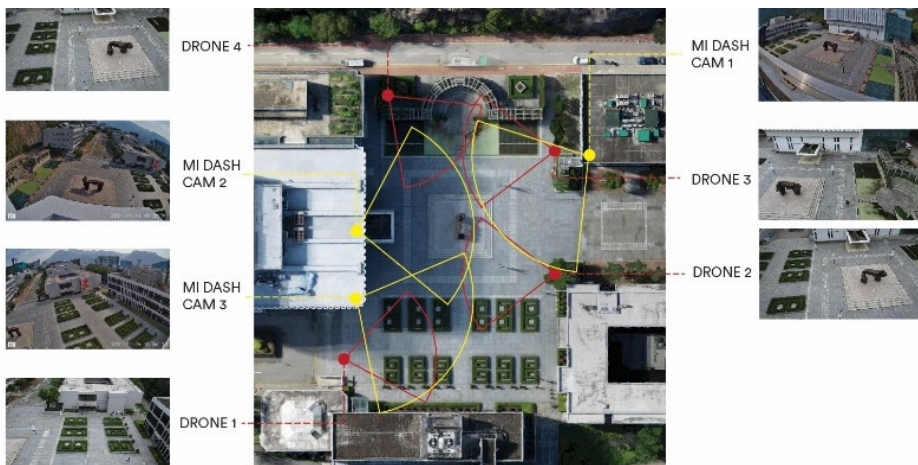


Figure 2. Multiple Camera System Deployment locations

### 2.1. PEDESTRIAN DATA COLLECTION

Two methods were developed and tested for obtaining raw bird's-eye view footage of the university mall. The first method used four drones, and the second method used three fixed location cameras (Figure 2). For both methods, the camera angle was calibrated to maximise ground coverage and to exclude the perspective vanishing point. In the first experiment, four DJI Phantom pro 4 v2.0 drones were used to obtain time-synced footages from multiple aerial angles for object tracking purposes to capture pedestrian locations during the evening rush hour. Four drones were dispatched to different holding positions above the square at a flight level of 18m above ground, in which the take-off and landing procedures were designed sequentially to avoid collisions. Within the available battery flight time, the total recording time was 8.5 minutes at 30 frames per second. The per minute record size was 700mb. For the

second methodology tests, three Mi Dash Cam 2 car camcorders mounted on tripods and connected to portable battery packs were placed on the roofs of two buildings adjacent to the site. The camcorders were chosen for its ability to loop-record. This type of recording saves one video file per minute, which eases file management and transfer. The Mi Dash Cam 2 has a F1.8, 140° field of view lens which records in a 2K resolution (2560 x 1600). A minute of footage equates to 120mb, therefore with a 128gb microSD card, the camcorder can record up to 17 hours of continuous footage at 30 frames per second. Despite the drones' higher resolution image quality and flexibility in controlling the viewing angle comparing to the building-mounted camcorders, the recording time was severely limited by the battery span. The video stability was also affected by the wind condition and required corrections by the drone operator, compared to the camcorders which remained stable on a tripod and required no human intervention during operation. However, the drone footage yielded a better object tracking result with existing pre-trained object tracking weights, while the footage from the camcorder required a transfer learning approach, training custom object tracking weights for improved accuracy with a one-stage object detection model.

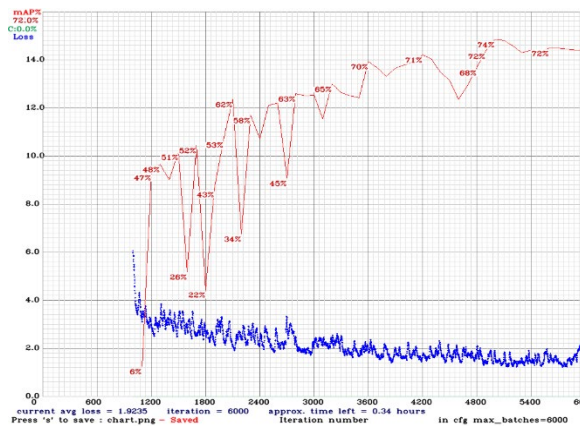


Figure 3. Training Loss (red line) and mean average precision chart (blue line) chart

## 2.2. CUSTOM DETECTION WEIGHTS

The type of transfer learning used was parameter transfer under inductive transfer learning, where pretrained parameters between the source domain and target domain model were shared (Pan and Yang, 2009). This training process adds a new detection class to the YOLO (You-Only-Look-Once) v4 pre-trained weights, using selected frames from the collected dataset. This methodology is more time-efficient compared to training a new set of weight from scratch. The Darknet convolution framework, exclusively designed for object detection, was used to process a set of 333 image frames. The images were extracted from the camcorder videos and selectively included different time and angles to maximise the diversity in the learning process. During the detection process, bounding boxes of the target object are drawn on the image using the open-source annotation tool OpenLabeling. This process generates a separate text file, containing the coordinates and its relative detection class number. To train the

model, 80% of the data set was used for the training process while 20% was used for evaluation. The return training accuracy rate reached its peak at 76% mean average precision (mAP) after 13 hours of computation time on a Nvidia Geforce RTX 2060 GPU (Figure 3). The dataset used was designed to recognize pedestrians from a bird's-eye viewing angle with a new custom class "human". This avoids the new custom detection class from merging with the pre trained detection class "people" to avoid negative transfer, as the specificities of the dataset it was trained on are unknown.

### 2.3. OBJECT TRACKING

The object tracking process was implemented with YOLOv4, DeepSort, and TensorFlow. YOLOv4 is a one-stage object detection model algorithm that uses convolutional neural networks to perform object detections. Compared to two stage object detection models like R-CNN which have a better detection accuracy but higher inference speed, YOLOv4 was preferred for its lower computational power requirements. YOLOv4 detects the pedestrian positions and draws the relevant prediction boxes per video frame. The detection output of YOLOv4 was then fed into DeepSORT (Simple Online and Realtime Tracking with a Deep Association Metric), to create a continuous object tracker ID (Wojke et al., 2017). It considers information within the bounding box parameters of the detection results of the current and previous frames, to make predictions about the current frame. DeepSORT enable the continuous tracking of the pedestrian even if the detection is lost in certain frames due to a) missed detection or b) the subject passing through covered areas. From the first frame of a successful detection, a unique track ID is assigned to each bounding box which represents the activated detection class. On the top of the box, a confidence value higher is also displayed. Detections which have a confidence value lower than the pre-set threshold will not be displayed. The Hungarian algorithm is used to assign the detections in a new frame to existing tracks. The deep learning process were implemented in TensorFlow, Google's deep-learning software. Comparing to Darknet used in weights training, TensorFlow provides a Python API and is compatible with the Python language, to which a perspective transformation process can be added to compute the real-world location of the tracking point.

### 2.4. TRACKING ACCURACY EVALUATION

To compare the tracking accuracy of the pre-trained YOLOv4 and the self-trained detection weights, both weights were tested on footage acquired by the two different camera types. A one-minute excerpt from both cameras covering the same area of the university mall was used for this purpose. Both excerpts were lowered to 1 frame per second at their original resolution, to allow for faster processing time and recording location datapoints at one snapshot per second. A total detection count was established by employing manual counting, combining all the detected pedestrians in each frame. For both excerpts, 360 counts were recorded which took 39-41 minutes to complete.

Afterwards, the YOLOv4 and self-trained detection weights were tested on both the excerpts under the Tensorflow Convolutional Neural Network (CNN) with the confidence score set to 50% and above. This filters out low probability detections. The YOLOv4 detection on the excerpt filmed by the DJI drone returned a total of 271

correct detections and 17 incorrect detections, while the self-trained weights returned 248 correct detections and 63 incorrect detections. On the footage excerpt from the Mi Dash Cam 2, the YOLOv4 returned 89 correct detections and 130 incorrect detections while the self-trained weight returned 249 correct detections and 404 incorrect detections. These detection processes took 40-57.5 seconds to complete (Table 1). Amongst the tested options, the YOLOv4 weight and the DJI drone combination performed the best with 75.3% of accuracy and fewer incorrect detections while YOLOv4 and the Mi Dash Cam 2 has the lowest detection accuracy. These differences revealed how the image resolution and colour balance affect the detection accuracy. The Mi Dash Cam 2 has a lower image resolution and quality compared to the DJI drone. Despite filming the same location, an area of plants was detected as ‘potted plants’ with YOLOv4 and the self-trained weights detected it as ‘human’ (Figure 4).

Table 1. Tracking accuracy data comparison

<b>SETUP FOR DETECTION ACCURACY COMPARISON</b>					
Detection Weights	YOLOv4	Self-	YOLOv	Self-	
Camera	DJI Phantom pro 4		Mi Dash Cam 2		
Mean Average precision (mAP) (%)	84	76	84	76	
Duration (min)		1		1	
Frames per Second (fps)	1	1	1	1	
Total Frame Counts	60		60		
Resolution	3840 x 2160		2560 x 1600		

<b>MANUAL PEOPLE COUNTING</b>				
Total Detection Count	360		360	
Processed Time (min)	41		39	

<b>COMPUTER VISION OBJECT DETECTION (CVOD)</b>				
Detection Class	‘Person’	‘Human’	‘Per-’	‘Human’
Confidence Score Threshold (%)	50	50	50	50
Total Detection Count	271	248	89	249
Detection Accuracy (%)	75.3	68.9	24.7	69.2
Incorrect Detection	17	63	130	404
Processed Time (min)	42.3	40	55.4	57.5

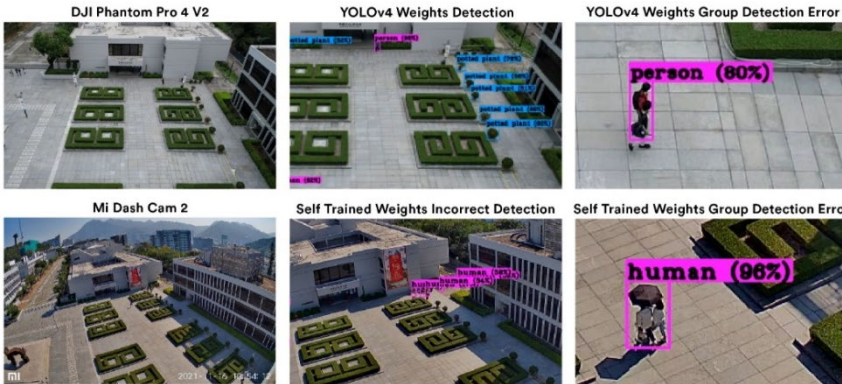


Figure 4. Tracking bounding box and detection confidence value

The low filming quality of the Mi Dash Cam 2 also directly impacted the performance of YOLOv4 as the COCO dataset that is utilized for the machine learning training process includes mostly close-up imageries of the human figure. However, the self-trained weights achieved a 69.2% detection accuracy on the Mi Dash Cam 2 footages, which could be further improved by enlarging the image data set that is being trained under the detection class ‘human’. Both detection weights underperformed in the detection of people that walk in groups. Inclusion of extra detection classes would improve the quality of the detection data. At the current stage of development, the YOLOv4 and DJI drone footage combination was chosen for their accuracy.

## 2.5. DATA SPATIAL PROJECTION

The object tracking process returns a set of U,V coordinates per tracking ID that represents the pedestrian tracked in pixel space. The set of coordinates are specified as the centroid of the bounding box generated by YOLOv4. To turn the coordinates from pixel space to real world coordinates, the OpenCV function *getPerspectiveTransform* is used to perform the calculation of a matrix operation to project a set of points from one 2D plane to the map of the site generated in the photogrammetry process (Figure 5). The process requires a set of non-collinear coordinate points that represents the corners of a quadrilateral to be defined in a custom Python script. To obtain the coordinate points required for the transform process, four identical points between the video and the map were identified as pixel locations. The accuracy of these two sets of coordinate points is crucial to the perspective transform process. In selecting references, the points should be as far apart as possible.

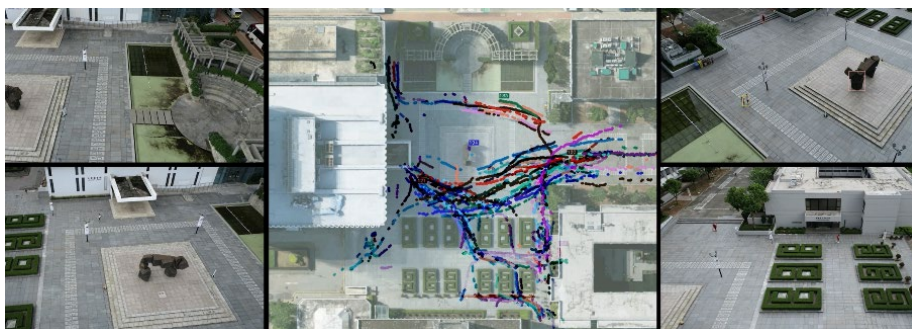


Figure 5. Tracked trajectory of 1730-1738, 28/07/2021 (510 frames, 4575 detection points)

The pixel-to-map perspective transformation performs only a linear translation of coordinates, it does not consider if a background pixel falls accurately on the 2D surface. Therefore, it is important to undistort the videos produced by the camcorder, which has 140 degrees field of view wide angle lens and produces a radial distortion that causes straight lines to appear curved. Using OpenCV’s camera calibration function *cv.findChessboardCorners*, the image points of a checkerboard grid were located in ten reference photos taken with the camcorder. A distortion coefficient was then generated to undistort the whole data set. The result of the calculation outputs the tracker ID, timestamp, and map pixel coordinates data as a list in CSV format, which could then be used for analysis in the Rhino/Grasshopper environment.

## 2.6. CO-PRESENCE ANALYSIS

A subsequent analytical process was developed to interpret the pedestrian location data, evaluating “the in-between space that facilitates co-presence and regulates interpersonal relationships” (Madanipour, 2003, p. 206). In the analysis of our case study spaces, we evaluate various distances between people based on proxemic interactions theory (Hall, 1966), which describes how people “perceive, interpret, structure, and (often unconsciously) use the micro-space around them, and how this affects their interaction and communication with other nearby people” (Marquardt & Greenberg, 2015, p. 33). In our workflow, we employed a customised algorithm to analyse the closeness of individuals using the discrete proxemic zones defined by Hall: intimate (0 - 0.5 m), personal (0.5 – 1 m), social (1 – 4 m) and public (> 4 m) (Hall, 1966). The tool analyses the distance to all other people within the public space, and groups and counts people who are within the thresholds of social and personal space. Figure 6 illustrates this analysis, using a single time-frame snapshot observation, mapped on the 3D model of the site.

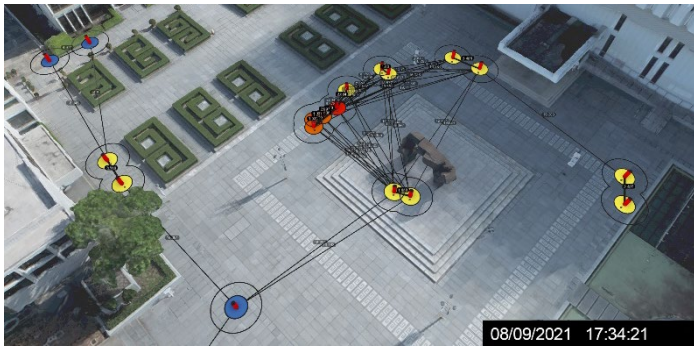


Figure 6: Analysis of people locations and co-presence, based on a snapshot observation



Figure 7: Space-centric analysis of people locations and closeness, based on the combined pedestrian location data of 8.5 minutes of drone scanning

In a final step of data translation, the human-centric analysis of people densities was translated into a space-centric analysis, defining the statistical occurrence of user presence and co-presence as a feature of the various locations within the case study space. The mapping process follows a basic logic of defining a spatial grid of cells,



defined by their size; in this test, the grid spacing was set to two meters. The number of people location markers is counted within each cell, and the social proximity value is also recorded. For this analysis, multiple datasets relating to various time intervals can be combined, to produce insights into the general statistical patterns of space occupancy as they occur over longer periods of time. Figure 7 illustrates a data mapping of the entire time period of recording through the four static drone locations described in section 2.1 and Figure 2, compiled into one analytical visualisation projected onto the photogrammetry model.

### **3. Conclusions and Future Directions**

This research has developed a methodology that employs UAV and building-mounted cameras, to construct a pedestrian Real-Time Locating System (RTLS), using Computer Vision Object Detection (CVOD) techniques to create digitalised pedestrian trajectory data. The methodology testing has demonstrated how multiple drone or building-mounted cameras can be used to capture different angles and segments of a public space, allowing for the method to be deployed in complex urban areas consisting of various spaces. The integration of multiple overlapping camera positions into a single analytical framework allows the methodology to be scalable across larger geographic areas, overcoming the limitations of traditional observational techniques for ethnographic study of space. Referring to concepts from humanistic geography and time geography, the methodology paves the way towards a system for automated detection of behavioural patterns related to social interaction and community forming, understanding specific behaviours such as avoiding, gathering, interacting, collaborating and dwelling in relation to specific environmental characteristics, public space morphology and configuration. Correlation analysis between behaviours and environmental attributes can produce insights in urban design principles that are conducive to socialising and placemaking, which would facilitate design guidelines for public spaces that support the social and mental health of individuals and communities.

As we will continue developing the methodology through case study applications, systematic analysis can reveal more detailed insights into which facilities are used more often, when and for how long, and how people move or interact around certain spaces. Public space design guidelines identified by urbanists such William Whyte and Jan Gehl will be able to be verified by data-driven and context-specific research, analysing interpersonal relationships, cultural norms and behaviour in relation to environment and context. The data-driven analysis of social dynamics can address important urban issues relating to safety and community when considering urban night-time settings, vulnerable people, and specific neighbourhood typologies such as old urban districts or public housing. It is our aim to continue developing research methodologies and digital visualisations of the social use of public spaces, to engage academics, policy makers and urban designers in conversations about improvements to existing, and the creation of new public spaces as part of the development of sustainable future cities.

### **Acknowledgements**

This project was supported by a Direct Grant for Research from the Social Science Panel, The Chinese University of Hong Kong, grant number 4052224.

## References

- Biljecki, F., & Ito, K. (2021). Street view imagery in urban analytics and GIS: A review. *Landscape and Urban Planning*, 215, 104217.
- Chen, C., Ma, J., Susilo, Y., Liu, Y., & Wang, M. (2016). The promises of big data and small data for travel behavior (aka human mobility) analysis. *Transportation research part C: emerging technologies*, 68, 285-299.
- De Stefani, E., & Mondada, L. (2018). Encounters in public space: How acquainted versus unacquainted persons establish social and spatial arrangements. *Research on Language and Social Interaction*, 51(3), 248-270.
- Friedmann, J. (2010). Place and place-making in cities: A global perspective. *Planning Theory & Practice*, 11(2), 149-165.
- Gehl, J., & Svarre, B. (2013). *How to study public life*. Washington, DC: Island press.
- Gou, Z., Xie, X., Lu, Y., & Khoshbakht, M. (2018). Quality of Life (QoL) Survey in Hong Kong: Understanding the Importance of Housing Environment and Needs of Residents from Different Housing Sectors. *International Journal of Environmental Research and Public Health*, 15(2), 219.
- Hall, E.T., (1966). *The Hidden Dimension*. Doubleday, Garden City, NY.
- Jacobs, J. (1961). *The Death and Life of Great American Cities*. New York: Vintage Books.
- Low, S. (2016). *Spatializing Culture: The Ethnography of Space and Place*. London, England; New York, New York: Routledge.
- Low, S., Simpson, T. and Scheld, S. (2019). *Toolkit for the Ethnographic Study of Space (TESS)*, Public Space Research Group, Center for Human Environments, The Graduate Center, City University of New York.
- Madanipour, A. (2003). Social exclusion and space. In R. LeGates, and F. Stout, *The city reader* (Third ed., pp. 181-189). London: Routledge.
- Marquardt, N., & Greenberg, S. (2015). Proxemic interactions: From theory to practice. *Synthesis Lectures on Human-Centered Informatics*, 8(1), 1-199.
- Mehta, V. (2014). Evaluating public space. *Journal of Urban design*, 19(1), 53-88.
- Pan, S. J., & Yang, Q. (2009). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10), 1345-1359.
- Schaumann, D., Kalay, Y.E., Hong, S. W., & Simeone, D. (2015). Simulating human behavior in not-yet built environments by means of event-based narratives. In *Proceedings of the symposium on simulation for architecture & urban design* (pp. 5-12).
- Sidla, O., Lypetsky, Y., Brandle, A., & Seer, S. (2006). Pedestrian Detection and Tracking for Counting Applications in Crowded Situations. *2006 IEEE International Conference on Video and Signal Based Surveillance*, 70.
- Talen, E. (1999). Sense of community and neighbourhood form: An assessment of the social doctrine of new urbanism. *Urban studies*, 36(8), 1361-1379.
- Tuan, Y. (1976). Humanistic Geography. *Annals of the Association of American Geographers*, 66(2), 266-276.
- Westerink, J., Haase, D., Bauer, A., Ravetz, J., Jarrige, F., & Aalbers, C. B. E. M. (2013). Dealing with Sustainability Trade-Offs of the Compact City in Peri-Urban Planning Across European City Regions. *European Planning Studies*, 21(4), 473-497. <https://doi.org/10.1080/09654313.2012.722927>
- Whyte, W. (1980). *The social life of small urban spaces*. Washington, D.C.: Conservation Foundation.
- Williams, D. R. (2014). Making sense of 'place': Reflections on pluralism and positionality in place research. *Landscape and Urban Planning*, 131, 74-82.
- Zakariya, K., Harun, N. Z., & Mansor, M. (2014). Spatial characteristics of urban square and sociability: A review of the City Square, Melbourne. *Procedia-Social and Behavioral Sciences*, 153, 678-688.

# URBAN SCALE 3 DIMENSIONAL CFD APPROXIMATION BASED ON DEEP LEARNING

*A quick air flow prediction for volume study in architecture early design stage*

YAHAN XIAO<sup>1</sup>, AKITO HOTTA<sup>2</sup>, TAKAAKI FUJI<sup>3</sup>, NAOTO KIKUZATO<sup>4</sup> and KENSUKE HOTTA<sup>5</sup>

<sup>1,2,5</sup>*Pollc.*

<sup>3,4</sup>*Mitsubishi Jisho Sekkei Inc.*

<sup>1</sup>*1023856216@qq.com, 0000-0002-5520-1163*

<sup>2</sup>*hottaakito@p-o.co.jp, 0000-0001-7000-243X*

<sup>3</sup>*takaaki@ty-fuji.info, 0000-0002-9182-2071*

<sup>4</sup>*nao.kikuzato@gmail.com, 0000-0001-5627-0905*

<sup>5</sup>*hotakensuke@p-o.co.jp, 0000-0003-3507-570X*

**Abstract.** The CFD generated by an object and its surroundings is critical during architectural design. The most common method of CFD calculation is to discretize the spatial region into small cells to form a three-dimensional grid or grid point and then apply a suitable algorithm to solve the equation iteratively until the steady state, which usually takes a significant amount of time before it converges to the exact solution of the problem. Deep learning is a subset of a Machine Learning algorithm that uses multiple layers of neural networks to perform in processing data and computations on a large amount of data. This paper presents a deep learning model CNN architecture to provide a quick and approximated 3-dimensional solution for the CFD. Our network speeds up 45 times compared to the standard CFD solver. Moreover, our network is able to predict a CFD in which the wind inlet and outlet appear at the same surface of a wind tunnel.

**Keywords.** Urban Microclimate; Machine Learning; 3D Unet; Residual Block; 3 Dimensional CFD Prediction; SDG 11.

## 1. Introduction

In architectural design, the interaction between a building and its surroundings is critical to a designer's decision-making. With the development of a variety of analytical tools to assist architectural design, architects can easily complete the analysis of various indicators such as light, wind, and illuminance without mastering professional knowledge. However, when there are multiple design options, a time-consuming simulation would be unable to be applied to each one of them. Of these analyses, airflow analysis also known as computational fluid dynamics is the most expensive one. Even so, many decisions need to be based on it, such as the shape design of super

high-rise buildings without bringing extra wind pressure to its surrounding buildings and creating a safe living environment for pedestrian and citizens. Therefore, a quick and accurate CFD simulation tool is needed. Since the middle of 2010, cloud environment simulation services have been launched one after another, and the speed of access to simulation results as well as the accuracy of the results have been attracting attention. There is a similar trend in Asia-Oceania, and the methods introduced in this paper will contribute to the development of the competitiveness of the services as well as the tools themselves.

Computational fluid dynamics (CFD) is a branch of fluid mechanics that uses numerical analysis and data structures to analyse and solve problems that involve fluid flows. The most common method of CFD calculation is to discretize the spatial region into small cells to form a three-dimensional grid or grid point and then apply a suitable algorithm to solve the equation iteratively until the steady state (Brunton et al., 2021). Since the number of iterations of the calculation cannot be specified, the computation often takes a considerable amount of time to converge to an exact solution to the problem.

Deep learning is an algorithm in machine learning based on characterization learning of data, which allows a system to automatically discover the representations needed for feature detection or classification from raw data. It uses mathematics to explain the relationship between data and applies the logic to new data to predict. The successful application of deep learning in computer vision, speech recognition, natural language processing, and biological information also proves its efficiency and potentiality to other applications. For example, a stable CFD pattern shows periodicity, which means the result at the same time in different periods should be the same. Then the results can be regarded as only related to the situation of objects, such as quantity, position and shape. Hence, we apply deep learning to map geometry and velocity fields efficiently and enable quick CFD prediction.

In this paper, the factors which impact CFD like seasons, temperate, humidity, sun radiation, site, and building materials are not taken into consideration.

## **2. Literature Review**

In recent years, the booming of convolutional neural networks research on computer vision and the applications like self-driving, body temperature detection has proven that it is successful in learning geometry recognition and representation. Meanwhile, many researchers (Ahmed et al., 2020) are trying to utilize neural networks to predict CFD since it is very dependent on the geometry's condition.

Researchers (Guo et al., 2016) approximated solutions for accelerating the results with a low error rate to show how neural networks can predict the velocity magnitude field in steady-state flows. While other works (Lui and Wolf, 2019), like aerofoil design optimization and acceleration of sparse linear system solutions, also provide further contributions of neural networks' application on fluid prediction in 3 dimensions. However, in these studies, although the prediction of the CFD on 2-dimensional can get a high degree of reduction in a short time, the default condition is that the shape of the other section of the objects is consistent with the shape of the observation plane or remains unchanged. Meanwhile, most studies on predicting CFD on 3-dimensional

only considered a single object, and the wind direction is fixed. These do not correspond to reality.

In researcher MD Ribeiro's work, the U-Net network has been proven to be capable of CFD prediction in 2D version. Therefore, we propose an approximation model based on the general U-Net structure to predict the 3-dimensional velocity field of multiple objects for conceptual study at the early stage. This paper visualizes the simulation results digitally during the conceptual design process, which enable more reasonable decision-making and then achieve sustainable cities.

### 3. Methodology

The approach is combined by 5 steps: 1) procedural geometry preparation; 2) CFD results collection; 3) dataset preparation; 4) ANN (Artificial Neural Network) training; 5) testing and evaluation.

#### 3.1. TRAINING DATASET PREPARATION

A city has many buildings that are interrelated and have a complex appearance. In order to approximate the actual situation, CFD under the mutual influence of multiple different objects will be the primary data source of this research. 20 random objects in different heights, shapes and scales inside a bounding box of 192\*192\*96 meters are generated by grasshopper to simulate city buildings. Buildings' height varies from 10 meters to 80 meters. Moreover, the grid size for CFD calculation is 1 meter. (Figure 1)

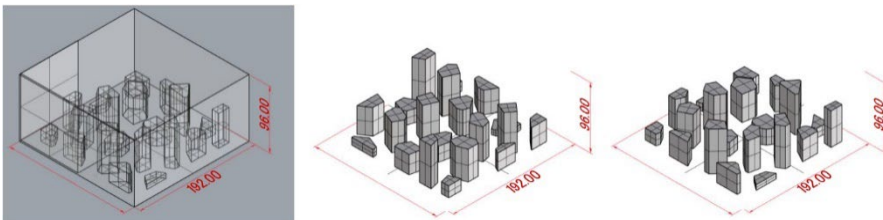


Figure 1. Wind tunnel scale (left) and other two cases

The representation of geometry needs to reflect the relative relationship between the building's shape and wind tunnel and the scale of the wind tunnel. So, we choose SDF (Signed Distance Function) (Guo et al., 2016) (Figure2 left), the shortest distance from a space point to the nearest surrounding objects, to reflect the relative position of the object and the wind tunnel. And we take reference from DeepCFD (Ahmed et al., 2020) and create WSD (non-slip wall side distance) value, distance from the wall of the wind tunnel to the central axis of the ground in the direction of the wind (Figure2 right). To make the neural network more accurately identify whether a point is inside of the objects, the SDF value of an inside point is set to -50. Both SDF and WSD were calculated by grasshopper, and for each case, SDF values remain the same regardless of the wind direction.

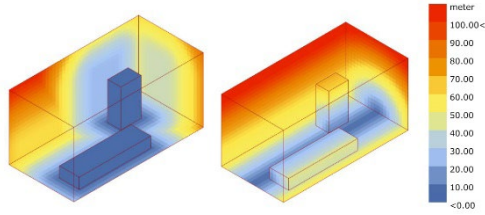


Figure 2. (left) Signed Distance Function in 3D domain; (right) non-slip Wall Side Distance in 3D domain

### 3.2. CFD SIMULATION

CFD result is solved by AKL FlowDesigner, a computational fluid dynamics simulation software that supports models from Rhinoceros and is able to export the velocity result on grid points. The generated objects are regarded as obstacles without any material, so the heat absorption and emission would not impact the wind speed and direction. Each case is simulated with the wind in 4 directions: south, east, north and west. The actual wind is different from the wind tunnel set in the experiment with only one inlet and outlet. It is affected by objects' location, shape and quantity that multiple inlets and outlets could exist on one plane. Considering this, a CFD result with a smaller scale of  $96*96*96$  meters which is taken from the overall results of each case will contain the feature. Twenty-five small boxes data are taken from one CFD simulation, where 75% of each small box is overlapped with other boxes (Figure3 left), which allows the network to thoroughly learn the relationship between the different boundary wind directions and the generated wind fields.

In this paper, we use LGR (latent geometry representation) (Ribeiro, 2020) to represent the flow region (0 for the obstacle inside condition, 1 for the flow region, 2 for the upper and surrounding no-slip wall condition, 3 for velocity inlet condition, 4 for velocity outlet condition, 5 for obstacle' surface condition and 6 for ground condition). For the LGR value of a point on a small-scale box's surface, the inlet and outlet condition are determined by the product of its wind direction and the average vector of the surfaces (Figure3 right).

Three hundred small boxes with a size of  $96*96*96$  meters' geometry data and its corresponding CFD data are prepared as the dataset of our neural network.

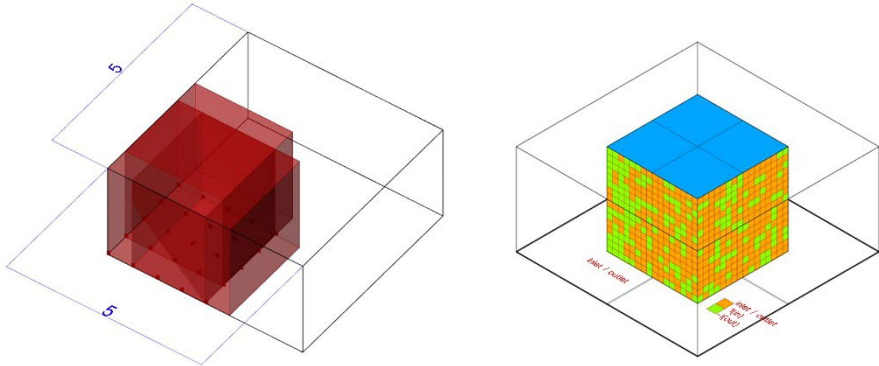


Figure 3. (Left) small-scale training dataset preparation; (Right) different LGR value: 1. inlet: orange; 2. outlet: green; 3. non-slip wall: blue; 4. Ground: brown

### 3.3. NETWORK INTRODUCTION

An autoencoder (Figure 4), a type of artificial neural network, learns a representation of a set of data and has two main parts: an encoder that maps the input into the code and a decoder that maps the code to a reconstruction of the input. The autoencoder network can discover a coordinate and a decoder to obtain a latent space containing minimal information to describe the whole system. So, it suits the great dimensional problem such as a fluid flow which would require billions of degrees of freedom to represent but only a few of them are essential and have low dimensional patterns.

The U-net network is similar to the autoencoder model, consisting of an encoder and decoder. The difference is that U-net uses a concatenation layer to connect the encoder and decoder, which recovers information loss in the down-sampling process and gives localization information to let the output back to the exact size of the input. Other researchers have proven that the U-net network is powerful enough for CFD and has achieved high-speed and accurate prediction on one single 2D object and one single 3D object (Guo and Li, 2016; Ribeiro et al., 2020). Since in reality, the CFD impacted by the interaction among buildings is more complicated to learn. To predict a result close to the real situation, a more generalized and more robust neural network is necessary. Then the structure of our network is based on the existing structure of 3D U-net, making every path from the network's input to its output is composed of a residual sub-network. The residual block can increase the depth of the network without losing accuracy.

We implemented our network based on Lee K's work, in which a 3D U-net network with residual block has been successfully applied to brain image segmentation tasks (Liu et al., 2020). The basic building block of our network consists of a single convolutional layer followed by the residual block, and a summation join replaces the concatenation joining. Our network-building block's order is group normalization layer, convolutional, and Relu activation function. Upsampling is implemented with stridden transposed convolution and downsampling with max pooling (Figure 7).

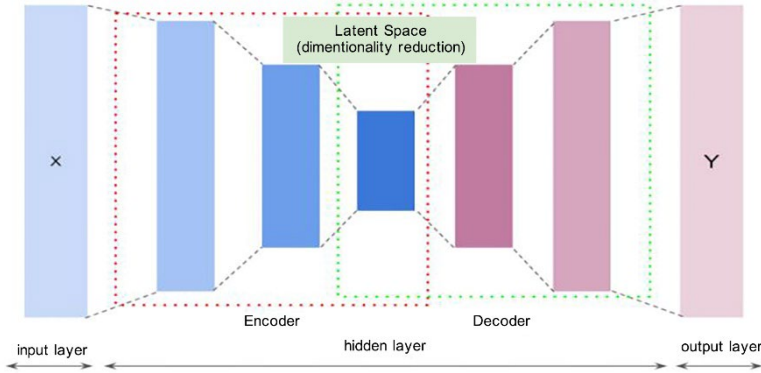


Figure 4. Autoencoder structure

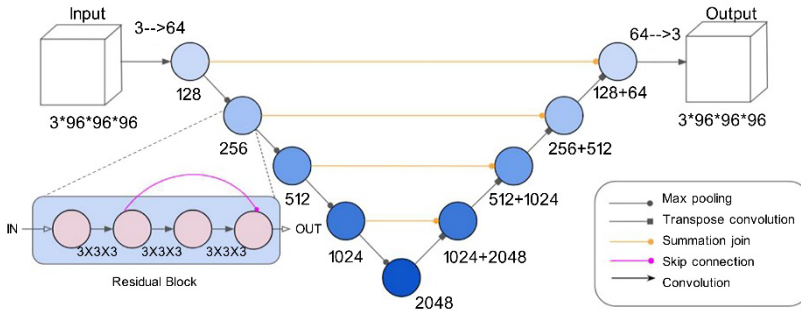


Figure 5. network structure

### 3.4. TRAINING

We implemented the network using Pytorch. Regarding the learning process, the Adam optimizer was employed with batch size, and weight decay was set to 0.05. Three hundred datasets are divided into training sets and validation sets at 80% to 20%. At the same time, the loss function is the mean squared error (squared L2 norm) which measures the output velocity in  $U_x$ ,  $U_y$  and  $U_z$  directions. The initial learning rate is set to  $1e-5$  and is gradually decreased by a learning rate scheduler which allows dynamic learning rate reduction based on the validation measurements. The training time for one epoch is approximately 6500 seconds.

## 4. Results and Performance Analysis

To evaluate the accuracy of the network, 50 cases were tested, and the velocity field prediction is compared with the generated actual value from FlowDesigner. We



visualize one prediction and true value in Figure 6 to understand the model's accuracy. The performance of our network is evaluated from three cross sections of the velocity fields on the X-Y, X-Z and Y-Z planes and shows the magnitude of the difference in wind's three directions in Figure 7.

Test wind tunnel size is the same as the training sets. However, considering the CFD solver only uses CPU for calculating, we compared our network's time-consuming prediction and simulation by Flow Designer on CPU. The average simulation time for the current wind tunnel is 3 minutes, and the prediction time is 4 seconds, which means our network has speed up 45 times.

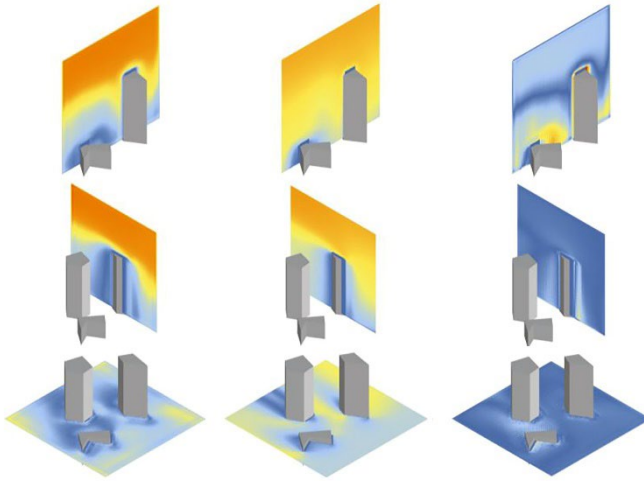


Figure 6. (left side images) CFD prediction(0-10m/s); (middle images) true value(0-10m/s); (right side images) differential(0-5m/s)

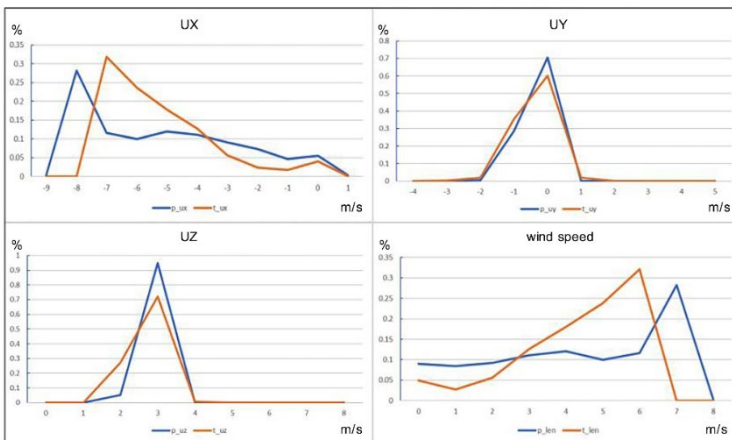


Figure 7. (horizontal axis: speed (m/s); vertical axis: distribution (%))The difference between the predicted value and the true value of the three wind directions and wind speed (blue: prediction; orange: true value)

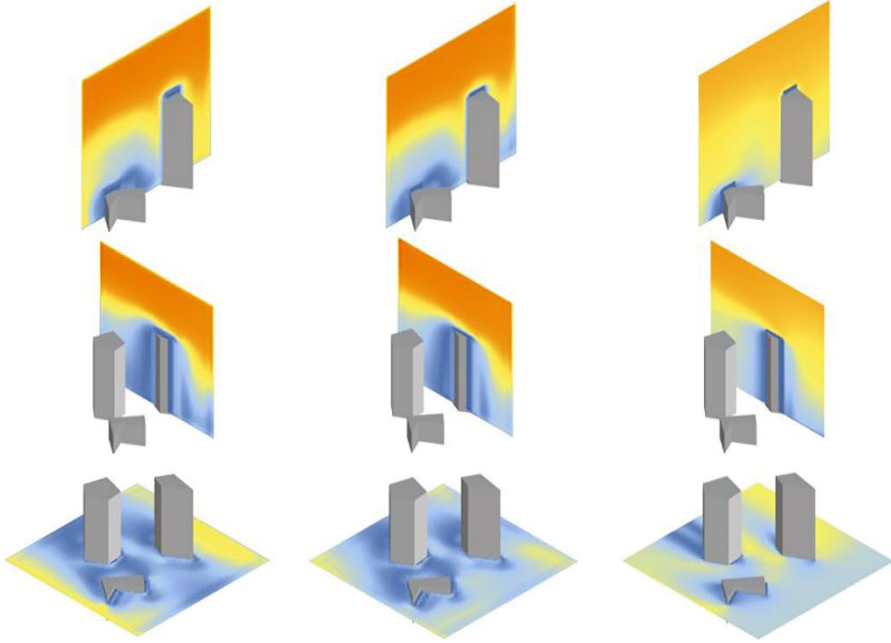


Figure 8. (left side images) prediction with actual boundary condition(0-10m/s); (middle images) previous prediction(0-10m/s); (right side images) true value

Figure 6 predicts that the boundary conditions are set as non-slip walls, two surfaces and their positions are parallel to the wind direction; the ground; the air inlet, one surface which is perpendicular to the wind direction and let the wind get in; the air outlet, one surface is positioned perpendicular to the wind direction and as the main exit; and the surface parallels to the ground. The prediction in Figure 6 shows that the network can learn low dimensional latent space like the variation law of wind speed along the height, the changing law of wind speed, and wind direction after passing through an object from the current data. However, the velocity field's feature has not been well-learned yet, for example, the changed wind direction and speed at objects' surfaces.

There is more than one wind condition on the surface of the wind tunnel. In other words, there are both air inlets and outlets on the same surface. Therefore, we predict the same objects with the acceptable boundary conditions extracted from FlowDesigner when preparing the actual value. As a result, the overall difference between the new prediction (the left image in Figure 8) and the actual value is reduced by 10% compared to the previous prediction, which shows that our network can provide more accurate results under the proper boundary conditions.

Usually, for different designs, the surrounding objects that impact the CFD simulation of the target area need to be taken into calculation every time. The more design options there are, the greater the total CFD calculation times. After obtaining

the conditions of the boundary which has the negligible impact to the target area, our current network can be applied to the partial design of large buildings or the volume study of small buildings.

## 5. Future Work and Conclusion

In this paper, we proposed an efficient and quick way for approximating a city-scale CFD in 3 dimensions which enables architects to get the velocity field of their design during the volume study. In addition, this article is the starting point for our future development of the grasshopper plugin, which provides instant feedback on CFD.

The basic equation describing the characteristics of fluid motion is the Navier-Stokes equation, a nonlinear differential equation, which expresses the relationship between fluid motion and the force acting on the fluid. It contains the fluid velocity, pressure, density, viscosity, temperature, and other variables closely related to the space position and time. Moreover, the computational difficulty of 3-dimensional is exponential times that of 2-dimensional, which means in order to predict a 3-dimensional CFD accurately, the network needs to be deep enough. We first tried with four encoders without applying residual blocks during the network construction. The prediction shows that objects' edges can be recognized, whereas the repeating pattern of swirling vortices behind the object is hard to be well predicted. After we added residual blocks and increased the depth of our network to six, the previous problem was solved. Nevertheless, compared with the ground truth, the wind direction and wind speed changed by the pressure field, which is impacted by objects' number, relative position and boundary situation before the wind reaches an object, has not been well learned yet. What is more, our current predictable scale is  $96*96*96$ , which is not enough for an accurate urban scale CFD prediction.

For our future work, we intend to enlarge our prediction scale and expand the dataset in terms of the more complex shape of objects and flow conditions, so that the prediction can be applied to the actual work. We have proved that 3D U-net with residual blocks is capable of 3-dimensional CFD prediction issues, but the number of the network's trainable parameters will increase sharply as the network deepens. Therefore, we intend to reconstruct the current network by combining the autoencoder and regression model. After the low-dimensional latent space is found by the autoencoder, the regression model will be built inside of it (Brunton et al., 2021).

## Acknowledgements

This work was funded by Mitsubishi Jisho Sekkei Inc.

## References

- Ahmed, S., Dengel, A., Rehman, A., & Riberio, M. D. (2020, April 19). *DeepCFD: Efficient Steady-State Laminar Flow Approximation with Deep Convolutional Neural Networks*. arXiv Physics. Retrieved September 7, 2021, from <https://arxiv.org/abs/2004.08826>.
- An, W., Liu, X., Lyu, H., & Wu, H. (2021). A generative deep learning framework for airfoil flow field prediction with sparse data. *Chinese Journal of Aeronautics*, 35(1), 470-484. <https://www.sciencedirect.com/science/article/pii/S1000936121000728>.

- Bhat, B., Huval, B., Manning, C. D., Ng, A. Y., & Socher, R. (2012). Convolutional-Recursive Deep Learning for 3D Object Classification. In *Advances in Neural Information Processing Systems, 25 (NIPS 2012)*. The Conference on Neural Information Processing Systems.
- Brunton, S. L., Callahan, J. L., & Loiseau, J. C. (2021). *On the role of nonlinear correlations in reduced-order modeling*. arXiv Physics. Retrieved September 7, 2021, from <https://arxiv.org/abs/2106.02409>.
- Chen, S. F., Hsiung, P. A., & Utomo, D. (2017). Landslide Prediction with Model Switching. *Applied Sciences*, 9(9). <https://www.mdpi.com/2076-3417/9/9/1839>.
- Davila, C. C., Mokhtar, S., & Sojika, A. (2020). Conditional Generative Adversarial Networks for Pedestrian Wind Flow Approximation. *The 11th annual Symposium on Simulation for Architecture and Urban Design, SimAUD2020* (pp. 469-476). The Symposium on Simulation for Architecture and Urban Design (SimAUD).
- Ding, C., & Lam, K. P. (2019). Data-driven model for cross ventilation potential in high-density cities based on coupled CFD simulation and machine learning. *Building and Environment*, 165, Article 106394. <https://www.sciencedirect.com/science/article/abs/pii/S0360132319306043>.
- Guo, X., & Li, W. (2016). Convolutional Neural Networks for Steady Flow Approximation. In *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 481-490). The annual ACM SIGKDD conference.
- He, Y., Liu, X. H., Mei, Y., Schnabel, M. A., Zhang, H. L., Zhao, F. Y., & Zheng, W. (2021). Hybrid framework for rapid evaluation of wind environment around buildings through parametric design, CFD simulation, image processing and machine learning. *Sustainable Cities and Society*, 73, Article 103092. <https://www.sciencedirect.com/science/article/abs/pii/S2210670721003759>.
- Huang, Y., Lu, X., Sun, C., Zhang, F., Zhao, P., & Zhao, X. (2021). Automated Simulation Framework for Urban Wind Environments Based on Aerial Point Clouds and Deep Learning. *Remote Sensing*, 13(12), 2383#. <https://www.mdpi.com/2072-4292/13/12/2383#>.
- Jain, V., Lee, K., Li, P., Seung, H. S., & Zung, J. (2017). *Superhuman Accuracy on the SNEMI3D Connectomics Challenge*. arXiv Computer Science. Retrieved September 7, 2021, from <https://arxiv.org/abs/1706.00120>.
- Liu, Ping & Dou, Qi & Wang, Qiong & Heng, Pheng-Ann. (2020). An Encoder-Decoder Neural Network With 3D Squeeze-and-Excitation and Deep Supervision for Brain Tumor Segmentation. *IEEE Access*, 8, 34029-34037. <https://ieeexplore.ieee.org/document/8998244>.
- Lui, H. F. S., & Wolf, W. R. (2019). *Construction of reduced-order models for fluid flows using deep feedforward neural networks*. arXiv Physics. Retrieved September 7, 2021, from <https://www.cambridge.org/core/journals/journal-of-fluid-mechanics/article/abs/construction-of-reducedorder-models-for-fluid-flows-using-deep-feedforward-neural-networks/ECEC52E32AEBBEA049CF26D6C79EE394>.

# SYNTHETIC MACHINE LEARNING FOR REAL-TIME ARCHITECTURAL DAYLIGHTING PREDICTION

RUTVIK DESHPANDE<sup>1</sup>, MACIEJ NISZTUK<sup>2</sup>, CESAR CHENG<sup>3</sup>,  
RAMANATHAN SUBRAMANIAN<sup>4</sup>, TEJAS CHAVAN<sup>5</sup>, CAMIEL  
WEIJENBERG<sup>6</sup> and SAYJEL VIJAY PATEL<sup>7</sup>

<sup>1,2,3,4,5,6,7</sup>*Digital Blue Foam.*

<sup>1,2,3,4,5,6,7</sup>{rutvik|maciej|cesar|ram|tejas|camiel|sayjel}@digital-bluefoam.com

**Abstract.** “Synthetic Machine Learning” offers a revolutionary leap in real-time environmental analysis for conceptual architectural design. By integrating automatic synthetic data generation, artificial neural network (ANN) training and online deployment, Synthetic Machine Learning offers two main advantages over conventional simulation; First, it reduces the analysis time for a reference simulation from minutes to seconds; Second, it is possible to deploy ANN as a web service in an online design environment, which therein increases accessibility, significantly reducing simulation costs and setup time. The application of Synthetic Machine Learning to perform Daylight Autonomy (DA) and Spatial Daylight Autonomy (sDA) studies to maximise building daylighting for a given use, window to wall ratio, and floorplan arrangement is showcased through a preliminary demonstration work. Comparatively the use of algorithmically generated synthetic data versus real-world data is becoming ubiquitous in other disciplines, the advantages of this approach to the building design process are further discussed.

**Keywords.** Daylight Autonomy, machine learning, building energy performance, synthetic data-sets; SDG 7; SDG 11.

## 1. Introduction

Synthetic Machine Learning is part of a revolution in environmental analysis for architectural design. For the first time, no-code Machine Learning (ML) platforms, such as Google Cloud, allow anyone to use their data to make fast, automatic design predictions. With the easing of barriers to ML model creation, non-experts will be able to access, create, share, and apply a vast set of bespoke ML tools and services.

Presently, simulations are the go-to solution for building performance analysis. Despite broad acceptance, there are several technical obstacles that designers must overcome when performing a simulation; without a domain-level understanding of the simulation mathematical models, the results are vulnerable to misinterpretation.

Furthermore, simulations require tremendous computational resources and can be extremely time-intensive. This conflicts with the early-stage design process, where designers perform several rounds of design iteration and refinement.

In recent years, research interest in using ML is rapidly gaining momentum. It is being explored as an alternative to simulation and as a method for building performance analysis (Fedorova et al., 2021, Seyedzadeh et al., 2018). This is because ML models rely on observations of known outcomes to make fast, automatic predictions of building performance metrics, rather than applying the expert-crafted, time-intensive mathematical models of simulations.

Despite these benefits, ML algorithms require access to large data sets to learn and predict outcomes. Unlike digital businesses like finance, which generate billions of gigabytes of data every day, the architecture, engineering, and construction (AEC) industries face a unique problem: a majority of building-related datasets are confined in the real world. As a result, new approaches for acquiring building-related data sets are intended to better utilize ML's potential. Synthetic Machine Learning aims to fill this gap by incorporating synthetic data (Rubin, 1993), which is algorithmically generated data via computer simulation, into the ML model's inception.

The conceptual framework and technological implementation of a novel Synthetic Machine Learning system are showcased in the paper. The proposed system offers two major advantages over conventional simulations by coupling automatic data-set generation with neural network training: First, it reduces analysis time for a benchmark simulation from several minutes to a few seconds; Second, it can be deployed as a web service within an online design environment, increasing accessibility by significantly reducing simulation cost and setup time. A preliminary concept of developing a web service to perform Daylight Autonomy (DA) and Spatial Daylight Autonomy (sDA) analysis (Reinhart et al., 2006) within an online design platform — Digital Blue Foam (DBF) — is presented to demonstrate its capabilities and features. DBF has the capabilities to run sustainability analysis based on CFD, generative design and urban insights. Existing daylighting tools are cumbersome and difficult to set up. The platform has been built with the intent of the integrated design, reduced computation time and collaboration in mind. The advantage of this approach is informed design recommendations with minimum effort and flexibility offered to the user in generating context-specific massing options.

## **2. Related Work**

The use of synthetic data and ML in architectural design is becoming more prominent. The background and recent work on the application of such methodologies in the assessment of building daylighting are presented in this section. It also encompasses a comprehensive research summary on the application of ML and daylighting study to assess building energy performance.

### **2.1. DAYLIGHT AUTONOMY**

The DBF platform can perform traditional analysis (shadow and sun hours studies) to understand the daylighting performance of the building exterior and its urban context. To better understand the building internal performance, the DA and sDA analysis were used. DA is a standard metric to describe the quality of building daylighting. DA is the percentage of time a given space remains above the threshold illuminance (300 lux for general office occupancy) throughout the year (Reinhart et al., 2001). sDA is a metric to support DA. sDA is the percentage of the floor area of the given space, which receives sufficient daylight annually. Sufficient daylight is generally a minimum of 300 lux for at least 50% of the annual occupied hours (Heschong et al., 2012). The DA and sDA analysis is directly linked to the size of the windows. The “window-to-wall ratio” (WWR) directly determines the amount of sunlight entering the building's interior. The connection of DA and sDA with the size of the windows allows for the practical use of the analysis into building's energy performance assessment as illustrated in Figure 1. At any time during a given day, the amount of daylight entering a room can be tailored by suggesting the optimal size, location, and relative positioning of windows. For example, placing smaller windows in hot-dry climates would be preferable for sufficient daylight and to avoid glare and solar heat gain. The impact of those factors combined, allows economically justified decisions from the start of the design stage.



Figure 1. Daylight Autonomy and Spatial Daylight Autonomy as metrics to quantify daylighting

## 2.2. DAYLIGHT SIMULATIONS

Presently building designers work with expert consultants who use advanced simulation software to calculate DA and sDA through computational simulation with the required parameters. This approach uses a local climate file such as an EnergyPlus weather file (EPW file), running hour-by-hour calculation of annual illuminance (amount of light received on the surface) in space with lighting simulation software. The typical workflow uses the existing libraries available in tools, like Rhinoceros-grasshopper's Ladybug plugin for Sketchup De Luminae (De Luminae, 2021) extension, to simulate the DA values. Both DA and sDA are quite common and used widely, but sDA has become popular due to its inclusion in LEED v4 and the WELL Building Standard. This approach can be time-consuming, laborious, and might require expertise because the user needs to design the building models, add materials, set targets, and correct radiance parameters. When the setup is ready, the simulation can be performed, which might take several minutes based on the input parameters and the complexity of geometry (Solemma LLC., 2021). The Synthetic Machine Learning system allows the calculation of DA and sDA to be performed in real-time.

## 2.3. SYNTHETIC DATA GENERATION

The physical nature of the construction industry limits access to data. This requires different methods of obtaining information. As a result, an example of novel methodology was devised for synthetic data generation pipeline that can generate an arbitrary amount of 3D data, as well as the relevant 2D and 3D annotations (Fedorova et al., 2021). Machine learning algorithms (MLAs) have been widely employed to predict various building performance parameters since the early 2000s. A few studies that deployed MLA for design appraisal in the field of architecture through performance parameters such as visual comfort (Chatzikonstantinou & Sariyildiz, 2016; Nasrollahi & Shokri, 2016; Navada et al., 2016), internal daylighting (Kazanasmaz et al., 2009; Lorenz & Jabi, 2017; Yoon et al., 2016), artificial lighting (Bellocchio et al., 2011; Şahin et al., 2015), and building energy performance (Chakraborty & Elzarka, 2019; Singh et al., 2019).

## 2.4. RESEARCH QUESTION

This paper presents a conceptual framework and an implementation of Synthetic Machine Learning. As a result, this paper explores several questions: What advantages does Synthetic Machine Learning offer over conventional environmental simulations? What are the trade-offs? What are possible applications of Synthetic Machine Learning within the building design process? What method allows for the creation and consolidation of useful synthetic data sets for analysis? How does Synthetic Machine learning unlock analysis to a broader audience of designers?

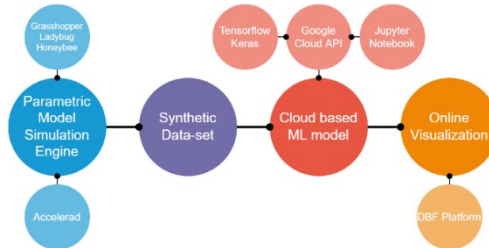


Figure 2. An overview of Synthetic Machine Learning system and summary of applied technologies

## 3. Methodology

### 3.1. OVERVIEW

Figure 2 provides an overview of the Synthetic Machine Learning system and summarizes the applied technologies. The system consists of four parts: [1] a parametric model; generating synthetic data based on environmental simulation; [2] a ML model, trained on the generated synthetic data, to predict building DA and sDA; [3] a web service containing a pre-trained ML model; and [4] a web-based user interface that allows the interaction with the ML mechanics.

### 3.2. TECHNICAL IMPLEMENTATION



This section describes the technical implementation of the Synthetic Machine Learning system for DA and sDA prediction. Figure 3 illustrates the data flow within the system.

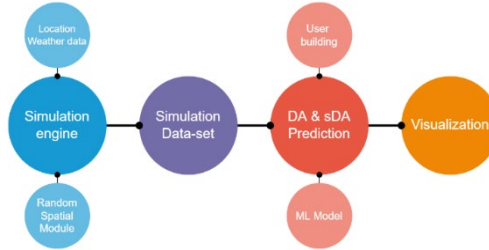


Figure 3. The framework of Synthetic Machine Learning system for DA and sDA prediction

### 3.2.1. Automatic Data Generation

To generate a synthetic dataset for ML model training, the first step is to set up a parametric model. Developed in Rhino 6, Grasshopper3D, with the LadyBug tools plugin, the parametric model automates the generation of DA and sDA data for different locations and design configurations. The method creates the simulation data based on a spatial module representing the simplified building. The system creates data that randomly cover the space of possible solutions, including the boundary conditions. The ML model can then map between different extreme inputs and create the predicted solutions. The spatial module used for simulation was defined by the following parameters: width, length, height, WWR ratio for cardinal directions (North, East, South & West), module rotation, and location (based on EPW file).

Complex typologies such as L or H shape are represented as a set of spatial modules. The analysis final value is then computed as an average value of sub-values of all spatial modules. With this representation, the system is also able to analyse individual units of the building as shown in Figure 5. This building volume description covers most of the design spectrum options at a scale of the early design stage process. Because of that, currently, there is no option to upload custom 3D models. As this interpretation is cuboid-based, currently the organic typologies are not supported.

One of the main challenges of the data generation process is optimizing computation time. This is done in several ways; First, the system implements an efficient pipeline that allows data to be created in randomly distributed batches and processed across several computers. Second, Accelerad, (Jones, 2019; Jones & Reinhart, 2017), a daylighting simulation package that makes use of GPU-compute architecture to accelerate simulation time. Finally, the system simplifies building designs to a cuboid unit-based representation.

### 3.2.2. Model Training

Based on research by (Lee et al., 2019), three different ML models have been tested: Linear Regression, Random Forest Regression, and Artificial Neural Networks (ANN) (Walczak & Cerpa, 2003). In terms of accuracy and performance with outliers and extremities, ANN outperformed the other two models, and similar findings have been reported in the literature (Ekici et al., 2021; Lorenz et al., 2019; Lorenz & Jabi, 2017).

When the simulation stage was finished, the Pandas library (McKinney, 2010; Reback et al., 2020) was used to import, clean, manipulate the data, and structure the ML model dataset. As mentioned, the ANN was applied as the core learning mechanism. Usage of Keras (Chollet, 2015), the deep learning API within the TensorFlow framework (Abadi et al., 2015) facilitated the ML model setup. The Keras API is a simplified interface between the user and core TensorFlow logics, minimizing steps that had to be performed to create functional ML models.

### 3.2.3. ML Web Service

After the training process, the ML model was ready to use. The model was then deployed on Google Cloud, a cloud computing service offered by Google. The platform offers existing infrastructure and interfaces created for these types of use cases. This significantly accelerates the implementation process of ML models. Using the Google Cloud API, the ML model was directly connected to the DBF platform which served as the user interface. Through the connection, the model evaluates the 3D building models generated by the DBF platform logic.

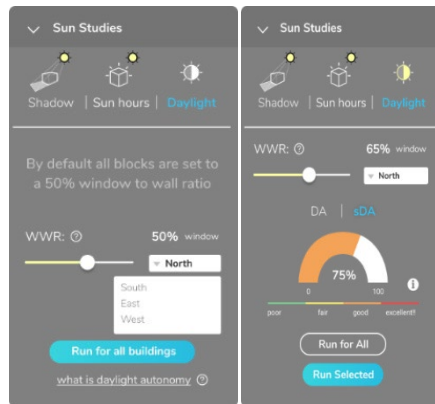


Figure 4. DBF tool GUI for varying WWR on various facades to perform DA and sDA analysis

### 3.2.4. Digital Blue Foam

The user can perform analysis on the buildings generated by the DBF platform. The prediction is presented as a clear 3D visualization and contains information on building energy efficiency. The colour-coded report presents the estimated DA and sDA values, for selected WWR ratios for each facade of the 3D building. In addition, the WWR values are represented graphically by window sizes on the corresponding building parts. Figure 4 illustrates how DBF visualizes the results of model prediction to the user. Figure 5 demonstrates the user interface which allows a designer to access the web service. Figure 6 shows the use of DBF with the DA score to evaluate 3 design options. The DBF tool was used to evaluate various forms and orientations that maximize daylight, reduce solar heat gain, and accommodate the high-angle sun in the tropical city of Singapore. Different design options were tested and the ML Daylight analysis was run to compare DA & sDA metrics for the solutions.



Figure 5. DBF building DA prediction system is based on ML and dramatically reduces daylight analysis time compared to traditional simulation-based approaches. The building score is given on the UI. The building parts performance is represented as coloured units visualized on the 3D model

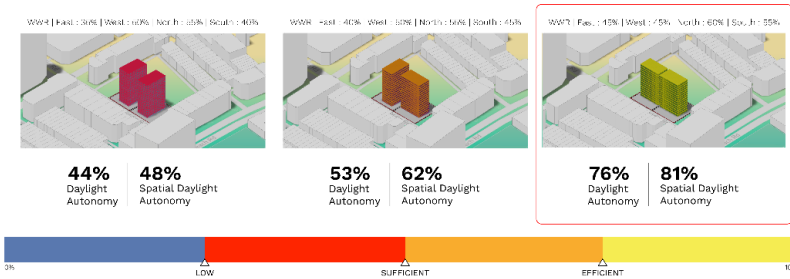


Figure 6. Illustration of an initial example study to demonstrate the application of the Synthetic Machine Learning system to a hypothetical design scenario. The case study objective was to design an office space with two 15 stories buildings located in the Downtown Singapore region. To find the optimal building orientation, size, and WWR facade ratios, the ML-Daylighting feature was used. The DA improved from 44% to 76% and the sDA improved from 48% to 81%. Selecting the last option would result in less energy consumption from artificial lighting during the daytime

## 4. Results

The section presents the effectiveness of the ML system in measuring the building solar energy performance. The accuracy of ML model prediction, as well as the execution time, were measured relative to simulation. The section also discusses how the results of model prediction are being visualized to the user.

### 4.1. ML PREDICTION EFFECTIVENESS

The system was created through two iterations, using different datasets and ML models. To formulate the problem correctly, our initial spatial module had only four features: width, length, height, and WWR ratio for the southern elevation. We performed the simulation for the city of Mumbai. As the number of data points, approximately 3500, is relatively low, we used Linear Regression as a learning mechanism. With this setup, we compared predicted results with actual, simulated ones using correlation metric, which was as good as 96.7%. Performance measurement is visualised in Figure 7.

In the second iteration, we added more parameters (WWR for each elevation, rotation of the spatial module) which assured more varied data. We used the Neural Network Regression ML algorithm as a learning method. The data was trained with 6 to 10 Neural Network layers with 100.000+ total trainable parameters. We achieved a training loss of Mean Absolute Error (MAE) of 0.009 and Mean Squared Error (MSE) of 0.0002. The performance kept on improving with more diverse data. We ensured higher variability of the dataset by the greater granularity of randomly generated inputs.

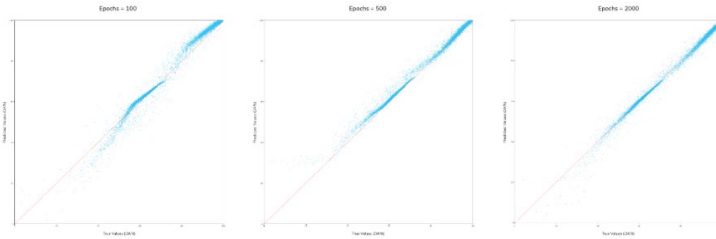


Figure 7. ML DA model's prediction accuracy

#### 4.2. COMPUTATION TIME AND SCALE-ABILITY

In the first iteration of the system, the computation time for a model was more than a month. Optimizing the computational pipeline, simplifying building representation, and leveraging Accelerad's GPU capabilities were shown to drastically improve computation time. The current system can automate synthetic data generation and model training for a single city within approximately 1 day, using a typical workstation computer: equivalent to around 3500 data points. This number ensures the high accuracy of our neural network model. To further improve the computation time, the next iteration of the system will utilize distributed computing using rendering farms.

### 5. Discussion

Our work aimed to develop a system that allows real-time assessment of the building's daylighting performance without the need for time-consuming simulations. We focused primarily on the universal application of the system and the ease of its expansion. The system uses a compact set of simulation data that can predict the DA and sDA correctly even with a small size. Our research has shown that the data-set size has the most significant impact on the accuracy of ML system predictions. It is possible that when simulating other parameters, the precision of the system will not be constant. If the accuracy is not satisfactory, we will increase the number of simulations for a given parameter. While the system answers to the general assumptions of the conceptual building design stage (operating on basic shapes), the current limitation can be the simplified building representation with the cuboid units. Further work is needed to correctly calculate daylight and generate data for irregular and organic typologies.

### 6. Conclusion

This paper demonstrates a Synthetic Machine Learning system to provide a real-time, quantitative prediction of daylighting for architectural floor plan designs. The research presented here makes the following contributions: (1) implementation of novel

automatic synthetic data generation pipeline, using a typical render farm configuration and architectural software; (2) conceptual framework for deploying ML models as web services online to assess the performance of architectural proposals; (3) comparison of Synthetic Machine Learning analysis versus benchmark simulation, showing significantly faster DA analysis from several minutes to seconds (Solemma LLC., 2021) with an accuracy of 96.7% predicted results; (4) an initial demonstration of the system to address daylight analysis of an architectural proposal in DBF. The novelty of automatic, synthetic data generation of DA pipeline relies on: (1) the aggregation of site-specific datasets created for different cities and implemented directly onto a generative design tool; (2) automatic selection of the appropriate dataset based on the chosen location; (3) the integration of this data with the DBF tenancy editor.

### 6.1. FUTURE WORK

This paper describes the application of Synthetic Machine Learning to create a real-time and scalable system for performing DA and sDA assessments. Future development of the project will expand the analysis to predict the maximum acceptable value of WWR for each orientation, and suggest how this influences other parameters like heat gain and cooling load. Additional evaluators of solar energy performance within the system will also enrich the analysis, including: information on primary solar data such as direct sun hours, annual irradiance, and radiation metrics; lighting based real estate planning and marketing; photovoltaics performance metrics; daylight parameters for floor plans; floor plans internal circulation patterns; urban heat island modelling; and finally pedestrian wind comfort. While this study focuses on daylight analysis, the approach can be applied to create a potentially limitless set of ML-based web services for building energy, comfort, daylighting, and CFD analysis.

### References

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., ... Zheng, X. (2015). *TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems*. 19.
- Bellocchio, F., Ferrari, S., Lazzaroni, M., Cristaldi, L., Rossi, M., Poli, T., & Paolini, R. (2011). Illuminance prediction through SVM regression. *2011 IEEE Workshop on Environmental Energy and Structural Monitoring Systems*, 1–5.
- Chakraborty, D., & Elzarka, H. (2019). Advanced machine learning techniques for building performance simulation: A comparative analysis. *Journal of Building Performance Simulation*, 12(2), 193–207. <https://doi.org/10.1080/19401493.2018.1498538>
- Chatzikonstantinou, I., & Sariyildiz, S. (2016). Approximation of simulation-derived visual comfort indicators in office spaces: A comparative study in machine learning. *Architectural Science Review*, 59(4), 307–322.
- Chollet, F. (2015). *Keras: The Python deep learning API*. <https://keras.io/>
- De Luminae. (2021). *Daylight Autonomy Extension*. De Luminae. Retrieved from <https://deluminaelab.com/dl-light-manual/en/daylightautonomy.html>
- Ekici, B., Kazanasmaz, Z. T., Turrin, M., Taşgetiren, M. F., & Sariyildiz, I. S. (2021). Multi-zone optimisation of high-rise buildings using artificial intelligence for sustainable metropolises. Part 1: Background, methodology, setup, and machine learning results. *Solar Energy*, 224, 373–389. <https://doi.org/10.1016/j.solener.2021.05.083>

- Fedorova, S., Tono, A., Nigam, M. S., Zhang, J., Ahmadnia, A., Bolognesi, C., & Michels, D. L. (2021). Synthetic 3d data generation pipeline for geometric deep learning in architecture. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B2-2021, 337–344.
- Heschong, L. (Chair), Wymelenberg, V. D., (Vice-Chair), K., Andersen, M., Digert, N., Fernandes, L., Keller, A., Loveland, J., McKay, H., Mistrick, R., Mosher, B., Reinhart, C., Rogers, Z., & Tanteri, M. (Ed.). (2012). *Approved method: les spatial daylight autonomy (Sda) and annual sunlight exposure (Ase)*. IES - Illuminating Engineering Society.
- Jones, N. (2019). *Accelerad*. <https://nljones.github.io/Accelerad/index.html>
- Jones, N. L., & Reinhart, C. F. (2017). Experimental validation of ray tracing as a means of image-based visual discomfort prediction. *Building and Environment*, 113, 131–150.
- Kazanasmaz, T., Günaydin, M., & Binol, S. (2009). Artificial neural networks to predict daylight illuminance in office buildings. *Building and Environment*, 44(8), 1751–1757.
- Lee, J., Boubekri, M., & Liang, F. (2019). Impact of Building Design Parameters on Daylighting Metrics Using an Analysis, Prediction, and Optimization Approach Based on Statistical Learning Technique. *Sustainability*, 11(5), 1474.
- Lorenz, C.-L., & Jabi, W. (2017). Predicting Daylight Autonomy Metrics Using Machine Learning. *Sustainable Design of the Built Environment SDBE 2017*, London, UK.
- Lorenz, C.-L., Spaeth, A., Bleil de Souza, C., & Packianather, M. (2019, June 28). Machine Learning in Design Exploration: An Investigation of the Sensitivities of ANN-based Daylight Predictions. *CAAD Futures 2019*, Daejeon.
- McKinney, W. (2010). *Data Structures for Statistical Computing in Python*. 56–61.
- Navada, S. G., Adiga, C. S., & Kini, S. G. (2016). Prediction of daylight availability for visual comfort. *International Journal of Applied Engineering Research*, 11(7), 4711–4717.
- Nasrollahi, N., & Shokri, E. (2016). Daylight illuminance in urban environments for visual comfort and energy performance. *Renewable and Sustainable Energy Reviews*, 66, 861–874. <https://doi.org/10.1016/j.rser.2016.08.052>
- Reback, J., McKinney, W., jbrockmendel, Bossche, J. V. den, Augspurger, T., Cloud, P., gyoung, Sinhrks, Klein, A., Roeschke, M., Hawkins, S., Tratner, J., She, C., Ayd, W., Petersen, T., Garcia, M., Schendel, J., Hayden, A., MomIsBestFriend, ... Mehryar, M. (2020). pandas-dev/pandas: *Pandas 1.0.3*. Zenodo.
- Reinhart, C. F., & Walkenhorst, O. (2001). Validation of dynamic RADIANCE-based daylight simulations for a test office with external blinds. *Energy and Buildings*, 33(7), 683–697. [https://doi.org/10.1016/S0378-7788\(01\)00058-5](https://doi.org/10.1016/S0378-7788(01)00058-5)
- Reinhart, C. F., Mardaljevic, J., & Rogers, Z. (2006). Dynamic Daylight Performance Metrics for Sustainable Building Design. *LEUKOS*, 3(1), 7–31.
- Rubin, B. (1993). Statistical disclosure limitation. *Journal of official Statistics*, 9(2), 461–468.
- Şahin, M., Oğuz, Y., & Büyüktümtürk, F. (2015). Approximate and three-dimensional modeling of brightness levels in interior spaces by using artificial neural networks. *Journal of Electrical Engineering and Technology*, 10(4), 1822–1829.
- Singh, M. M., Singaravel, S., & Geyer, P. (2019). Improving Prediction Accuracy of Machine Learning Energy Prediction Models. *Proceedings of the 36th CIB W*, 78, 2019.
- Seyedzadeh, S., Rahimian, F. P., Glesk, I., & Roper, M. (2018). Machine learning for estimation of building energy consumption and performance: A review. *Visualization in Engineering*, 6(1), 5. <https://doi.org/10.1186/s40327-018-0064-7>
- Solem LLC. (2021). *Why is ClimateStudio so fast?* Solemma. Retrieved August 26, 2021, from <https://www.solem.com/climatestudio/speed>
- Walczak, S., & Cerpa, N. (2003). Artificial Neural Networks. In *Encyclopedia of Physical Science and Technology* (pp. 631–645). <https://doi.org/10.1016/B0-12-227410-5/00837-1>
- Yoon, Y., Moon, J. W., & Kim, S. (2016). Development of annual daylight simulation algorithms for prediction of indoor daylight illuminance. *Energy and Buildings*, 118, 1–17. <https://doi.org/10.1016/j.enbuild.2016.02.030>

# AUTOCOMPLETION OF FLOOR PLANS FOR THE EARLY DESIGN PHASES IN ARCHITECTURE: FOUNDATIONS, EXISTING METHODS, AND RESEARCH OUTLOOK

VIKTOR EISENSTADT<sup>1</sup>, JESSICA BIELSKI<sup>2</sup>, BURAK METE<sup>3</sup>,  
CHRISTOPH LANGENHAN<sup>4</sup>, KLAUS-DIETER ALTHOFF<sup>5</sup> and  
ANDREAS DENGEL<sup>6</sup>

<sup>1,5</sup>*University of Hildesheim, Germany*

<sup>1,5,6</sup>*German Research Center for Artificial Intelligence (DFKI)*

<sup>2,3,4</sup>*Technical University of Munich, Germany*

<sup>1</sup>*viktor.eisenstadt@dfki.de 0000-0001-6567-0943*

<sup>2</sup>*jessica.bielski@tum.de 0000-0003-4936-1993*

<sup>3</sup>*burak.mete@tum.de 0000-0001-8896-6691*

<sup>4</sup>*christoph.langenhan@tum.de 0000-0002-6922-2707*

<sup>5</sup>*klaus-dieter.althoff@dfki.de 0000-0002-7330-6540*

<sup>6</sup>*andreas.dengel@dfki.de 0000-0002-6100-8255*

**Abstract.** This paper contributes the current research state and possible future developments of AI-based autocompletion of architectural floor plans and shows demand for its establishment in computer-aided architectural design to facilitate decent work, economic growth through accelerating the design process to meet the future workload. Foundations of data representations together with the autocompletion contexts are defined, existing methods described and evaluated in the integrated literature review, and criteria for qualitative and sustainable autocompletion are proposed. Subsequently, we contribute three unique deep learning-based autocompletion methods currently in development for the research project metis-II. They are described in detail from a technical point of view on the backdrop of how they adhere to the proposed criteria for creating our novel AI.

**Keywords.** Artificial Intelligence, Architectural Design, Floor Plan, Autocompletion, SDG 8, SDG 9.

## 1. Introduction

The recent technological advances established artificial intelligence (AI) as an essential direction of computer science. Being applied in industry and research as the leading computational method for a number of innovations, AI became ubiquitous in everyday life. Some examples are the personal assistants in mobile devices or automatic translation services. In architecture, however, AI cannot be seen as a leading computational method, mostly due to its absence in the established design software. AI in architecture almost exclusively takes place in the context of research for computer-aided architectural design (CAAD), e.g., for knowledge management approaches of

building information modelling (BIM), where AI in general and its methods deep learning (DL) or case-based reasoning (CBR) in particular, are permanently present at major conferences, journals, and special book editions (As and Basu, 2021).

In this paper, the research area of AI dedicated to autocompletion of missing parts in an architectural design will be discussed. Similar to the suggestions of the next word in a sentence on the modern mobile devices, AI-based autocompletion methods can be applied to the design process in architecture as well, especially in the early conceptual phases where the ideas for the future architectural design are vague and incomplete. The architect can benefit from AI and let it suggest the most probable next design steps based on own and other architects' design experiences of already designed and built architecture stored, e.g., in a case base and used to train neural networks. In the research project metis-II funded by German Research Foundation (DFG), we investigate such autocompletion methods and develop our own using DL and CBR. This paper explores these methods to provide a basis by outlining the current stage of our research and define foundations and possible future developments for AI-supported design process.

## 2. Foundations

To start with the definition of foundations for the autocompletion for floor plans, few sentences should be spent on the background of its application during the early design phase (also: *ideation* phase) in architectural design. During this phase, the concepts and ideas are generated, discussed, and evaluated for the potential of being implemented in the final building (Achten 2002). The early design phase provides a significant impact on the further design process influencing the final layout and utilization (Çavuşoğlu and Çağdaş, 2017). In the construction programs of the Official Scale of Fees for Services by Architects and Engineers in Germany (HOAI) and the American Institute of Architects (AIA) the ideation phases take approx. 30% of the actual design activities. One of the most essential characteristics of the ideation phase is the review of sketches, models, or otherwise represented references of the previously built projects. A common use case for such past references is to provide inspiration, e.g., show integration of current design in a similar context. However, inspiration is “not usually a complete solution to the problem”, but “points to the direction in which the complete solution may be found” (Mahmoodi, 2001, p. 101).

While the review (search, retrieval) can be performed manually, it is more efficient and time-saving to delegate it to a computational method based on CBR or DL (Sharma et al., 2017). In the former research project metis-I, such search methods were investigated and developed (Eisenstadt et al., 2020). In the successor project metis-II, the research goals were extended to the use of past references for deep learning-based autocompletion of architectural designs. By suggesting one or more design step alternatives, the resulting autocompletion methods aim at making the design process more efficient, maintaining or even improving the quality of the designs and the future built environment while accelerating the decision-making of the early design phases, ultimately reducing the work-related stress and catalyse economic growth. The resulting longer-lasting and better building lifecycle can contribute to a more resource-efficient use of building materials (Haeusler et al., 2021) and thus to more sustainable and cost-efficient construction process and maintenance.



2.1. DATA REPRESENTATION

In both *metis* projects, the early designs are interpreted as graph-based spatial configurations derived from the floor plan sketches, where rooms (e.g., Living or Sleeping) represent the nodes and spatial relations (e.g., Door or Passage) represent the edges. To cover multiple design cases and contexts, semantic building fingerprints (SBF, see Figure 1) are applied that use semantic information extracted from the layout graphs. An SBF is considered a data representation type as well as a search pattern.

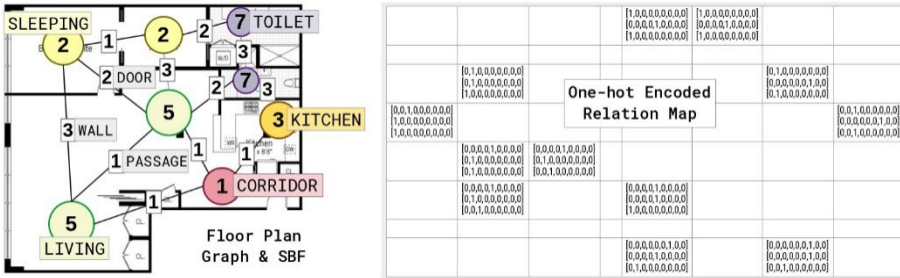


Figure 1. An example of a floor plan SBF graph converted into a one-hot-vector encoded map using a numerical typology for semantic entities room and connection (overlay of the floor plan). (Author illustration)

For retrieval, the pure graph-based representation together with its attribute-value derivations (both based on SBFs) were used. For autocompletion using DL, however, these representations are not applicable directly, as they are not compatible with the popular DL development frameworks, such as TensorFlow. That is, to apply existing SBF-based graph representations for deep learning, two general possibilities exist:

- Development of a *specific intermediate representation* that converts semantic information from the fingerprint into the numerical form processable and properly interpretable by the methods of the popular/all-purpose deep learning frameworks
- Use of a *specific deep learning library*, e.g., Deep Graph Library (DGL), which however does not contain architecture-specific methods and uses generic representations which do not guarantee proper interpretation of the semantics

We investigate both methods for DL-based autocompletion and proposed a specific intermediate representation “relation map” in which SBF-based graphs are converted into 2D or 3D tensors keeping all the relevant semantic information on the spatial layout intact and are compatible with the popular DL frameworks. A relation map is a modified adjacency matrix of the graph in which a specific relation code is used instead of the binary connection value. Relation codes consist of the entries for connected room types (nodes) and the connection type, both based on a provided semantic typology. Three types of relation map were proposed: *one-hot-vector encoded map* (2D, see Figure 1), *numerical multilayer map* (3D), and *textual map* (2D). In the evaluations performed to investigate if the deep learning models of the popular frameworks in the form of convolutional neural networks (CNN) properly interpret the encoded semantic information, the one-hot encoded map showed the best results (Eisenstadt et al., 2021).

## 2.2. APPLICATION CONTEXT

In order to demonstrate the application contexts for the autocompletion methods, a case study based on an existing residential housing design scenario example will be used as a basis throughout this paper. We apply the housing context due to its simplicity and flexibility in comparison to other design domains, e.g., medical buildings that are usually designed by large and specialised architecture offices while residential housing can be designed by big as well as smaller firms. Our research targets are all kinds of architectural practitioners, including architects in academia.

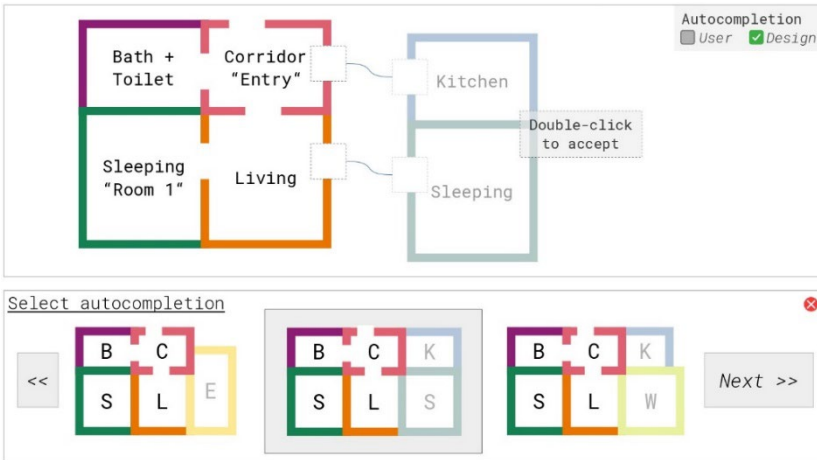


Figure 2. Mock-up of the autocompletion visualization. (Author illustration)

Consider the following design task example: Due to the acute shortage of student housing, new housing units should be built in the Olympic Village of Munich. Two units should be constructed in the first step. The one-storey high unit complex should be detached from the already existing surrounding buildings. The main facade with the entrance of the 1st unit faces North, while the orientation of the 2nd is east-bound. The area size of each parcel is approximately 52 sqm. The room program consists of 1 living room, 1 kitchen, 1 bathroom, and 2-3 sleeping rooms.

The task of the system that implements AI-based methods would be to utilize the selected autocompletion context and provide recommendations for the next design steps or automatically filling in the missing parts of the spatial configuration. That is, we differentiate between two foundational autocompletion contexts:

- **User context:** AI suggests the most probable next design step (e.g., add or delete space) based on the past design steps made by the architect and other architects
- **Design context:** AI fills in the missing parts (rooms, connections) of the spatial configuration graph based on the previous complete graphs that it has learned from

In Figure 2, a mock-up of the visualization of the autocompletion according to the student dormitory design task example is shown: Living room, bath, and the 1st sleeping room are already set, using the design context selected by designer, the system suggests adding the missing kitchen and the 2nd sleeping room and other alternatives.

### 3. Literature Review of Existing Methods

In order to follow the research goals of the metis projects and the SDGs 8 and 9 (see Section 2), we propose criteria for qualitative AI-based floor plan autocompletion to enhance sustainability in digitization of the early phases of architectural design process:

- The AI methods can be *extended with further features* and are *fit for use in future*
- The results returned by the methods are *justified* by being *compliant with the design requirements* of the architectural domain they were developed for (e.g., housing)
- The data should be *properly obtained* and/or *reasonably synthesized*

The existing floor plan autocompletion methods, found in the contemporary literature, will be reviewed in the current section to estimate if they adhere to these criteria.

#### 3.1. SHALLOWDREAM

Bayer et al. (2018) presented an approach for generation and auto-completion of floor plans using recurrent neural networks (RNN) based on LSTM (Long Short-Term Memory) structure. The authors propose a *ShallowDream* structure that uses sequences of feature vectors that contain information blocks that describe the floor plans in terms of functionalities of the rooms, connections, and geometrical features. Blocks can be generated by the LSTM “Block Generation Sequencer” and the features of the vectors can be predicted using the “Vector Prediction Sequencer”. The *ShallowDream* structure can be used for room connection generation and room layout generation, i.e., effectively being an autocompletion approach based on augmentation of architectural data. This approach can clearly be extended with further architectural features for the vectors and blocks, and considered sustainable as it is based on the basic neural network structure LSTM. However, its evaluation showed that the approach is error-prone which might lead to the poor quality of generated data. Thus, the methods’ results cannot be considered justifiable for practitioners and/or domain experts.

#### 3.2. CITYMATRIX

From the related domain of urban configuration generation, the approach *CityMatrix* (Zhang Y. et al., 2018) uses different methods of machine learning to produce a suggestion for optimization of the urban design for different domain-related issues, such as solar access or traffic performance. The designs in *CityMatrix* are created with a physical user interface and then used as input into ML methods for optimization. The ML methods used include linear regression, K-Nearest-Neighbor, Random Forest and convolutional neural networks. Specifically, the CNNs were used for user guidance providing real-time feedback and so allowing for quicker decision-making. *CityMatrix* allows for addition of further methods as well, and the use of CNNs can be considered future-proof, due to the same reasons as LSTM. The *CityMatrix* methods, however, need to be transferred to semantic graph-based floor plans as data representation differs.

In this regard, *CityMatrix* is very similar to the approach for interior prediction using activity-associated relation graphs (Fu et al., 2017) and the walls number estimation method (Fafoutis et al., 2015). All of them require significant modification and data representation transfer to be relevant for the autocompletion of floor plans.

### 3.3. BHK-BASED APPROACH

In the most recent approach, Rajasenbagam et al. (2020) present a methodology for space probability prediction using user input in the form of total area and a *BHK* object (*Bedroom, Hall, Kitchen*). A deep CNN is applied to generate and combine room types in context of already generated room types and predict feasible spatial configurations using filters, i.e., user requirements. A specific characteristic of the approach is that the cultural aspects in the form of “*vaasthu*” (architectural science teachings in India and China) can be included in the input data as well. The authors consider their approach time-saving and platform-independent, claiming that it is able to generate design suggestions that differ from “normal designs”. While no detailed information on applied CNN is available in the publication, it still can be considered extendable and future-fit like both previously described methods. The approach also provides a user interface; however, the decisions of the system seem not to be transparent. Quality of generated plans is questionable as the described evaluation seems to be incomplete.

### 3.4. FPDX

The approach *FPDX* (*Floor Plan Design Expert System*) (Yau et al. 1994) uses then popular methods of model-based reasoning (a method closely related to CBR) to suggest modifications to the residential building designs in their early stage of development if they do not fit the official constraints and regulations (so-called “government codes”). The system consists of the following components:

- *User Interface* (AutoCAD- and AutoLISP-based) for floor plan drawing input
- *Floor Plan Compiler* for representation of object relations recognized in the floor plan in the form of syntactic data based on the provided typology (e.g., habitable or non-habitable rooms and their respective sub-types such as living room or kitchen)
- *Design Data Base* for intermediate saving of object relations representation
- *Inference Engine* that validates the fulfilment of the official regulations by the recognized relations and returns feedback for improvement
- Knowledge Base of relation representation references as the basis for validation

FPDX can be seen as the most complete approach presented so far. Its application context and domain are clearly outlined and the authors were able to justify the application of model-based methods, considering them more robust and flexible than CBR. This allows for proper extension of the models and components. Consistency of the final output is ensured by the government codes. However, the approach cannot be considered sustainable as the AI methods applied are outdated by today's AI standards.

## 4. Research Outlook for Our Work

Existing methods, as found in contemporary literature (see Sections 3.1-3.4), lack one or more features to accurately correspond to the AI-based autocompletion quality criteria proposed (see Section 3). In the next sections, AI-based solutions in active development by us for project *metis-II* will be individually detailed out, accompanied by our views on how these solutions plan to adhere better to the quality criteria.

#### 4.1. CLUSTER COMPLETION WITH RECURRENT NEURAL NETWORKS

The first autocompletion method presented applies established methods of *graph clustering* and *sequence learning* with *recurrent neural networks* to predict the *most probable next neighbours* in an incomplete (“problematic”) spatial cluster. It represents the design context (see Section 2.2) and consists of the following steps (see Figure 3):

- *Identify problematic clusters* in the current design using graph clustering methods compatible with the domain of architectural design
- *Complete the cluster* using its sequential representation as a query for an RNN that is trained on a dataset of complete clusters to predict the next sequence members

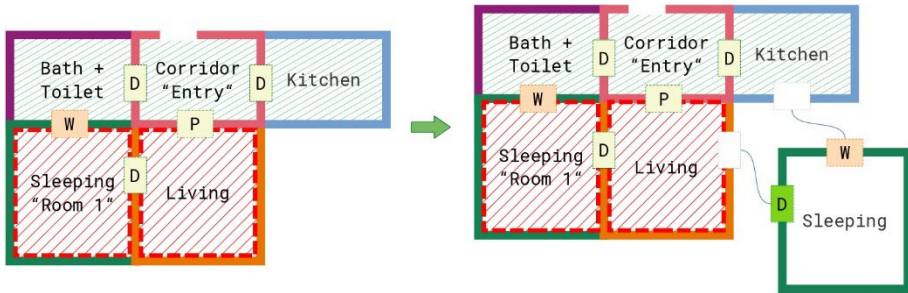


Figure 3. Application of cluster completion with RNNs. In the design scenario example from Section 2.2, a problematic cluster is identified on the left side (red-hatched rooms), used as the input into the RNN in a relation map form, and completed with a new room and connections on the right side as the output. Relations are encoded as follows: W (wall), P (passage), D (door). (Author illustration)

For the first step, a number of graph clustering methods can be applied. For example, the Girvan-Newman clustering algorithm (Girvan and Newman, 2002) can be used to cluster the current design graph using the highly frequented pass-through connections (e.g., Door, Passage, Entrance) as cluster boundaries, whereas the non-pass-through connections (e.g., Wall or Slab) will be excluded from such clustering analysis. Also, distance-based clustering, e.g., DBSCAN (Ester et al., 1996), can be used to find clusters based on density between the cluster candidates. The problematic clusters can be then identified using cluster consistency rules developed as an extension of an already existing component “*Consistency Checker*” (Arora et al., 2021) that checks if the entire graph is following the housing consistency rules.

For the second step, an RNN can be applied that uses the sequential representation of a relation map (see Section 2.1) for data it will be trained on (previous consistent clusters). The data for training is compiled using the spatial configuration data from evaluations of relation maps (see Section 2.1), clustered with the method applied for the autocompletion process. Multiple rooms at once (see Figure 2) and the corresponding connections using a stepwise check, can be suggested.

The combination of clustering and RNNs has the potential to increase the quality of the design suggesting valid solutions based on knowledge acquired from the consistent data. It recommends rooms that do not exist in the design yet, providing the architect with assistance in completion of the problematic housing spaces.

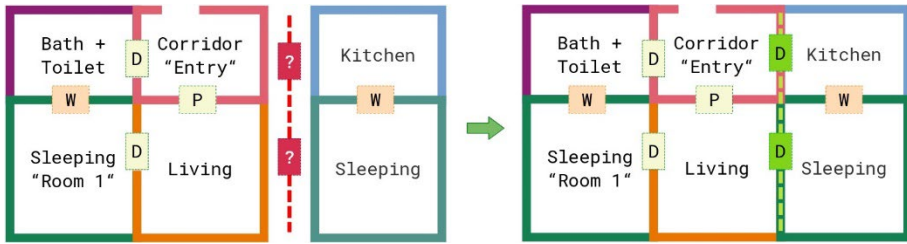


Figure 4. Application example of link prediction with graph-based neural networks. Green-coloured relations on the right side (output) represent the estimated connections of the two spatial areas isolated from each other (input on the left side). (Author illustration)

#### 4.2. LINK PREDICTION WITH GRAPH NEURAL NETWORKS

The next autocompletion approach proposed uses the *link prediction (LP)* methods based on *graph neural networks (GNN)* to estimate the most probable connections within the spatial configuration graph at hand. Link prediction is an established AI method that looks for the *most common neighbours* of the nodes in the network. Usually, LP is applied in recommender systems for e-commerce or online social networks, in which relations between items or users are estimated to recommend related products or make friend suggestions. GNNs were already applied to investigate DL for link prediction, e.g., in the approach SEAL (Zhang M. et al., 2018).

For autocompletion of room configurations, link prediction with GNNs can be applied to predict the connections between the rooms that are not yet connected with each other. For example, if the current design contains two spatial areas isolated from each other, GNN can predict the most probable connections between them (see Figure 4 based on design task example from Section 2.2). That is, link prediction should be applied in the design context, for which the semantic fingerprint of the graph will be converted into a relation map representation compatible with the GNN model. The designer can benefit from link prediction by receiving timely relation suggestions, based on the references and experiences made and saved by architects in the training dataset of the developed GNN. A further development of this method may suggest changes and support the architect through drawing attention to potentially flawed relations. Thus, the system provides alternatives and consequently further insight, and the overall quality of the design can be improved.

#### 4.3. DESIGN STEP PREDICTION USING RNN

Third and last autocompletion method proposed is a representative of the user context. That is, the design structure and content will not be used as foundation for prediction, instead, the *sequential design process models* for the early design phase will be investigated for the *recommendation of continuation of the current design direction*. The design actions made by the architect should be evaluated and used as a query to the *contextually suitable RNN* in order to get the most probable next design step.

Based on methods implemented in an already existing system developed for the *metis* projects, an extended and more advanced version of the design step recommendation system (see Figure 5 based on design task example from Section 2.2)

applies grouping of design steps into reasonable sequences based on recordings of an initial study with representatives from the CAAD / architecture domain whose design sketching processes will be evaluated and the corresponding actions assigned into one of the phases of the common design model *ASE* (*Analysis, Synthesis, Evaluation*) (Lawson, 2005). Based on the determined design phase, using e.g., CBR or an additional classification neural network, the action sequence will be completed with the step predicted by the RNN specific for this design phase.

One of the key challenges of this method is the acquisition of proper data for the training of the RNNs. As currently no sufficient, publicly available, datasets of design steps exist, data augmentation methods will be applied to synthesize the design step sequences, based on design process data acquired during the initial study. During the data synthesis process, a consistency method derived from the consistency checker evaluates the intermediate design states and so ensures the quality of the process data.

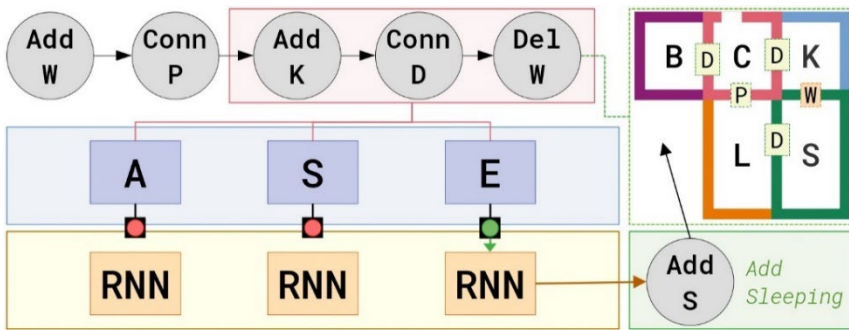


Figure 5. A sub-sequence of the design process is used as the input to the most suitable contextual RNN that produces a design step prediction as the output. (Author illustration)

## 5. Discussion on Quality Assessment of Our Autocompletion Methods

In order to accurately correspond to the proposed quality criteria (see Section 3), each of our autocompletion methods is conceptualized to be fit for use in future work, as the DL models apply basic neural network structures upon which many other ANN types are developed. That is, each method should be in theory usable with any other new ANN type from the same category, e.g., RNNs based on GRU (Gated Recurrent Unit) should work as good as LSTM-based ones. This includes addition of new features to the data representations, e.g., new semantic features (shape, area) for the relation maps.

Besides the consistency assurance methods for housing, results of each DL-based autocompletion method can be checked for compliance with other domains, governmental restrictions/regulations (see FPDx), or cultural aspects (see BHK). Being trained on consistent datasets only, the methods should in theory return consistent results only, however, more evaluations are required to support this theory.

Training datasets for the methods are based on real data collected from the design process recordings and spatial configuration references of existing architecture or manually validated floor plan concepts (e.g., from the student projects). If the amount is not sufficient for training, data augmentation is applied to increase it. The results of data augmentation are validated by the architectural consistency checker as well.

Summarizing, it can be concluded that meeting of the autocompletion quality criteria as described above contributes to the achievement of the research goals of the project metis-II and sets the basis for sustainable support of the ideation phase and its outcomes (see Section 2). While the research is still in its initial phase with expectable problems ahead, it adds a promising option to the discovered potential of AI in architecture.

## References

- Achten, H. H. (2002). Requirements for collaborative design in architecture. In *6th Sixth Design and Decision Support Systems in Architecture and Urban Planning*. (pp. 1-13)
- Arora, H., Bielski, J., Eisenstadt, V., Langenhan, C., Ziegler, C., Althoff, K. D., & Dengel, A. (2021). Consistency Checker-An automatic constraint-based evaluator for housing spatial configurations. *ECAADE 2021*.
- As, I., & Basu, P. (Eds.). (2021). *The Routledge Companion to Artificial Intelligence in Architecture*. Routledge. ISBN 9780367824259.
- Bayer, J., Bukhari, S., & Dengel, A. (2018). Interactive Design Support for Architecture Projects During Early Phases Based on Recurrent Neural Networks. *ICPRAM 2018*.
- Çavuşoğlu, Ö. H., & Çağdaş, G. (2017). Why Do We Need Building Information Modeling (BIM) in Conceptual Design Phase?. *CAAD Futures 2017*.
- Eisenstadt, V., Langenhan, C., Althoff, K. D., & Dengel, A. (2020). Improved and visually enhanced case-based retrieval of room configurations for assistance in architectural design education. In *International Conference on CBR* (pp. 213-228). Springer, Cham.
- Eisenstadt, V., Arora, H., Ziegler, C., Bielski, J., Langenhan, C., Althoff, K. D., & Dengel, A. (2021). Exploring optimal ways to represent topological and spatial features of building designs in deep learning methods and applications for architecture. *CAADRIA 2021*.
- Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *KDD 96*(34) (pp. 226-231).
- Fafoutis, X., Mellios, E., Twomey, N., Diethel, T., Hilton, G., & Piechocki, R. (2015). An RSSI-based wall prediction model for residential floor map construction. In *2015 IEEE 2nd World Forum on Internet of Things (WF-IoT)* (pp. 357-362). IEEE.
- Fu, Q., Chen, X., Wang, X., Wen, S., Zhou, B., & Fu, H. (2017). Adaptive synthesis of indoor scenes via activity-associated object relation graphs. *ACM TOG*, 36(6), 1-13.
- Girvan, M., & Newman, M. E. (2002). Community structure in social and biological networks. *Proceedings of the national academy of sciences*, 99(12) (pp. 7821-7826).
- Haeusler, M.H., Butler, A., Gardner, N., Sepasgozar, S., & Pan, S. (2021). 'Wasted... again: Or how to understand waste as a data problem and aiming to address the reduction of waste as a computational challenge'. *CAADRIA 2021*.
- Lawson, B. (2005). *How Designers Think*. 4th edition, Routledge. ISBN 9780080454979.
- Mahmoodi, A. S. (2001). *The design process in architecture: a pedagogic approach using interactive thinking*. Doctoral dissertation, University of Leeds.
- Rajasenbagam, T., Jeyanthi, S., & Uma Maheswari, N. (2020). Floor Plan Designer Application by Predicting Spatial Configuration Using Machine Learning. *ICTCS 2020*.
- Sharma, D., Gupta, N., Chattopadhyay, C., & Mehta, S. (2017). Daniel: A deep architecture for automatic analysis and retrieval of building floor plans. In *14th International Conference on Document Analysis and Recognition*, 1 (pp. 420-425). IEEE.
- Yau, M. Y., Lai, E. K., & Chun, H. W. (1994). FPDx: A knowledge-based system for architectural floor plan design. In *Proceedings of International Conference on Expert Systems for Development* (pp. 309-314). IEEE.
- Zhang, M., & Chen, Y. (2018). Link prediction based on graph neural networks. In *Advances in Neural Information Processing Systems*, 31 (pp. 5165-5175).
- Zhang, Y., Aubuchon, A., Lyons, K., & Larson, K. (2018). Machine learning for real-time urban metrics and design recommendations. *ACADIA 2018*.



# DESIGN INTENTS DISENTANGLEMENT

*A Multimodal Approach for Grounding Design Attributes in Objects*

MANUEL LADRON DE GUEVARA<sup>1</sup>, ALEXANDER SCHNEID-  
MAN<sup>2</sup>, DARAGH BYRNE<sup>3</sup> and RAMESH KRISHNAMURTI<sup>4</sup>

<sup>1,2,3,4</sup>*Carnegie Mellon University, USA.*

<sup>1</sup>*manuelr@andrew.cmu.edu, 0000-0002-4585-3213*

<sup>2</sup>*amschnei@andrew.cmu.edu, 0000-0002-3788-1193*

<sup>3</sup>*daraghb@andrew.cmu.edu, 0000-0001-7193-006X*

<sup>4</sup>*ramesh@andrew.cmu.edu, 0000-0002-6327-8286*

**Abstract.** Language is ambiguous; many terms and expressions convey the same idea. This is especially true in design fields, where conceptual ideas are generally described by high-level, qualitative attributes, called design intents. Words such as "organic", sequences like "this chair is a mixture between Japanese aesthetics and Scandinavian design" or more complex structures such as "we made the furniture layering materials like a bird weaving its nest" represent design intents. Furthermore, most design intents do not have unique visual representations, and are highly entangled within the design artifact, leading to complex relationships between language and images. This paper examines an alternative design scenario based on everyday natural language used by designers, where inputs such as a minimal and sleek looking chair are visually inferred by algorithms that have previously learned complex associations between designs and intents—vision and language, respectively. We propose a multimodal sequence-to-sequence model which takes in design images and their corresponding descriptions and outputs a probability distribution over regions of the images in which design attributes are grounded. Expectedly, our model can reason and ground objective descriptors such as black or curved. Surprisingly, our model can reason about and ground more complex subjective attributes such as rippled or free, suggesting potential regions where the design object might register such vague descriptions. Link to code: <https://github.com/manuelladron/codedBert.git>

**Keywords.** Natural Language Processing; Multimodal Machine Learning; Design Intents Disentanglement; SDG 9.

## 1. Introduction

Language can be ambiguous and similar ideas can be expressed in many different expressions. This is especially true in design fields, where conceptual ideas are generally described by high-level, qualitative attributes, called design intents. Even

though these descriptors are highly used in everyday language by designers—"the dining table should look more organic", "this chair is lightweight and minimal"—, they have complex visual associations due to partial subjective and conceptual components and thus, finding visual representations is a challenge. While humans might be able to identify design intents from an image of a chair with attributes such as "organic" or "minimalist" and differentiate between a "heavyweight" and a "lightweight" stand-lamp, they might also face challenges differentiating design intents such as "dynamic", "organic" or "functional". Current machine learning literature is unable to recognize these types of high-level attributes but has potential to understand them. Resolving such task would have a major impact in design communities, opening new scenarios where natural human language could directly be used in the process of design.

For computational linguistics, resolving this problem can challenge the status of theoretical understanding, problem-solving methods, and evaluation techniques (Alm, 2011). For computer vision, this presents a complex challenge of disentangling qualitative attributes—sleek, elegant, minimal—from images. Beyond its relevance in pushing machine learning research boundaries, this would significantly impact creative practice—designers, architects, and engineers. Real-time design intents understanding could open new design scenarios (e.g., voice-assisted natural language input), that reduce procedures based on intent reinterpretation as imperative commands—move, circle, radius, extrude, vertical—required by digital design engines, like AutoCAD or Rhinoceros. Such methods require a lengthy sequential process of deterministic commands that manually shape the design object.

Research on identifying high-level attributes has been done in other tasks. For instance, for selecting font types using attributes by (O'Donovan et al., 2014), or for search tasks on fashion objects with relative attribute feedback by (Kovashka et al., 2012).

In this work we aim to ground such relationships between modalities. We expand upon FashionBert (Gao et al., 2020), a framework that tackles a similar problem of disentangling design elements. The fashion descriptions from their data, however, are purely focused on the designs themselves, and most descriptions are objective rather than conceptual. In addition, we modify their image encoding strategy to make our model completely agnostic to the object itself. Finally, we adopt a modified token masking scheme to place more weight on masking adjective tokens because most keywords in design intents, such as "minimal" or "organic" are adjectives, and we hypothesize that these should provide a better training signal with respect to using the visual modality to infer language.

## 2. Related Work

This section firstly reviews work related to high-level attributes and design. We then review prior work done using the CODED dataset.

### 2.1. HIGH-LEVEL ATTRIBUTES

Research in understanding the relationships between high-level attributes and objects has not received much attention in comparison with objective or quantitative attributes. Some previous works have focused on image composition, particularly on

high-level attributes of beauty, interest, and memorability and some authors described methods to predict aesthetic quality of photographs. (Datta et al., 2006) represent concepts such as colorfulness, saturation or rule-of-thirds with designed visual features, and evaluated aesthetic rating predictions on photographs. (Li & Chen, 2009) used the same approach for impressionist paintings. (Gygli et al., 2013) predicted human evaluation of image interestingness, building on work by (Isola et al., 2011), who uses various high-level features to predict human judgements of image memorability. Similarly, (Borth et al., 2013) performed sentiment analysis on images using object classifiers trained on adjective-noun pairs.

Object attributes have been explored for image search using binary attributes in works by (Kumar et al., 2011; Tao et al., 2009). Other work for searching interfaces has been done by (Parikh & Grauman, 2011), which estimates relative attributes. Whittlesearch (Kovashka et al., 2012) allows searching image collections using also relative attributes. In the following year (Kovashka & Grauman, 2013) improved their technique by using an adaptive model.

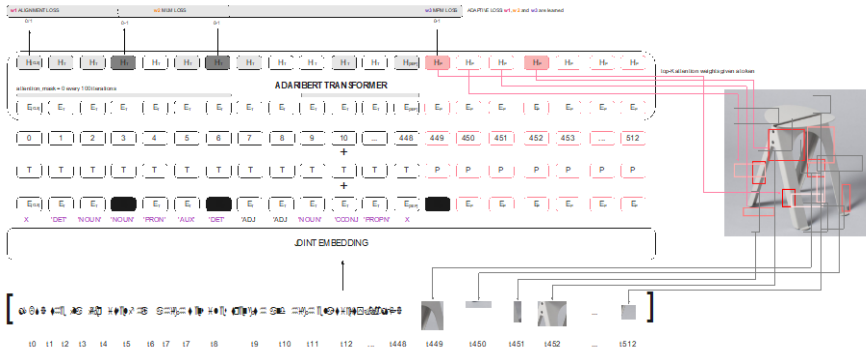


Figure 1. Overview of the CODEDBERT model. All images subjected to copyright. Used here for research purpose. Image: Dumbo, by Studio Pesi

## 2.2. PRIOR WORK ON THE CODED DATASET

Preliminary explorations that visually disambiguate vague terms in the context of design have been done by (Ladron de Guevara et al., 2020). The authors use a multimodal approach that combines a pretrained convolutional neural network to get the representation for images with general word indexes into a common joint subspace. A bidirectional Long Short-Term Memory (biLSTM) decoder—which models the labels co-occurrence information—learns semantic relationships between words and images.

To our knowledge, our work is the first attempt to scale work on high-level attributes on very complex unsupervised scenarios, where images do not have ground truth associating descriptors. Furthermore, CODED's language modality does not come from third party workers annotating the data, but the very own natural description from professional designers, which is a key element to integrating this knowledge and comprehending designs more naturally. While (Gao et al., 2020) explore similar

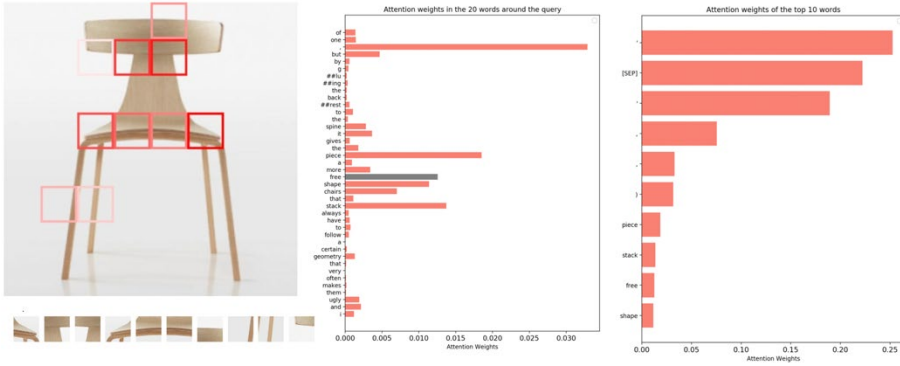


Figure 2. CODEDBERT at inference. (Left) Attention windows for the design intent "free"; higher color intensity corresponds to higher attention. (Middle) Attention weights for 20 neighboring words. (Right) Attention weights for top 10 words.

problems, their dataset only contains valid descriptors and most of them are objective attributes, our approach differs in two main aspects: unstructured language modality and high-level attributes. This motivates different masking and patching schemes.

### 3. Method

#### 3.1. BERT

BERT was introduced by (Devlin et al., 2018) and uses the Transformer architecture (Vaswani et al., 2017) with a word-piece approach that divides words into tokens to be fed into the model. Every subword is projected into a set of embeddings  $E$ , and each embedding in  $E$  is computed as a sum of token embeddings particular to each subword. Each segment embedding indicates the part of the text that comes from, and a position embedding encodes the position of each token. This is fed into the multilayer BERT transformer, which generally has 12 or 24 layers, and outputs contextualized representations of each token. BERT is generally trained in two steps, pre-training and fine-tuning, where the former step is done using a combination of two objectives, mask language modeling and next sentence prediction. The latter step generally applies BERT to a particular task using different objectives according to the training task.

#### 3.2. CODEDBERT

Our approach is to use the self-attention mechanisms of the BERT model to address two main tasks: (1) ground high-level attributes in images in an object-agnostic approach and (2) use within-language attention mechanisms as means to find relevant parts of our unstructured text and filter out those sentences that do not relate to the images, such as contextual information.

In addition to the text components in BERT, we introduce a set of visual tokens to model an image and learn joint representations of design intents and design images. The CODEDBERT model is illustrated in Figure 1. Generally, vision-and-language BERT models like VisualBert or ViLBert use object-detection methods to extract

objects within the image and pass the entire isolated objects through the multimodal Transformer. Our research differs from such methods in a fundamental way: we build an object-agnostic model that disentangles design intents. For this approach, we propose two strategies to process the images. Following the FashionBert model, we resize images to 64x64 pixels and break them into 64 patches of 8x8 pixels each. We pass this sequence of patches along with a position embedding which theoretically gives the model the option to reconstruct the entire image. We called this strategy normal patching (NP). With the intention of destroying the object representation completely, our second approach takes a step further and produces random patches of positions and sizes, with a minimum of 4 and a maximum of 16 pixels in width and height. In this case, since the patches are randomly generated, there is no explicit order, and therefore, we do not use any position embeddings. We called this approach random patches (RP)—see Figure 4. The model is then tasked with grounding these patches with the design intent in the original description. We use the publicly available transformer library by HuggingFace, as a backbone for our implementation. Link to code: <https://github.com/manuelladron/codedBert.git>

### 3.2.1. Training CODEDBERT

Our training schema for CODEDBERT consists of three main objectives:

**Ground Masked Language Modeling (MLM):** This is a regular BERT training task in which text tokens, encoded using the wordpiece strategy, are masked with a probability of 15% and the model must minimize the negative log likelihood of predicting the original mask token, using surrounding language tokens and vision patch tokens. Given a sequence of text tokens  $t_i = t_1, t_2, \dots, t_N$  the masked-out sequence is denoted by  $t_k = t_1, t_{\text{MASK}}, \dots, t_K$ . The last layer of the transformer model output is fed into a linear classifier with vocabulary size the standard BERT model. The objective is defined as follows:

$$l_{\text{MLM}}(\theta) = -E_{t \sim D} \log P(t_i | t_k, \theta)$$

where  $\theta$  corresponds to the CODEDBERT parameters,  $D$  is the training set and  $P(t_i | t_{/i})$  is the probability of the masked-out token  $t_i$  predicted by the model given the rest of the language and vision tokens.

We propose a variant of the MLM approach that focuses on masking those words that their part of speech tag corresponds to adjectives with a probability of 13.5%, while masking non-adjectives with probability of 1.5%. We hypothesize that most design intents are captured by adjectives, and this strategy spends more training time predicting potentially more useful words for design grounding. Use within-language attention mechanisms as means to find relevant parts of our unstructured text and filter out those sentences that do not relate to the images, such as contextual information.

**Text-Patches Alignment (TPA):** We pre-process each sample in the dataset by pairing each image with a random text sample with a uniform distribution over the dataset. The model must predict whether any given text-image sample is paired or not. To do that, we use the pooled output from the BERT model, which is a dense representation of the [CLS] token of the entire sequence and pass it through a binary linear classifier. TPA is trained using binary cross entropy loss as follows:

$$l_{atig} - \frac{1}{n} \sum_{i=1}^n y \log(P(\hat{y})) + (1 - y) \log(1 - P(\hat{y}))$$

**Masked Patch Modeling (MPM):** Like the MLP task, we randomly mask out patches with probability 10%, setting the image encoder features to zero. We treat the output features as distributions over the original patches' features, and the model tries to minimize the KL divergence between the true patch features and the output masked-out features by:

$$l_{MPM}(\theta) = E_{KLp \sim D}(Distr.(p_i | p_k, \theta) | Distr.(p_i))$$

### 3.2.2. Adaptive Loss

Following the fashionBERT model training strategy, we employ an adaptive loss algorithm to learn each of the three weights corresponding to each loss. Given the initial total loss function as

$$\mathcal{L}(\theta) = \sum_{i=1}^L w_i l_i(\theta)$$

where  $L = 3$ . We let the model learn the weights  $w_i$  as a new optimal problem:

$$\begin{aligned} \operatorname{argmin} & -\frac{1}{2} \sum_{i=1}^L \|w_i \nabla l_i\|^2 + \frac{1}{2} \sum_{i,j=1}^L \|w_i - w_j\|^2 \\ \text{s. t.} & \sum_{i=1}^L w_i = 1 \text{ and } \exists w_i \end{aligned}$$

This formulation aims to minimize the total loss while fairly treating the learning of all tasks. Considering the Karush-Kuhn-Tucher (KKT) conditions, we obtain the solution to  $w_i$  as:

$$w_i = \frac{(L - \nabla l_i^2)^{-1}}{\sum_{i=1}^L (L - \nabla l_i^2)^{-1}}$$

## 4. Experimental Setup

The CODED dataset is unique in that the text modality is not annotated by third-party workers but it is indeed leveraged from design conversations about the work in question. On the one hand this is an excellent resource for understanding design intents, as the dataset contains the original design intents rather than external labels. The downside is that the text includes rejections, negations and contextual information

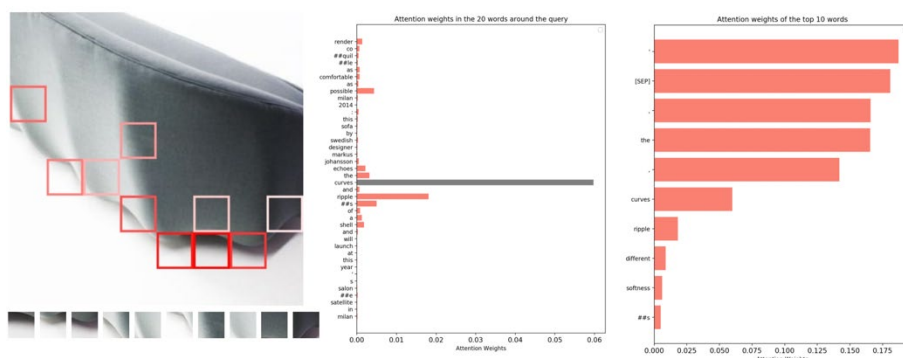


Figure 3. CODEDBERT at inference. (Left) Attention windows for the design intent "curves"; higher color intensity corresponds to higher attention. (Middle) Attention weights for 20 neighboring words. (Right) Attention weights for top 10 words.

about the designer's philosophy that can be vaguely applied to the objects we see in images. Our research examines whether it is possible to ground design intents under an object-agnostic schema, and therefore, disentangle such high-level attributes from the objects in the image modality.

#### 4.1. DATASET AND INPUT MODALITIES

To address the issue of disentangling design intents in the context of creative practice, we use the CODED dataset, first presented in (Ladron de Guevara et al., 2020). The self-annotated CODED dataset contains a total of 33,230 samples of contemporary creative works represented by 264,028 raw sentences—provided by the original creators and by art curators—that describe 241,982 images. This dataset was assembled by collecting articles that include editorial descriptions along with associated images of the creative visual work. CODED is an organized dataset divided into seven categories: "architecture, art, design, product design, furniture, fashion" and "technology". CODED is the first dataset of pairs of images and language that, besides containing objective information of the elements in the images such as "wooden chair" or "black table", focuses on high-level attributes that correspond to design intents, such as "minimal, elegant and sleek looking chair". CODED also contains contextual information that can or cannot be indirectly applied to the images, and other more complex structures such as analogies. Note that there are ground truth labels in each image pair, but they are not indexed.

For all the experiments shown in this paper, we work with the furniture domain within the CODED dataset, and we use word modality for all our experiments. The Furniture-CODED dataset contains 17,532 images of contemporary workpieces.

#### 4.2. MULTIMODAL BASELINE MODELS

We test 4 main experiments and provide a series of ablation studies to compare their performance. Namely, these experiments are:

**Normal Patching and Normal Masking.** The image is equally divided into 64 patches of 8x8 pixels, and we apply a normal common language masking strategy. The features of the patches are extracted using a pretrained ResNet50.

**Random Patching and Normal Masking.** The image is randomly divided into 64 patches of varying dimensions. We apply common masking and use pretrained ResNet50 as feature extractor.

**Normal Patching and Adjective Masking.** The images are divided by the NP approach, and we use the adjective masking approach defined in section 3. To extract the features of the patches, we fine-tuned ResNet152 trained on CODED under a contrastive learning strategy for cross-modal retrieval tasks. We also experiment with and without image-attention only by enforcing the attention mask every 100 iterations. In our model, the language modality is defined by a sequence of 448 tokens, whereas the image modality contains only 64. By looking only at the image attention weights, enforce the model to predict the masked tokens and alignment only from the visual context.

**Random Patching and Adjective Masking.** Like the prior experiment but using the RP approach.

#### 4.3. IMPLEMENTATION DETAILS

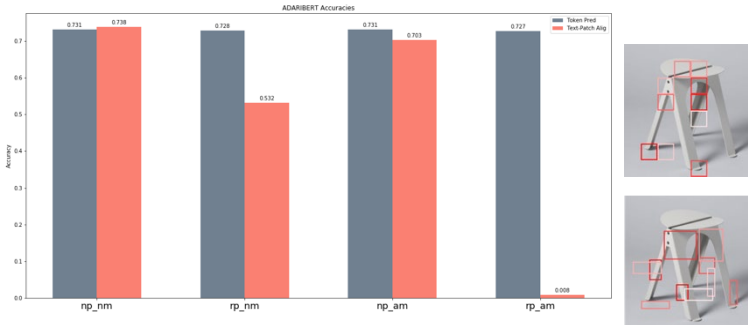


Figure 4. Accuracy between models (left). Random vs Normal patches approach (right)

We test the performance of our CODEDBERT model on two tasks, text and image alignment and masked token prediction. For text-patch alignment and token prediction, we employ accuracy as a metric to test our experiments, on a subset of 1,000 samples drawn from the test set.

We use a pre-trained 12-layers Bert-Base uncased model, which has 12 attention heads, and each hidden unit has 768D, with a total sequence length of 512. We set the text sequence to 448D, leaving the remainder 64D for the image input. For each patch, we discard the last fully connected layer of the ResNet model to have a 2,048-dimensional feature representation. We train the models for 5, 10 and 20 epochs. We discard results from the 20 epochs approach due to high overfitting issues. We set batch size to 4 due to VRAM memory constraints. We use p2.xlarge for training and g4dn.xlarge for evaluation. We use Adam optimizer with a learning rate of  $2e-5$ ,  $\beta_1 =$



0.95,  $\beta_2 = 0.999$ , weight decay of  $1e-4$  and a linear schedule with warmup set to 5,000 steps. We clip gradients at 1.0.

## 5. Results and Discussion

We compare our four main experiments. The notation for our ablation studies is as follows: our baseline is defined by equal or normal patches and normal masking (NPNM), and it is compared against random patches and normal masking (RPNM), equal patches and adjective masking (NPAM), and random patches and adjective masking (RPAM).

We measure the performance of our models with the accuracy metric, as seen in Figure 4. For masked token prediction, slicing the image into equal patches yields slightly better results than using the random patches approach. Our best performing model in this task is our baseline (NPNM) with an accuracy of 73.15% followed very closely by NPAM with 73.10% accuracy. Likewise, the random patches approach yields almost same performance. We hypothesize that for predicting masked language tokens, the model is agnostic to the vision modality, ignoring the slicing approach. This might be due to the imbalance in the sequence length of the two modalities, or that the attention mechanism is not as strong in the vision modality as it is in the language part. We observe that the fact that skewing the masking towards adjectives and masking from a uniform distribution across the text sequence does not impact significantly in the performance of the language token prediction task.

In the alignment task, however, the slicing approach has a significant impact. We see how our baseline outperforms the rest of the ablative experiments with a 73.80% accuracy. The equal patches approach is very superior to the random patches approach. RPNM performs near to random guessing with a 53.20% accuracy, while that RPAM significantly underperforms a random classifier with a 0.8% accuracy. A reason for such low performance of the random patch scheme is that random patches do not assure covering the entire distribution of the pixels in the images, contrary to the equal patch strategy. The random patches might be scattered about a noisy background, providing a weak signal.

## 6. Conclusion and Future Work

We developed a transformer-based model called CODEDBERT, a joint model for image and text descriptions to address two main tasks: grounding design intents in images under an object-agnostic approach and finding the meaningful parts of the lengthy and noisy text leveraging the use of the attention mechanisms. Figure 2 and Figure 3 show how our model reasons between a chair and corresponding intent "free" and a couch and the attribute "curves", respectively. Figure 2 (left) visualizes with higher intensity those regions of the chair that are more likely to ground the design intent "free". As difficult as this task is, the model focuses the attention on the unusual kinks and convex parts of the chair. Likewise, Figure 3 correctly associates the attribute "curves" in the image.

A challenging and unique aspect of CODED is the descriptions are from long-form interviews with designers. Thus, a large portion of a given description may not be relevant to the design. We had hoped BERT could learn to distinguish which parts are

relevant, but our training tasks may not have encouraged this enough. One possible direction would be to take a hierarchical approach, encoding entire sentences into single embeddings which are then passed through attention layers. This could act as a filtering mechanism to determine which sentences should be focused on.

## References

- Alm, C. O. (2011). Subjective natural language problems: Motivations, applications, characterizations, and implications. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, 107–112.
- Borth, D., Ji, R., Chen, T., Breuel, T., & Chang, S.-F. (2013). Large-scale visual sentiment ontology and detectors using adjective noun pairs. *Proceedings of the 21st ACM International Conference on Multimedia*, 223–232.
- Datta, R., Joshi, D., Li, J., & Wang, J. Z. (2006). Studying aesthetics in photographic images using a computational approach. *European Conference on Computer Vision*, 288–301. [https://doi.org/10.1007/11744078\\_23](https://doi.org/10.1007/11744078_23)
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *ArXiv Preprint ArXiv:1810.04805*.
- Gao, D., Jin, L., Chen, B., Qiu, M., Li, P., Wei, Y., Hu, Y., & Wang, H. (2020). Fashionbert: Text and image matching with adaptive loss for cross-modal retrieval. *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2251–2260.
- Gygli, M., Grabner, H., Riemenschneider, H., Nater, F., & Van Gool, L. (2013). The interestingness of images. *Proceedings of the IEEE International Conference on Computer Vision*, 1633–1640.
- Isola, P., Xiao, J., Torralba, A., & Oliva, A. (2011). What makes an image memorable? *CVPR 2011*, 145–152.
- Kovashka, A., & Grauman, K. (2013). Attribute adaptation for personalized image search. *Proceedings of the IEEE International Conference on Computer Vision*, 3432–3439.
- Kovashka, A., Parikh, D., & Grauman, K. (2012, June). WhittleSearch: Image Search with Relative Attribute Feedback. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
- Kumar, N., Berg, A., Belhumeur, P. N., & Nayar, S. (2011). Describable visual attributes for face verification and image search. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(10), 1962–1977.
- Ladron de Guevara, M., George, C., Gupta, A., Byrne, D., & Krishnamurti, R. (2020). *Multimodal Word Sense Disambiguation in Creative Practice*. <http://arxiv.org/abs/2007.07758>
- Li, C., & Chen, T. (2009). Aesthetic visual quality assessment of paintings. *IEEE Journal on Selected Topics in Signal Processing*, 3(2), 236–252. <https://doi.org/10.1109/JSTSP.2009.2015077>
- O'Donovan, P., Libeks, J., Agarwala, A., & Hertzmann, A. (2014). Exploratory Font Selection Using Crowdsourced Attributes. *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 33.
- Parikh, D., & Grauman, K. (2011). Relative attributes. *2011 International Conference on Computer Vision*, 503–510.
- Tao, L., Yuan, L., & Sun, J. (2009). Skyfinder: attribute-based sky image search. *ACM Transactions on Graphics (TOG)*, 28(3), 1–5.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Transformer: Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998–6008.

# RHETORIC, WRITING, AND ANEXACT ARCHITECTURE

*The Experiment of Natural Language Processing (NLP) and Computer Vision (CV) in Architectural Design*

YUXIN LIN

<sup>1</sup>*University of Michigan*

<sup>1</sup>*linyuxin@umich.edu, 0000-0002-3867-820X*

**Abstract.** This paper presents a novel language-driven and artificial intelligence-based architectural design method. This new method demonstrates the ability of neural networks to integrate the language of form through written texts and has the potential to interpret the texts into sustainable architecture under the topic of the coexistence between technologies and humans. The research merges natural language processing, computer vision, and human-machine interaction into a machine learning-to-design workflow. This article encompasses the following topics: 1) an experiment of rethinking writing in architecture through anexact form as rhetoric; 2) an integrative machine learning design method incorporating Generative Pre-trained Transformer 2 model and Attentional Generative Adversarial Networks for sustainable architectural production with unique spatial feeling; 3) a human-machine interaction framework for model generation and detailed design. The whole process is from inexact to exact, then finally anexact, and the key result is a proof-of-concept project: Anexact Building, a mixed-use building that promotes sustainability and multifunctionality under the theme of post-carbon. This paper is of value to the discipline since it applies current and up-to-date digital tools research into a practical project.

**Keywords.** Rhetoric and Writing; Natural Language Processing; Computer Vision; GPT-2; AttnGAN; Human-computer Interaction; SDG 3; SDG 11.

## 1. Introduction

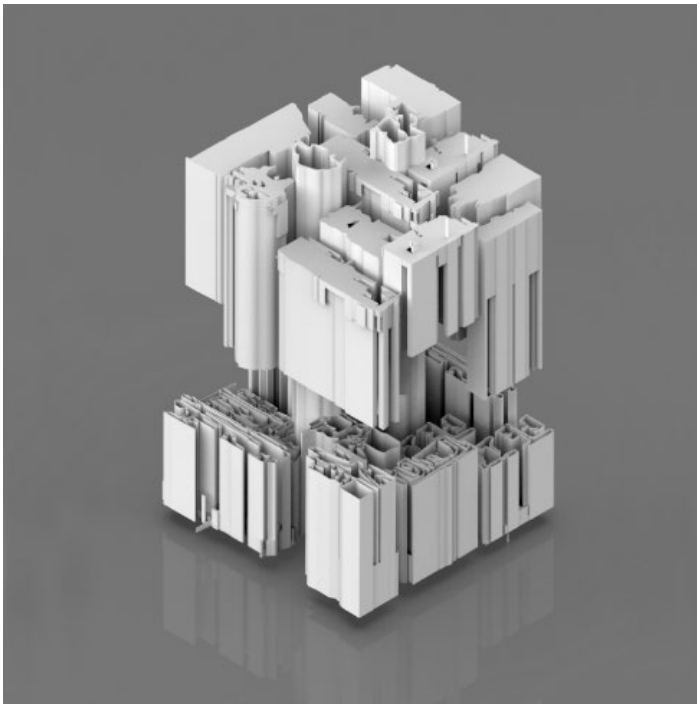
### 1.1. WRITING AND RHETORIC

Language has the power to create novel, generative architecture that could change the shape of the world, and writing aids the iterative process of creation (Stenson, 2017). The necessity of writing in architecture is mainly embodied in its ability to convey abstract feelings of architectural space imaginatively through words and effectively extend the aesthetics of architecture. Writers use various rhetorical devices to compose

narratives, and architects use forms that can be constructed with geometries to shape architecture and create space. ‘Geometry is the preferred language for architectural communication, and its interrogation has become the dominant form of writing in architecture’ (Lynn, 1998, p. 79). For architects, geometry or form is rhetoric in architectural writing. By writing down words to describe designs in spatial terms, architects can use geometry rhetoric as a strategy that connects different spatial feelings and functional needs to form a complete architecture.

## 1.2. ANEXACT FORM

Just as writers use different rhetorical devices, architects’ understanding of geometric rhetoric changes from time to time. Recently, Greg Lynn used ‘anexact’ to establish a new geometric understanding of architecture (Lynn, 1998). Anexact forms can be described with local precision yet cannot be wholly reduced, so they differ from exact forms that are precise and can be reduced eidetically or inexact forms that cannot be fixed or reduced because their contours cannot be described (Derrida, 1989). In Lynn’s view, anexact form may supply architecture with the ability to measure amorphousness and undecidability, creating a new model of architectural practice. Figure 1 shows that Project Anexact Building uses anexact form to detail the feeling of space visually and fix it into an architectural expression from the new perspective of artificial intelligence (AI) and synthesize the theory of form and space using rhetoric devices.



*Figure 1. The conceptual model of 'Anexact Building' consists of anexact forms. (Image: Yuxin Lin, Yuchun Huang, Taubman College Thesis 2021)*

1.3. MACHINE LOGIC

‘The language is a universalizing concept meant to explain every single act of building in the world’ (Alexander, 1978, p. 192). Christopher Alexander here sees language as a structuring system that gives each person who uses it the power to create an infinite variety of new and unique buildings (Alexander, 1978), and there must be a sense of order inside language. Deep Learning for Natural Language Process (NLP) was first introduced in 2012. Moreover, research proved that machine learning (ML) methods in NLP can achieve state-of-the-art results in many natural language tasks, such as learning and recombining patterns of association between several parts within text datasets (Goldberg, 2016; Goodfellow et al., 2016). ‘The more we understand the complex nature of form and the complex nature of the function, the more we shall have to seek the help of the computer when we set out to create form’ (Christopher, 1964, p. 54). ML in computer vision helps to give computers human-like capabilities to sense data, understand data, and act based on instances and experiences (Khan and Al-Habsi, 2020). The ML model makes it possible to recognize and detect objects in images then label them using text. Conversely, identifying objects in descriptive texts and seeking the relations between them to do text-to-image translation is also becoming feasible.

1.4. POST-CARBON THEME

Population growth and sustainability goals require us to seek effective ways to address the reality of high-density cities, as well as prepare for profound future challenges. Corresponding to the post-carbon framework, anexact forms with the ability to measure amorphousness and undecidability mentioned above can describe and organize spaces across sustainability and multifunctionality, helping architecture increase its resilience and respond to the internal and external environment effectively. The application of AI makes it possible to evaluate and understand the trends and consequences of sustainable correlations through big data, leading to the production of post-carbon architecture.

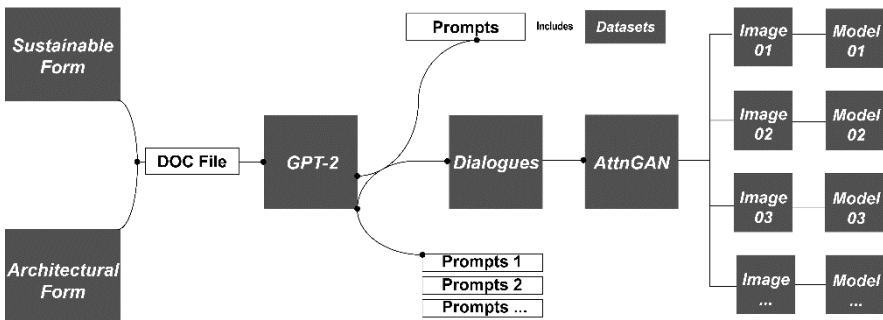


Figure 2. Workflow. (Image: Yuxin Lin, Yuchun Huang, Taubman College Thesis 2021)

2. Method

Figure 2 shows the whole process step by step. The NLP process takes the first step to write descriptive texts based on form language. Then the generated texts will be put

into a text-to-image translation algorithm. Here CV helps to visualize form rhetoric within the texts into 2D images. The third step is to extract the images by geometric recognition and do 2D-to-3D modeling through human-machine collaboration.

### 2.1. ALGORITHMIC MODEL

The selection of neural networks that can achieve NLP and image-based translation is critical in this article. Previous research about NLP succeeded in using the Generative Pre-trained Transformer 2 (GPT-2) to generate synthetic text samples in response to arbitrary input (Radford et al., 2019). 'Attentional Generative Adversarial Network (AttnGAN) allows attention-driven, multi-stage refinement for fine-grained text-to-image generation' (Xu et al., 2018, p. 1). Microsoft COCO is a new dataset with the goal of object recognition in the context of scene understanding (Lin et al., 2014). In the field of architecture, the application of the AttnGAN with the COCO dataset can serve as a successful design method for a complex architecture program by translating written form into visual output and interrogating shape through language (del Campo, 2021; Yuan et al., 2019). In their research, the algorithm helps visualize the dialogues from GPT-2 so every sentence can be related to an image containing complex shapes and forms that imply space. Instead of manually writing the detailed description of the whole space then translating each sentence into architectural components separately, the combination of GPT-2 and AttnGAN allows the inputs of fragments. Then it can predict the next word and generate coherent paragraphs of text by looking for inter-relationships and connecting the same symbols to organize the compositions through images.

### 2.2. DATASET

The content and structure of the dataset are under consideration since the ML process depends highly on a valid dataset. For this paper, dataset gathering focused on all the factors to a post-carbon future by articulating specific form language datasets, including A) architectural form and B) sustainable form. There are 4,720 sentences for database A describing the space configuration and spatial organization of different parts in architecture in a format that organizes all the parts through relational patterns. Database B contains 2,250 sentences that explain spatial form in terms of natural sustainability and multifunctionality. For example, unique forms of termite nests enable the whole structure to maintain a stable internal temperature and humidity through natural ventilation in the face of changing external environmental conditions. Mixing sustainable form from natural structures with architectural form through words is a new vision to use to think of sustainable design methods, and a total of 7,000 sentences is enough to fine-tune the GPT-2 model to achieve the desired result.

### 2.3. HUMAN-MACHINE COLLABORATION

Generating shape is a bottom-up process; it needs to collaborate with a top-down process where we make decisions as architects. Design is a multidimensional process, and it makes possible the emergence of intellectual activities which consist of a combination of creativity and perception (Kahveciođlu, 2007; Oxman, 2006). When it comes to multiple variables, AI can now help link these factors point-to-point in a linear

and potentially interlinked way. This paper does not see AI as a fully automated tool but as a catalyst to stimulate intuitive human thinking. So human-machine collaboration is inevitable and necessary to turn unexpected outputs into conceptualized designs.

### 2.4. WORKFLOW

The first step was to fine-tune the GPT-2 model with the collection of two text datasets describing architecture form and sustainable form in DOC files. After training, the model generated coherent and continuous descriptions of space by looking for inter-relationships to organize the compositions, merging two forms. Then system prompts which refer to the content in the dataset aid to trigger dialogue generation. For example:

- Prompt 1: when arranging living groups, ventilation is required in our pursuit of comfort and should be designed in a more sustainable way.

*Prompt 2: A vibrant community must be composed of different households. Small spaces are the small elements that make up the entire architecture.*



Figure 3. GPT-2 model and AttnGAN algorithm results. (Image: Yuxin Lin, Yuchun Huang, Taubman College Thesis 2021)

Figure 3 shows the detailed results generated by GPT-2: texts in bold italic come from the architectural form dataset, and white texts come from the sustainable form dataset, proving that the algorithm combines two datasets and generates realistic and coherent continuations or descriptions about one specific design problem.

The next step was to interpret or visualize output sentences generated by the GPT-2 model as input to the AttnGAN. The algorithm visualizes the generated dialogues from GPT-2 by permutating and combining shapes and forms that imply space. Figure 3 also shows that the AttnGAN was used several times to get multiple outcomes by switching prompts, and similar features were selected to identify forms such as walls, rooms, and gardens.

The final step was to transfer 2D pictures into 3D models. The output features are perspective images that get extruded into 3D assembly unit models, rationalized by the architect as a building design. This prospective form-finding pipeline with a taste of collaborative intelligence resulted in a specific inexact shape recognizable as exactly geometric lines in Figure 4. Then human intelligence continued to piece together the blobs into the final building, guided by the functionality of living, working, and entertaining spaces as well as our understanding of sustainable architecture.

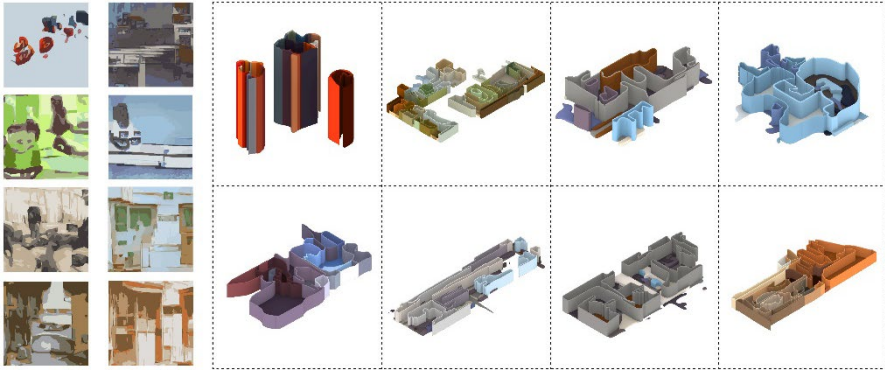


Figure 4. 3D assembly models. (Image: Yuxin Lin, Yuchun Huang, Taubman College Thesis 2021)

### 3. Result

The results of this research are visually presented as a complete design project called Anexact Building. We rely on the detailed architectural drawing of plans, sections, and renderings to describe how AI helped to seek unexpected opportunities and reassemble them into innovative and sustainable designs. We have proved that this method can be practically used by implementing a novel design method into an actual design project. The project is located close to the Roosevelt Loop along the Chicago River, and we aim to build a liveable and sustainable new community. This project, which serves as a mixed-use building, will rethink how to promote the growth and sustainability of a community towards a post-carbon future.

#### 3.1. ARCHITECTURAL DRAWING

Figure 5 represents the four typical plans of the building. The first plan is at the bottom of the building, offering retail and shopping areas for public use. The second plan is at the hotel lobby level along with different office buildings. We start to see some roof gardens that can be entered from different levels on this plan. The third plan represents the leisure and cultural space in the middle, transitioning from more public to more private. The fourth plan shows the residential area at the top. The surrounding apartments and the recreation areas in the middle are mainly grouped by massing blocks that have anexact boundaries. Diverse functional parts with different heights are distinguished both by layer and mass. In general, each part is relatively independent but works together.



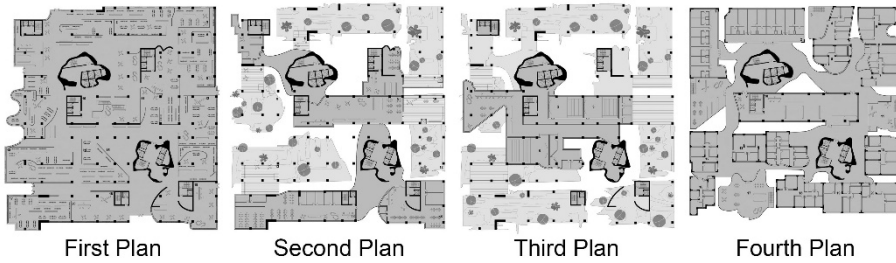


Figure 5. Plans. (Image: Yuxin Lin, Yuchun Huang, Taubman College Thesis 2021)

The primary language for the facade is vertical lines created from concrete and glass. The colour changes were turned into the blocks and kept as the ceiling or floor patterns. Figure 6 (Left) is the view from the middle part of the building, showing the roof garden, city balcony, and the hotel lobby, referring to the AI-generated sentence:

- At larger scales, structures become special places which invite people to come together and feel closer in a communal space.

Figure 6 (Right) represents the view from the rooftop showing all the gardens generated out of the colour patterns responding to the AI-generated sentence:

- Make each house a distinct piece of open land, with a garden at its front and back, and perhaps a small garden at its side.



Figure 6. Renderings. (Image: Yuxin Lin, Yuchun Huang, Taubman College Thesis 2021)

### 3.2. SUSTAINABILITY

Sustainability in this project includes 1) ecologically sustainable design that emphasizes environmental consciousness in energy and natural resources, and 2) socioculturally sustainable design that improves the quality of life and enhances comfortable living conditions (Yavasbatmaz and Gultekin, 2013).

#### 3.2.1. Ecologically sustainable design

Figure 7 shows that the overall shape successfully brings in the passive stack effect to promote the resilient and sustainable design. In Chicago, generally cool conditions in

winter require interior heating, and the vertical stack effect will help reduce some of the heating usages. Also, natural ventilation will help reduce the usage of the mechanical cooling system in moderately warm conditions during summer. Two core areas that work as the circulation systems ensure overall system efficiency, and their massive concrete structures can store natural heat to reduce energy consumption.

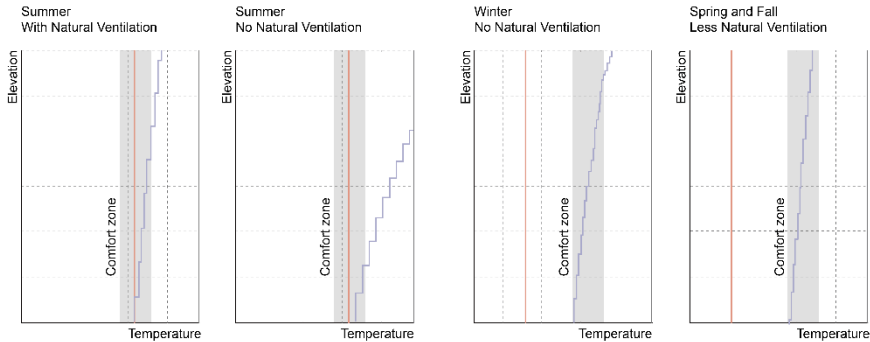


Figure 7. Stack effect and natural ventilation diagram. Summer: Reduce the interior temperature by drawing outside air and delay the use of air conditioning. Winter, spring, and fall: Reduce the difference between the actual interior temperature and the comfort set points to control heating or cooling usage.

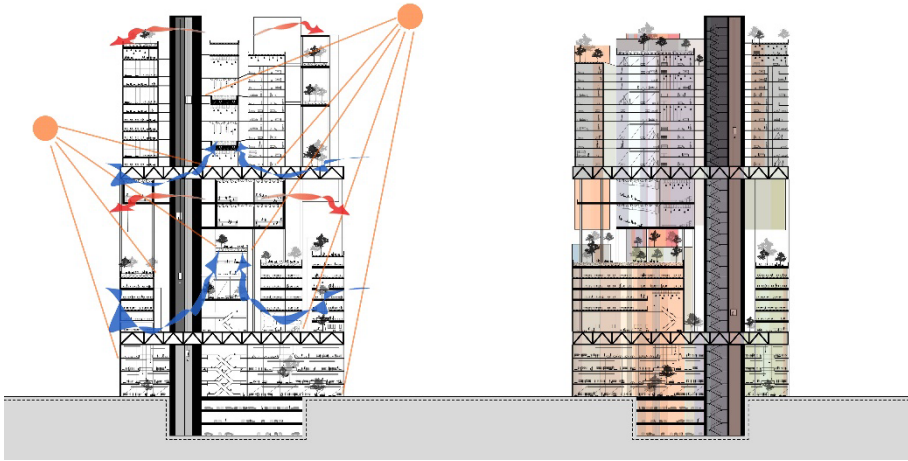


Figure 8. Introduce daylight, nature air, and views into the building.

### 3.2.2. Socioculturally sustainable design

Figure 8 shows that there is a glass room on each side of this building to ensure sunlight, and the distance between different groups enables natural ventilation as well as the maximize daylight penetration. Roof gardens with plants inspired by the algorithm create a high-quality living environment by keeping the users in touch with the outside space, and residents can have communal activities here. To some degree, the colours from the images will be kept inside to retain the original design of the building and

bring a unique aesthetic feeling through visual expression. The logic to provide the building occupants with a connection between indoor and outdoor space by introducing daylight, nature air, and views into regularly occupied areas can improve the health and wellbeing of all community members.

#### **4. Discussion**

The output of Anexact Building is an experiment of combining texts, visualizing machine logic, and artificial intelligence. This research succeeds in 1) proving that AI can synthesize the sense of order inside the architectural language that describes different spatial forms, 2) giving architecture the ability to measure and visualize uncertain, heterogeneous writing characteristics through form rhetoric, and 3) integrating anexact aesthetics with indeterminacy in sustainability computation.

In the context of anexact, the AI-generated description with translated images brings architecture the possibility to be constantly manipulated and interpreted by different viewers, continually generating new meanings. In this process, architecture, like writing, has the characteristic of inexactitude. AI image processing generates 3D assembly models that are then made into the entire building by human interpretation, ensuring exact and rigorous methods using reasonable functional partition and integration. This process is precise and can be induced eidetically based on the design requirements. The result obtained by refining these two processes is anexact that is recognizable as precisely functional units yet cannot be wholly reduced to a style.

For post-carbon, the form can lead to a sustainable design with the assistance of our understanding of sustainable architecture. Human agency here helps to enhance sustainability in the final form of the building. However, this cognition does not question or undermine AI's agency and authorship in sustainable design but emphasizes its creativity. We see AI as a tool to speculate on the conventional method with the thinking in a different field when faced with a similar problem. The role of AI here does not work as automation but as automatism (Schmitt and Weiß, 2018) to stimulate human imagination.

#### **5. Conclusion**

In the 'second digital turn,' which Mario Carpo defines as designers now seeing new kinds of digital tools as tools for thinking (Carpo, 2017), new ways of thinking derived from new technologies can result in new styles and aesthetics. AI brings to the architectural field a new design thinking that shifts from integrating all parts of a design problem in advance to breaking them apart along natural lines to better understand and recombine the parts during the process. In this project, AI provides a way to organize a sustainable system by composing a set of interacting elements within. From text to image as a sustainable form-finding logic, it is a combination of theoretical, historical, and technical conversations within the discipline.

The next generation of the algorithm should focus on translating input prompts into a series of scene images consisting of forms directly, and image discrimination can be automated by semantic segmentation or object detection. In the future, AI could even generate 3D models automatically based on arbitrary text inputs and conduct quantitative calculations on energy consumption simultaneously.

## Acknowledgements

This research was part of the thesis studio “Visionary machine”, instructed by Prof. Dr. Matias del Campo, at the University of Michigan Taubman College. I would like to acknowledge Yuchun Huang for her contributions to this project. In addition, thanks to Elsa Mathew, Danish Sayed, Janpreet Singh and Alexandra Carlson for providing technical support in coding and AI training.

## References

- Alexander, C. (1978). *Timeless way of building*. Cary, NC: Oxford University Press.
- Alexander, C. (1964). ‘A Much Asked Question about Computers and Design,’ in *Architecture and the Computer*, 52.
- Carpo, M. (2017). *The second digital turn: Design beyond intelligence*. London, England: MIT Press.
- del Campo, M. (2021). Architecture, Language and AI-Language, Attentional Generative Adversarial Networks (AttnGAN) and Architecture Design.
- Derrida, J. (1989). *Edmund Husserl’s origin of geometry: An introduction*. U of Nebraska Press.
- Goldberg, Y. (2016). A primer on neural network models for natural language processing. *The Journal of Artificial Intelligence Research*, 57, 345–420.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. London, England: MIT Press.
- Kahveciođlu, N.P. (2007). Architectural design studio organization and creativity.
- Khan, A. I., & Al-Habsi, S. (2020). Machine learning in computer vision. *Procedia Computer Science*, 167, 1444–1451.
- Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L. (2014). Microsoft COCO: Common objects in context. In *Computer Vision – ECCV 2014* (pp. 740–755). Cham: Springer International Publishing.
- Lynn, G. (1998). *Folds, bodies & blobs: collected essays*. Books-by-architects. [Bruxelles]: La Lettre volée.
- Oxman, R. (2006). Theory and design in the first digital age. *Design Studies*, 27(3), 229–265.
- Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language Models are Unsupervised Multitask Learners.
- Schmitt, P., & Weiß, S. (2018). The Chair Project: A Case-Study for using Generative Machine Learning as Automatism.
- Stenson, M. W. (2017). *Architectural intelligence: How designers and architects created the digital landscape*. The MIT Press.
- Xu, T., Zhang, P., Huang, Q., Zhang, H., Gan, Z., Huang, X., & He, X. (2018). AttnGAN: Fine-grained text to image generation with attentional generative adversarial networks. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE*.
- Yavasbatmaz, S., Gultekin, A. B. (2013) Sustainable design of tall buildings, *GRAĐEVINAR*, 65 (5), 449-461. <https://doi.org/10.14256/JCE.772.2012>
- Yuan, Y., Chen, Y., Liu, Z. (2019). *Architecture of {AI} Language*. Unpublished master’s thesis, Taubman College of Architecture and Urban Planning, University of Michigan, Ann Arbor, Michigan.

# IS LANGUAGE ALL WE NEED? A QUERY INTO ARCHITECTURAL SEMANTICS USING A MULTIMODAL GENERATIVE WORKFLOW

DANIEL BOLOJAN<sup>1</sup>, EMMANOUIL VERMISSO<sup>2</sup> and SHERMEEN YOUSIF<sup>3</sup>

<sup>1,2,3</sup>*Florida Atlantic University.*

<sup>1</sup>*dbolojan@fau.edu, 0000-0003-2060-367X*

<sup>2</sup>*evermiss@fau.edu, 0000-0001-9116-8877*

<sup>3</sup>*syousif@fau.edu, 0000-0003-4214-1023*

**Abstract.** This project examines how interconnected artificial intelligence (AI)-assisted workflows can address the limitations of current language-based models and streamline machine-vision related tasks for architectural design. A precise relationship between text and visual feature representation is problematic and can lead to “ambiguity” in the interpretation of the morphological/tectonic complexity of a building. Textual representation of a design concept only addresses spatial complexity in a reductionist way, since the outcome of the design process is co-dependent on multiple interrelated systems, according to systems theory (Alexander 1968). We propose herewith a process of feature disentanglement (using low level features, i.e., composition) within an interconnected generative adversarial networks (GANs) workflow. The insertion of natural language models within the proposed workflow can help mitigate the semantic distance between different domains and guide the encoding of semantic information throughout a domain transfer process.

**Keywords.** Neural Language Models; GAN; Domain Transfer; Design Agency; Semantic Encoding; SDG 9.

## 1. Introduction

Artificial intelligence (AI) models that focus on language and vision have progressed exponentially, featuring recent neural language models (NLMs) like DALL-E and CLIP (2021). State-of-the-art literature has speculated that this progression in computer vision is only the beginning of introducing NLMs’ potential for design. Such neural networks (ANNs) match text prompts with corresponding visual representation, helping circumvent computationally expensive and labor-intensive tasks involved in machine vision. These models can expedite and streamline machine-vision related tasks for architectural design, due to expanded, generic training datasets. However, a precise relationship between text and visual feature representation is problematic and leads to “ambiguity” in the interpretation of the morphological/tectonic complexity of architecture. For example, a building’s phenomenological aspects are especially

difficult to capture and concisely encode through natural language supervision.

Design is likely the most sophisticated aspect of human intelligence (Gero, 1991). It is a multimodal endeavour, where different design tasks can be addressed at various degrees of abstraction through mediums like text, drawing, painting, 3D models, etc. Direct application of AI models where one model tackles the overall design complexity is problematic. Textual encoding of a design concept only allows a reductionist understanding of spatial complexity, since the outcome of the architectural process is co-dependent on multiple interrelated systems (Alexander, 1968). Consequently, applying NLMs into the architectural domain, offers an inadequate interpretation and handling of the various layers entangled in architectural design. In response, this work investigates the use of NLMs into a design workflow, addressing aspects of: (1) oversimplification of design complexity and flattening architectural layers when applying NLMs discreetly, (2) models' ambiguity, particularly for training with non-architectural datasets, and (3) designers' agency-control within AI-assisted workflows.

In order to employ AI (i.e., NLMs) to tackle design complexity and the multiple interrelated architectural layers, we propose a workflow to disentangle the process into its interconnected component parts. Disentanglement into architectural layers or design tasks (i.e., organization, composition, structure) breaks down the architectural problem, allowing other AI models to deal with discrete layers. This way, the suggested design system supports feature disentanglement and accommodates design complexity. The NLMs -DALL-E, VQGAN+CLIP, Diffusion models (DMs)- used herewith were pre-trained with a generic dataset. Additional models (i.e., StyleGAN; CycleGAN) can help clarify the architectural domain, substituting for other pre-trained models. A protocol of two experiments was herewith introduced: (1) Discrete networks: three NLMs (DALL-E, VQGAN+CLIP, and DMs) were separately used to test a design scenario and, (2) Interconnected networks: the NLM became a component into a design workflow with multiple interconnected AI models (GANs, NLMs), where architectural layers were disentangled, incorporating domain-specific datasets. While the first experiment demonstrated limited agency over the process, the second experiment showed how NLMs can increase human agency, enabling the encoding of design intentionality at various design task and process levels. Our primary contribution to Creative AI research features the development of a hybrid workflow of GANs-NLMs that compensates for the latter's limitations by using domain-specific architectural data. This treated a complex layered problem in separate stages, leveraging a less expensive computational strategy (NLMs) when incorporating it into an interconnected workflow of multiple deep learning models. In light of a need to upgrade design processes for the built environment, our adoption of state-of-the-art technology (AI) in architectural design aimed to foster innovation in both research and pedagogy, augmenting human decision making in the AEC industry sector (in line with UN SDG 9).

## 2. Background of neural language models and State-of-the-Art

Neural Language models (NLM) were first proposed in 2001. (Mansimov et al., 2016) pioneered modern machine learning techniques for text-to-image synthesis by demonstrating that the DRAW generative model (Gregor et al., 2015) could also produce unique visual scenarios when expanded to condition on image captions. Later, Reed et al. (2016) showed that GANs (Goodfellow et al., 2014) increased picture

quality in place of recurrent variational auto-encoders (Ramesh et al., 2021). Vaswani et al. introduced transformers as a novel Natural Language Processing (NLP) architecture (2017), before their adoption in computer vision (McCann et al., 2017).

## 2.1. AI MODELS: DALL-E, VQGAN+CLIP, DIFFUSION MODELS.

The NLMs used here rely on the *CLIP* classification network, a zero-shot ANN trained on 400 million “text-image” pairs. *DALL-E* (1) is a CLIP-assisted generative network with a 12-billion parameter version of GPT-3 that creates images from text descriptions of a broad range of natural language topics. Trained on a dataset of text–image pairs, DALL-E can anthropomorphize animals and objects, combine unrelated concepts in believable ways, generate text, and apply alterations to existing images (Ramesh et al. 2021). *VQGAN+CLIP* (2) builds on earlier attempts to leverage the generative capacity of GAN architecture for NLP models like CLIP. The VQGAN is guided by CLIP, which encodes a text prompt and searches the correlation between text-image output, until a perfect match between image and text is identified. *CLIP-Guided Diffusion Models* (3) generate images from text by connecting a class-conditional ImageNet diffusion model (Dhariwal et al., 2021) to a CLIP model. DMs are probability-based models with the potential for creating high-quality images, while demonstrating characteristics such as easy scaling, distribution coverage, and stable training.

## 2.2. STATE-OF-THE-ART

NLMs have lately yielded a new generation of AI-based art and design synthetic imagery. Since NLM models were only released recently, their application in architectural research remains experimental. In one case, the work of (Yang & Buehler, 2021) used CLIP+VQGAN to produce images from text prompts and subsequently generate three-dimensional models for manufacturing. The work of Theodoros Galanos on training DALL-E with 150k floor plan layouts represents an early attempt to use text for generating design layouts (2021). In (Rodrigues et al.)’s work, VQGAN+CLIP was employed to offer design inspiration, using text prompts in literal (analogy) and abstract (metaphorical) ways, combined with photographs and sketches (2021). NLMs demonstrate potential and challenges; as seen in Galanos’ work, solving spatial relationships in floor plans remains difficult challenging. Huang et al. 's work attempts to address the subtleties of distinct GAN models (SAGAN vs DCGAN) for developing new designs. Here, NLPs are limited to conceptual interpretation of findings (2021), without extending the discussion to the generative role of NLMs.

## 3. Methods and Design Experiments

Adopting John Gero’s schema of ‘design prototypes’ (1990), a new approach to architectural design was pursued, employing interconnected deep learning models. To demonstrate the system and test its functionalities, a prototype was formulated through a series of experiments. Experiment 1, which is regarded as the ‘routine’ process (Gero, 1990), applied discrete NLMs to design scenarios. Experiment 2, the ‘innovative’ process - interconnected multiple DL models with NLMs into a “hybrid” workflow that seems to emulate the ‘multimodal’ thinking structure in human perception.

3.1. DESIGN EXPERIMENT 1: DISCRETE NEURAL NETWORKS

Experiment 1 employed a workflow to showcase opportunities and limitations present in NLMs like DALL-E, VQGAN/CLIP and DMs, when these are used in isolation. Various tests were conducted using a series of text prompts, ranging from archetypal semantic descriptions like “Building”, “Plan”, “Section”, “Façade” to more specific like “A modernist glass façade”, or “Building-Gaudi-Sagrada Familia-interior”.

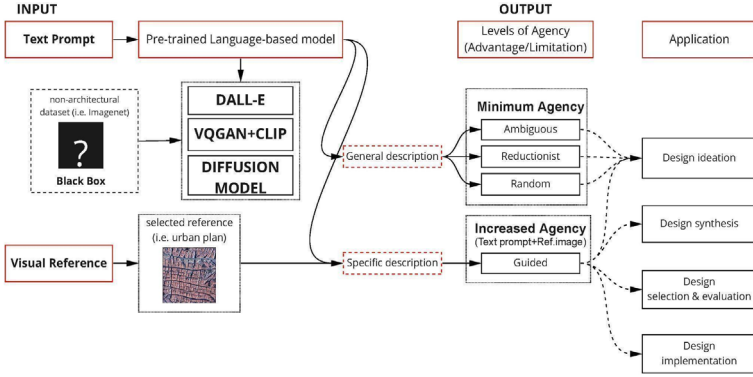


Figure 1: Design Experiment 1 (reductionist approach) workflow involves discrete neural language models, which rely on text prompts alone or text prompts and image reference.

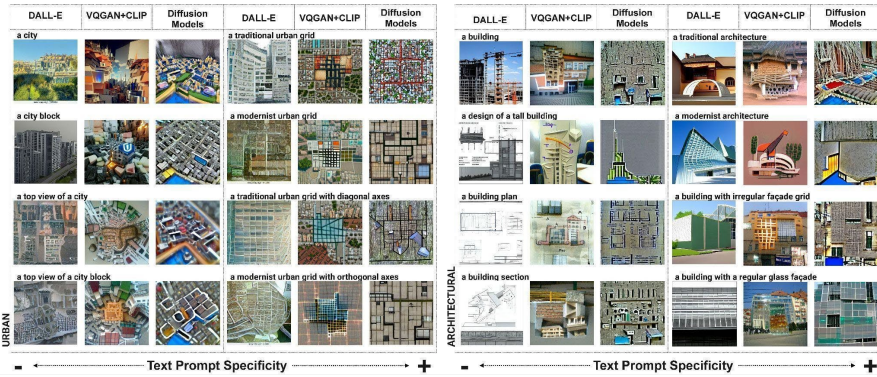


Figure 2: Experiment 1 provided three discrete NLMs with progressively more specific text prompts. Horizontal rows compare results given the same text prompt, while vertically showing the output of each model, given different text prompts relative to urban (left) and architectural (right) scales.

These trials were also attempted with additional visual references. Figure 1 demonstrates that the input-output sequence in NLMs relies primarily on text information, and the agency available to the designer is dependent on the curation of the text prompt (i.e., text length/specificity). We have argued here that this operational process, while potentially interesting, lacks the affordance of an overall flexibility to the user/designer. Our work with such models supports those findings and further provides observations in the following subsections.



### 3.1.1. Observations on Prompt Specificity in Experiment 1

As observed (Fig.2), specificity worked well in all models, up to a certain level while beyond a point, models like DALL-E performed well with less specific text prompts like “a city block”, while more specific descriptions like “a traditional urban grid” did not necessarily improve results. Other models performed better when text prompt specificity was higher; they failed to recognize more generic prompts such as "a building plan" (VQGAN+CLIP; DMs) "a building facade", and "a building section" (DMs) resulting in low resolution, undefined results. Semantic features described by the prompt can still be observed in the image outcomes, but they are vague. Some models seemed to output satisfactory results when prompts directed the search towards certain (more specific) semantic categories, i.e. “A Baroque facade”. This might depend on how well semantic domains in the text prompts, were represented in the original training dataset (i.e., Imagenet). When prompt specificity increased by additional descriptors (“A building with a colonnade in the Baroque style of Francesco Borromini”), the low level features (i.e., overall legibility of an architectural style) satisfied the requirement for “Baroque” characteristics, without fully or successfully resolving the high-level features (i.e., details like windows, colonnade). This was also true with more archetypical descriptors (i.e., "top view of a city"), allowing reading of features as a whole but with low resolution. Sometimes, in spite of acceptable visual outcomes, output variety suffered, leading to similar results. We know that these models were pre-trained on large datasets which are comprehensive, and therefore, not domain specific. Using general text descriptions alone (i.e., a building, a city) risks “flattening” individual architectural layers, leading to reductionist, ambiguous and random results, as the network relies on prior supervised learning from broad datasets.

### 3.1.2. Observations on additional references & hyperparameters in Experiment 1

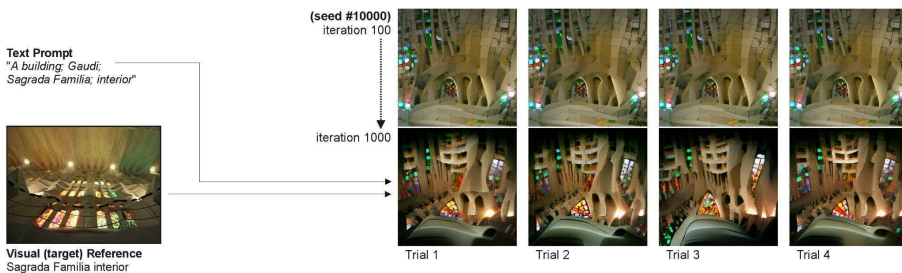


Figure 3: VQGAN+CLIP trials with prompt: “A building; Gaudi; Sagrada Familia; interior” (Seed #10,000; reference image). The 100th and 1000th iterations illustrate high structural similarities.

Fine-tuning hyperparameters like "seed selection" (DALL-E; VQGAN+CLIP) and “clip guidance scale” (DMs) settings in the NLMs enabled more agency in the process. Seed manipulation allowed access to different areas in the model’s latent space with possibly localized visual properties. In DALL-E’s case, modifying the hyperparameter did not affect the results’ success, but simply offered another comparable result. In the case of VQGAN+CLIP, different seeds were observed to output slightly different results of the same resolution. Nevertheless, processing the same seed while

keeping the same text prompt returned almost identical results, regardless of the number of repetitions (Fig.3). An alternative approach with certain NLMs (VQGAN/CLIP; DMs) allowed complementing the text prompts with visual references in the form of “source” or “target” images when possible. The DALL-E version used was “DALL-E mini”, and it did not accept reference images as input. While a visual reference slightly increased the user’s agency (more types of input) and guided the process (Fig.3), specific feature preservation was difficult to guide.

### 3.1.3. General Observations in Experiment 1

Once the prompt specificity was increased to very detailed (a paragraph, Fig.4), DALL-E’s performance featured more realistic qualities (similar to the earlier urban and architectural categories in Fig.2) than the “artistic” representations observed in the other two models (VQGAN+CLIP, DMs). VQGAN+CLIP indicated improved contextual representation of results, with the paragraph description, while adding a visual reference did not influence the compositional structure significantly, as was the case with DMs. Although results from the tested NLMs usually belong to the correct domain (i.e., building; façade; sculpture), they were not necessarily innovative, nor could they spark a kind of inspirational thinking, as they remain generic (Fig.2).

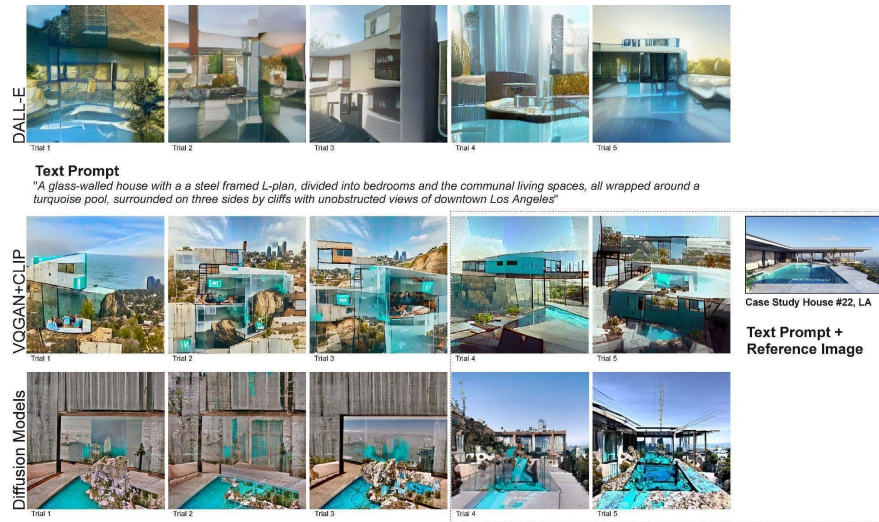


Figure 4: Several trials were performed for each NLM (DALL-E; VQGAN+CLIP; DMs), increasing text prompt specificity to a full paragraph. Reference images were used with VQGAN/CLIP; DMs.

Any localized specificity to a particular domain dataset (i.e., Architecture) is unknown, so the latent space of such models remains a “black box” to their users, making it difficult for designers to develop a sensibility to overcome the models’ limitations. It also reveals the minimal agency which the designer has over the design process parameters. Generally, a difficulty was observed in selecting the right text prompts to describe meaning as this became more entangled. The integrated language classification model (CLIP) in all tested models was trained in a supervised way, using

pairs of text-to-image, so it has assimilated particular correlations of verbal to visual data. It is therefore difficult to stir the models to specific resolution of low-level features without further ability to curate the process.

In Experiment 2, we proposed a more curated process, introducing ANN combinations (i.e., StyleGAN, CycleGAN, StyleGAN-NADA), to allow modification and controlled query into the latent space. Our intent was using this workflow to span longer “Semantic Distances”, thereby increasing the chance of creative solutions: “The role of semantic distance in creativity is intuitively embedded in the theory of creativity, through the notion that the farther one moves from a concept in a semantic space, the more novel or creative the new concept will be.” (Kenett, 2018).

### 3.2. DESIGN EXPERIMENT 2: INTERCONNECTED NEURAL NETWORKS

The experiment used an interconnected network of image-based models (StyleGAN, StyleGAN-nada), and NLMs (VQGAN, DMs) (Fig.5). As NLMs text to image translation relies on non-architectural specific domain training datasets, the architectural domain representation is constrained. Strategic curation of interconnected networks workflow improved representation of specific architectural domains at various scales and levels, yielding superior results to NLMs alone. Furthermore, domain-specific pre-trained networks enabled a more targeted and flexible conceptual space search. Designers typically work at various levels of abstraction, intermittently requiring specificity or ambiguity. Experiment 1 established that feature disentanglement, architectural layers disentanglement, and understanding interdependencies across multiple architectural layers is problematic when using NLMs alone. When inputting text, NLMs could not focus image generation on low level compositional structure, while simultaneously generating high level esthetical structure (Fig.2). As such, architectural levels and abstraction levels had to be specified, architectural domain-specific datasets generated, and task-specific models defined.

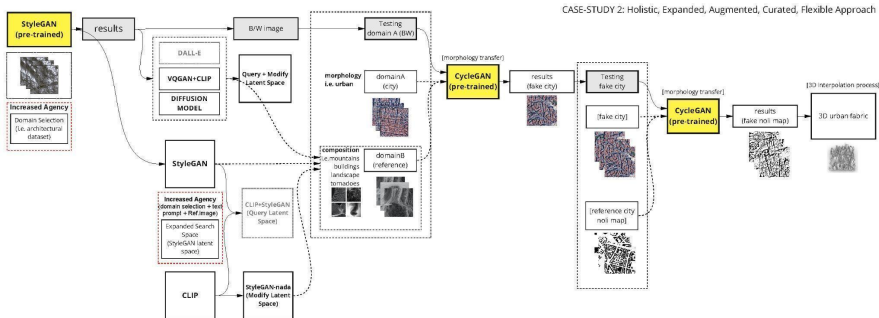


Figure 5: Experiment 2 workflow: interconnected ANNs and NLMs modify-query the latent space.

Furthermore, the experiment addressed levels of designer agency in the process, multimodality of design intention input, and meaningful disentanglement of architectural features. The phenomenological, semiological, tectonic, or organizational features of a building are particularly challenging and their concise encoding through language/text or NLMs results in a reductionist process, as different features necessitate different encoding mediums. As humans, we perceive the world in a variety of ways:

we hear noises, see things, smell odours, feel texture, and taste flavours. Design is also a multimodal endeavour where tasks of variable abstraction are addressed through various mediums (i.e., text, drawing, painting, 3D models). Relying on one modality alone to describe a concept is limiting. “Language is an imperfect, incomplete, and low-bandwidth serialization protocol for the internal data structures we call thoughts.” (LeCun, March 5, 2021) Certain concepts benefit from being described in text, while others benefit from being described in the form of a drawing/sketch, and still others from being described as a 3D model, and so forth. The experiment addressed the aspect of designer agency at multiple levels moving from low-level agency (i.e., input text prompt), to an increased agency (inserting reference images), by curating the dataset specifying the training domain (architectural data), to control the overall process by determining when to use NLMs (for example, for low level composition), and when to use generative DL models for other features (high level composition).

3.2.1. Observations on Experiment 2 Results: NLMs+GANs

To demonstrate an interconnected design approach, an architectural concept was formulated to generate an urban branching composition. For comparison reasons, two examples were generated; a first example where two text prompts (“a branching composition”, “an urban fabric with branching street compositions”) were defined for NLMs, and in a second example an interconnected process was defined. Outcomes of the first example (Fig.6) were constrained and ambiguous as aforementioned.

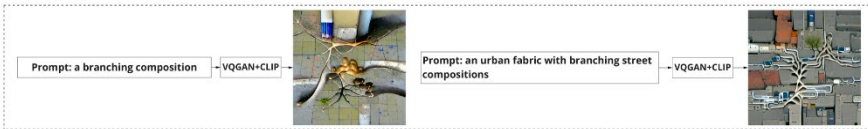


Figure 6: Experiment 1 process outcomes used as comparison with Experiment 2 workflow.

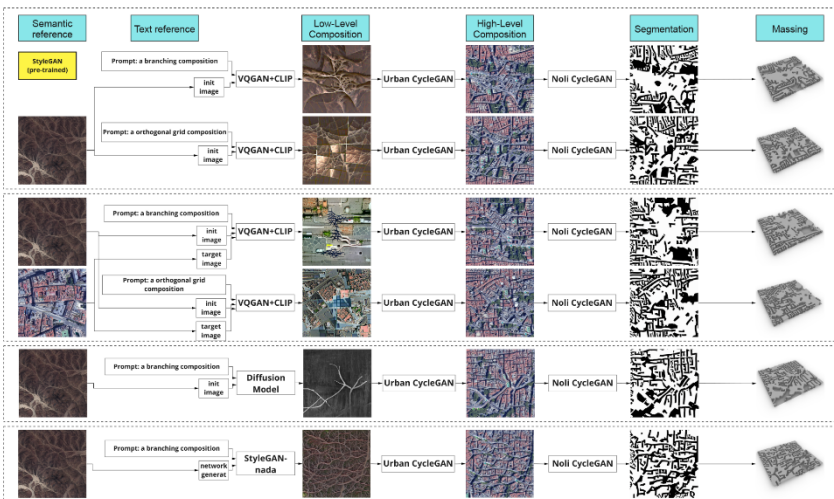


Figure 7: Experiment 2 integrates NLMs into an interconnected workflow to assist domain-transfer in GAN models by reducing semantic discrepancy across different domains.

The second example was divided into six discrete tasks, addressing aspects of agency, multimodality and architectural feature disentanglement. For each task (semantic reference, text reference, low and high-level composition, segmentation, massing), an AI model, modality, and relationships between the models were defined. A StyleGAN2-ada network (N1) learned compositional aspects of mountain topography to define the semantic reference level. The network N1 semantic features were then utilized to guide the NLM's (N2) generation. The aim of connecting the two networks (N1, N2) was to generate a low-level composition that is architecturally or urbanistically meaningful. Once this was defined, a CycleGAN image-based network (N3) was utilized to perform domain transfers between the low-level composition (mountain features) and high-level composition (Madrid urban fabric features).

Once the semantic and text reference was encoded, as well as the low-level and high-level compositional features, an additional network (N4) was used to perform segmentation of building blocks and free spaces. A Midas network (N5) was utilized to determine depth and building heights, informing the 3d massing. Within the sequence of other GAN models, the NLMs intervened to help clarify and guide the encoding of semantic information. For example, the branching mountain patterns (Fig.7) generated by the StyleGAN, belong to a very different domain than that of a typical urban morphology. Executing a domain transfer using the CycleGAN was assisted by the presence of the VQGAN+CLIP model, which reduced the semantic distance between the two domains, by acting as a mediating step.

#### **4. Conclusions and Future Work**

Working with image-based AI models for architectural design is hard; unlike other computational design workflows, designers' input relies primarily on visual data; this makes encoding design intentionality problematic because task complexity needs to be reduced to 2-dimensional data. In addition, current NLMs isolate the designer to one modality (verbal) which potentially exacerbates this reductionist interpretation, since they were also trained from 2-dimensional data in the first place. An important issue for applying AI to design, according to John Gero is "Representation in Design". What does a designer know that a computer doesn't? (Gero, 1991). A human can think in both verbal and non-verbal ways, unlike a computer. Several prominent thinkers have concluded that verbalization is not necessary for productive or even creative/innovative thinking. Non-Verbality manifests for example, as visual information, music, and mathematical language. The translation of "non-verbal" thinking into a "verbal" modality often severely constrains the meaning captured in the raw thinking process, as Roger Penrose recalls Albert Einstein's deliberations on his own thinking (Penrose, 1989). It is important to address this by offering complementary AI models to individual NLMs, so they circumvent limitations of verbal representation. This is important considering that such individual models are a priori constrained, because their training is dependent on a large, but generic dataset which cannot capture the complex nuances of architectural design. Specifically, the experiments performed demonstrated that increased specificity in individual NLMs (Experiment 1) did not produce visual outcomes which are useful to the designer. It is, therefore, necessary to maintain multimodality in design thinking, in line with the way human perception works. The proposed prototype (Experiment 2 multimodal workflow) enabled the

curation of visual references in the generative AI process, reaching beyond the limited search space of current NLMs. The proposed prototype offers the potential to support domain flexibility, by introducing other conceptual spaces, rather than working with individual NLMs, which are pre-trained in non-architecture specific domains. Furthermore, this line of work supports the UN SDG #9, pushing forward current technological capabilities of the AEC industry by offering innovative workflows. Moving forward, we are interested in situating such workflows into a pedagogical framework, exposing interconnected models to a broad student body.

## References

- Alexander, C. (1968). Systems generating systems. *Architectural Design*, 38, 605-610.
- Dhariwal, P., & Nichol, A. (2021). Diffusion models beat gans on image synthesis. *arXiv preprint arXiv 2105.05233*.
- Galanos, T. (2021). *DALL-E-pytorch*. Retrieved from <https://github.com/lucidrains/DALL-E-pytorch>
- Gero, J. S. (1990). Design prototypes: a knowledge representation schema for design. *AI magazine*, 11(4), 26-26.
- Gero, J. S. (1991). Ten problems for AI in design. *Workshop on AI in Design, IJCAI-91*.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., & Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, 27.
- Gregor, K., Danihelka, I., Graves, A., Rezende, D., & Wierstra, D. (2015). Draw: A recurrent neural network for image generation. *In the International Conference on Machine Learning*.
- Huang, J., Johanes, M., Kim, F. C., Doumpioti, C., & Holz, G.-C. (2021). On GANs, NLP and Architecture: Combining Human and Machine Intelligences for the Generation and Evaluation of Meaningful Designs. *Technology| Architecture+ Design*, 5(2), 207-224.
- Mansimov, E., Parisotto, E., Ba, J. L., & Salakhutdinov, R. (2015). Generating images from captions with attention. *arXiv preprint arXiv:1511.02793*.
- McCann, B., Bradbury, J., Xiong, C., & Socher, R. (2017). Learned in translation: Contextualized word vectors. *arXiv preprint arXiv:1708.00107*.
- Penrose, R. (1989). *The Emperor's New Mind*. Oxford: Oxford University Press.
- Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., & Clark, J. (2021). Learning transferable visual models from natural language supervision. *arXiv preprint arXiv:2103.00020*.
- Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C., Radford, A., & Sutskever, I. (2021). Zero-shot text-to-image generation. *arXiv preprint arXiv:2102.12092*.
- Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B., & Lee, H. (2016). Generative adversarial text to image synthesis. *In the International Conference on Machine Learning*.
- Rodrigues, R. C., Alzate-Martinez, F. A., Escobar, D., & Mistry, M. (2021). Rendering Conceptual Design Ideas with Artificial Intelligence: A Combinatory Framework of Text, Images and Sketches. *In the ACADIA 2021*.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., . . . Polosukhin, I. (2017). Attention is all you need. *In the Advances in neural information processing systems*.
- Yang, Z., & Buehler, M. J. (2021). Words to Matter: De novo Architected Materials Design Using Transformer Neural Networks. *Frontiers in Materials*, 8(417). doi:10.3389/fmats.2021.740754

# DEEP LEARNING-BASED SURROGATE MODELING FOR PERFORMANCE-DRIVEN GENERATIVE DESIGN SYSTEMS

SHERMEEN YOUSIF<sup>1</sup> and DANIEL BOLOJAN<sup>2</sup>

<sup>1,2</sup>*Florida Atlantic University.*

<sup>1</sup>*syousif@fau.edu, 0000-0003-4214-1023*

<sup>2</sup>*dbolojan@fau.edu, 0000-0003-2060-367X*

**Abstract.** Within the context of recent research to augment the design process with artificial intelligence (AI), this work contributes by introducing a new method. The proposed method automates the design environmental performance evaluation by developing a deep learning-based surrogate model to inform the early design stages. The project is aimed at automating performative design aspects, enabling designers to focus on creative design space exploration while retrieving real-time predictions of environmental metrics of evolving design options in generative systems. This shift from a simulation-based to a prediction-based approach liberates designers from having to conduct simulation and optimization procedures and allows for their native design abilities to be augmented. When introduced into design systems, AI strategies can improve existing protocols, and enable attaining environmentally conscious designs and achieve UN Sustainable Development Goal 11.

**Keywords.** Deep Learning; Artificial Intelligence; Surrogate Modeling; Automating Building Performance Simulation; Generative Design Systems; SDG 11.

## 1. Introduction

Generative design has already been established as a commonly used method to computationally generate and examine a large set of designs and often lends itself to designers' exploration and aesthetic evaluation. Building performance simulation (BPS) is an established method, required to target building design resiliency and adaptability to rising environmental problems (Attia, 2011). BPS has been pursued by modelers and engineers, assisting in attaining successful environmental design. Also, for optimization, it is important to couple BPS with parametric generative systems, to evaluate all design options. However, BPS remains poorly integrated into generative design processes (Danhaive & Mueller, 2021). More than before, environmental performance simulation, at both architectural and urban scales, has become essential to achieve resilience and adaptability to rising environmental problems as targeted in the UN Sustainable Development Goals 11. Within this goal, re-thinking the design process and prototyping new approaches to (facilitate) achieving high environmental performance in the built environment became important, which motivated this

research.

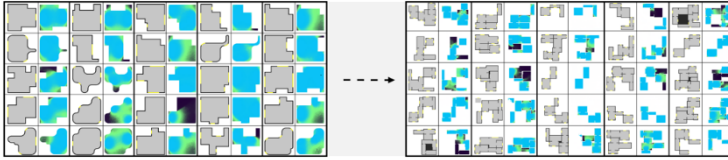
Simulation is still perceived as a complicated process that requires expert knowledge. Studies show that non-experts' simulation studies contain so many errors and shortcomings that the value of these studies in the design process becomes questionable (Reinhart & Wienold, 2011). For lay architects to perform BPS, the task requires a level of competence and technical abilities that they often do not possess. Another problem with conducting BPS studies is the lack of their seamless integration into generative design frameworks and associated issues of designer interaction. Also, there is a need to compensate for the normal idle computation time in the generative design process to support parallelized data exploration and synthesis while retaining close human-in-the-loop engagement (Urban Davis et al., 2021). More importantly, BPS is still computationally expensive and takes a long simulation time. Daylight simulation is time-consuming due to ray-tracing and requires especially an extended time for annual studies. The computation time issue becomes particularly problematic in generative parametric systems, where thousands of design alternatives emerge iteratively and require instantaneous feedback. Within BPS, energy modeling and daylight analysis are often targeted. Daylight performance has a direct impact on the building form, i.e., daylight is a form giver (Reinhart, 2019). Compact forms have fewer surfaces and thus fewer opportunities for daylighting, while extended forms perform better due to their higher surface-to-volume ratio, allowing more opportunities for window surfaces (Caldas, 2001). This means designers need to know about each design's daylight performance to make informed decisions in regard to morphology and design revisions early in the design process.

The recent introduction of AI has leveraged design protocols, marking the transition to the second generation of generative systems (Chen & Stouffs, 2021). Since AI can analyze, learn, and synthesize data, it can aid designers in making successful decisions by enabling the prediction of environmental parameters of their designs. Yet, developing accurate, predictive real-time techniques for environmental analysis remains difficult (Rahmani Asl et al., 2017). A solution to achieve high-fidelity real-time prediction of environmental analysis in generative design is to employ deep learning-based surrogate modeling. A surrogate model is "an approximation of the input-output data obtained by the simulation" (Kim & Boukouvala, 2019, p. 2). Deep learning (DL) models are computational models with multiple processing layers that learn representations of data at multiple levels of abstraction. Generative Adversarial Networks (GANs) are DL-based models that allow machines to learn structures and semantics by extracting features from input datasets (Goodfellow et al., 2016), and thus can be used for surrogate modeling.

In this paper, we address the environmental performance of the enclosure system where we target daylight analysis with the use of a deep learning-based surrogate model. The proposed surrogate model was developed in three stages: (1) generative modeling and daylight simulation for data acquisition; (2) DL-model training for building the surrogate model; and (3) assessment and validation. In our previously published Deep-Performance framework (Yousif & Bolojan, 2021), the investigation was limited to single-space floor plans without interior walls, to test the initial prototype of the proposed method. In this paper, we present further development to our method by improving the surrogate model and conducting additional experiments to include



more complex spatial configurations for the input datasets, as well as interior wall partitions and multi-room floor plan layouts. Our improved method shows promising results with regards to accurate predictions of daylight performance for complex floor plan designs with interior wall partitions and multi-space configurations (Figure 1).



*Figure 1. Left: The previous dataset with simple single-space layouts; right: the new dataset with complex multi-space floor plan layouts.*

## 2. Background

It is speculated that AI is most likely to have the biggest influence on performance-based aspects, particularly in architectural practice and urban design, where data-driven approaches and performance-informed design are becoming increasingly important (Leach, 2021). Reviewing existing literature, performative AI has seen an exponential increase in architectural research addressing multiple building performance aspects. An important reference for AI research in design is the City Intelligence Lab in Austria, offering a digital platform that utilizes AI technologies for urban planning workflows and processes. The lab examines technologies that involve generative design and AI solutions for data-driven design (2019). In design practice, the XKool firm investigates how deep learning can be used to construct numerous environmental scenarios, based on training datasets that combine geographical variables and building codes (2019). Another pioneer is Spacemaker, a design firm that offers AI-driven design and planning solutions for early design stages in pursuing sustainable built environments (2020). In approximating building energy modeling, the work of (Singaravel et al.) uses a method of component-based machine learning for mimicking BPS (2018). Papadopoulos et al. (2018) employ machine learning techniques combined with genetic algorithm-based optimization to offer energy use evaluations of building designs. Research is also expanding in using AI for automating daylight analysis (i.e., Ngarambe et al., 2020; Shaghaghian & Yan, 2019) Despite such progress in performative AI, these daylight-related approaches represent undergoing experimentation and do not yet offer validated methods for real-time daylight performance prediction in generative systems, which has motivated this project. We extend the state-of-the-art by introducing a new deep learning-based surrogate model that predicts daylight performance of floor plan designs with high accuracy.

Before describing the research methodology, we offer here definitions and explanations of the implemented methods. Surrogate models are prediction models that seek to approximate the output of simulation models as closely as possible and can offer compact and instantaneous performance information instead of simulation (Forrester et al., 2008). GANs are techniques for training a machine to perform complex tasks in a generative process measured against a set of training images (Goodfellow et al., 2016; Leach, 2021). A sub-class of GANs are the conditional or

"supervised" models such as Pix2Pix. In this image-to-image translation model, synthesizing images is based on labeling or pairing corresponding datasets. Reconstructing and producing images occurs based on the labels of one image of the pair (Isola et al., 2017). We have adopted the Pix2Pix model and further improved it, formulating a revised strategy to be applied to building designs, and their corresponding datasets. To verify and validate the accuracy of the DL prediction results, we implemented two quantified metrics, the Structural Similarity Index Model (SSIM), and Perceptual Similarity (PS). SSIM is a method for assessing AI-synthesized image quality that involves collecting structural data and evaluating the degradation of that data for the images in question (Wang et al., 2004). PS is another method for evaluating deep learning-generated images in a manner comparable to how humans make perceptual decisions (Zhang et al., 2018).

### 3. Research Methods

As explained above, this study was aimed at developing a new framework that incorporates an accurate approximation method, a surrogate model for predicting daylight studies of design options in generative design protocols. The methodology included experimentation with the surrogate model techniques, developing the model to be integrated into a performance-driven generative framework, prototyping the overall framework, application, and testing. In the development phase (authors' framework), the prototype was formulated into three tasks, as illustrated on the left side of Figure 2. First, (1) dataset acquisition was pursued using a parametric system with daylight simulation integrated, (2) the DL-based model was trained for prediction of daylight performance, and (3) assessment and validation studies were conducted, comparing prediction with actual simulations. For the system users (designers), in the application phase, the system becomes a two-process workflow that consists of a generative process (with floor plan design options) and a real-time daylight performance prediction offered by our trained model, as shown in the right part of Figure 2. The framework prototype comprises a package of algorithms that is under development and will be released as a Grasshopper add-on.

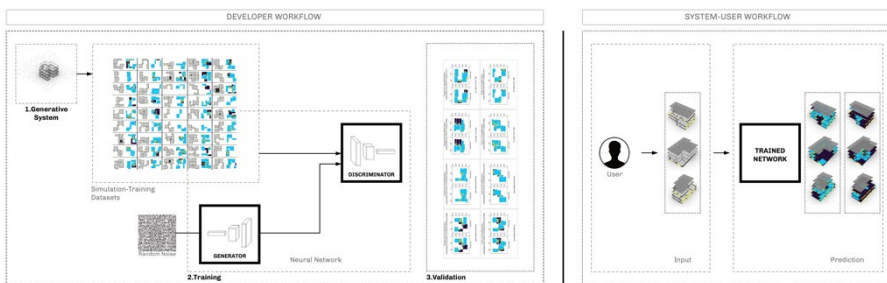


Figure 2. The workflow of the proposed DP framework, real-time daylight performance evaluation predictions using deep learning. The left side is the authors' workflow, and the right side is the users' workflow.

#### 3.1. SURROGATE MODEL DEVELOPMENT

In our previous test-case application (Yousif & Bolojan, 2021), the method was applied

to a sample of simplified floor plan layouts. To improve the prototype and apply the DL model to real (existing) floor plan designs, this experiment was conducted. The framework was developed using different methods for the required tasks. Dataset generation was done using the Rhino/Grasshopper® environment for parametric modeling and LadyBug® and HoneyBee® plugins (Roudsari, Pak, and Smith 2013) were used for daylight simulation. PyTorch® and Tensorflow® deep learning packages were used to develop and train the surrogate model. Validation was performed using machine learning algorithms such as Structural Similarity Index Metric (SSIM) and Perceptual Similarity (PS).

### 3.1.1. Dataset Generation

To generate the required data for training, we defined a parametric model to create 1815-floor plan design options that represent typological design layouts (Figure 4), and to perform daylight simulation for those generated designs. In the testing phase, the "CubiCasa" (Kalervo et al., 2019) dataset was used. Each design option was subjected to five annual daylight simulation metrics: (1) spatial Daylight Autonomy (sDA), (2) Direct Light Access (DLA), and (3) Useful Daylight Illuminance ranges of UDLI0-100, (4) UDLI100-2000, and (5) UDLI2000\_more in lux. sDA is an annual metric that quantifies the ratio of the area within a space for which the daylight autonomy exceeds a specified value (IES, 2020). DLA measures the space's access to direct daylight for a specific duration, while UDLI is a ratio of time across a daylighting study period where the illuminance at a point is between certain minimum and maximum levels. Each daylight simulation was conducted using a grid resolution of 0.2m (the distance between the test points/daylight sensors) and at a typically used working plane height of 0.76m. As shown in Figures 4, the floor plans were annotated by defining three major classes/labels. The light gray color represents the floor area class, the black label is for the wall class, and the yellow color signifies the window class. In this instance, the yellow color saturation determines the window height (in this instance, we used a uniform ceiling-to-floor height for all windows), and the gray shading denotes the program of the space, which was an open office in this experiment. Future developments and additional test-case applications will use different colors to define program activities, such as office, retail, as well as considering shading devices.

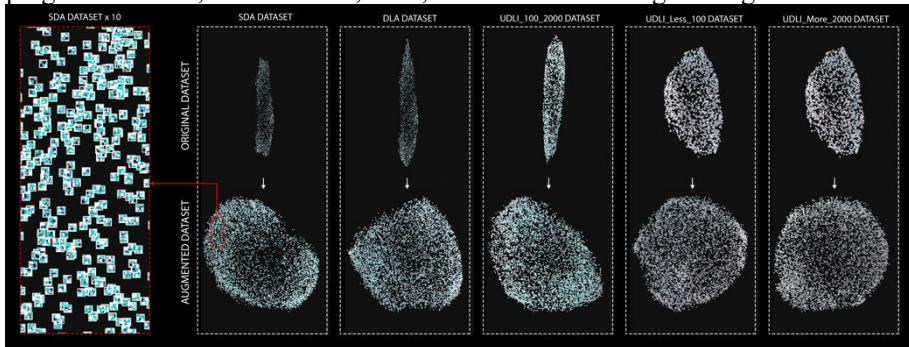


Figure 3. Topological Data Analysis. Dataset augmentation is represented through a U-Map algorithm. The top row represents the original dataset distribution, while the bottom row illustrates the augmented data distribution.

In the dataset management phase, 1815 floor plan samples were augmented by applying various transformation procedures (preserving the north orientation) to the parametric model prior to running the corresponding daylight simulation, increasing the dataset to 5400 pairs. Daylight metric simulations were carried out with the newly augmented dataset and paired with their corresponding floor plan layouts. For training, the data was structured as follows: 85% of the paired images were used for training and 15% were used as testing sets. Thus, we used 4590 pairs for training and another 810 for testing. Five Pix2Pix models were trained, with each model corresponding to one daylight metric. The dataset floor plans comprised different complexities, multiple/single spaces, and areas varying between 50 and 200 square meters. Data augmentation was a crucial task before model training. Normally, the performance of the deep learning model depends on two aspects: the neural network model's architecture, and the quality and quantity of data. Even when using the proper model for the intended task, the results can be unsatisfactory. The reason for this can be attributed to the dataset quality and diversity; therefore, understanding dataset topology is crucial. We used a U-Map algorithm to reduce the dimensionality of the features' manifold and to understand and analyze the topology of the dataset. As the upper row of Figure 3 shows, the initial dataset of 1815 paired images had a very limited (almost linear) distribution and limited variation (diversity). Therefore, domain-specific augmentation techniques were applied to create new and different data points (training samples), resulting in a more diverse and enhanced distribution of the dataset, as shown in the lower row of the figure.

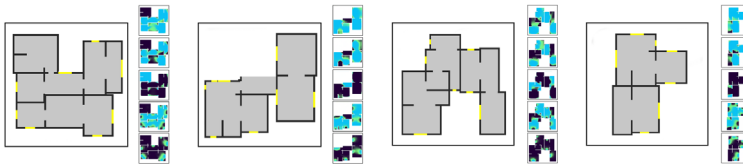


Figure 4. A sample of 4 design options was evaluated against 5 daylight simulation metrics of sDA, DLA, UDLI100, UDLI100-2000, and UDLI2000\_more.

### 3.1.2. Model Training

For training the model, the network architecture used is based on the Pix2Pix model, yet we improved the networks' filters to create feature maps that summarize the identified features in the input data (Figure 4). All five models were trained for a total of 200 epochs, which required approximately 24 hours of training. For the first half of the training, the Adam optimizer was utilized with a learning rate of 0.0005, which was then decreased linearly to 0.0000 throughout the remaining iterations. For both generator and discriminator networks, the optimizer momentum term was set to 0.65. The resolution was set to 1024x1024 pixels for both the input and output. Due to GPU training device memory constraints, a batch size of 40 was used, and we normalized the input layer with batch normalization (Ioffe & Szegedy, 2015). For the discriminator network, the basic network architecture of 70x70 PatchGAN was used, and the generator network used was a RESNET 9blocks network architecture. To reduce the model's oscillation, we used a pool size of 50 (Shrivastava et al., 2017). Each network's layer is followed by instance normalization and a ReLU layer.

3.1.3. Results and Validation

In the testing phase, when new floor plan designs were introduced to the trained model, it predicted the daylight simulation successfully. The accuracy of the synthetic (predicted) daylight meshes was assessed against the HoneyBee®-based (actual) simulation results. In one assessment study, quantifying the accuracy comparison between the surrogate model and the simulation model results. The results were evaluated by quantifying the ratio of the area of the space for which the daylight metric exceeds or lies within a specified value. This represents the same method used by the HoneyBee® add-on to calculate the daylight metrics. The comparison results showed that the network predicted the sDA metric with an average prediction accuracy of 89% when presented with a new batch (unseen by the trained model) of 15-floor plan alternatives of real floor plans, the CubiCasa dataset, (Figure 5). The figure demonstrates a comparison between the simulation sDA values and meshes performed using HoneyBee® and the DP-Prediction of sDA values and meshes.

<b>Simulation sDA</b>	89.98	84.4	17.2	48.0	90.84	95.0	86.6	28.9	31.0	46.2	50.3	98.2	67.1	85.6	84.3
<b>Prediction sDA</b>	84.2	79.7	14.8	45.7	88.24	86.2	76.1	22.6	28.2	45.0	43.4	94.9	59.86	80	74.9
real_B															
fake_B															

Figure 5. Prediction results of testing the trained model with the (CubiCasa) dataset- comparison between the simulation sDA values and DP-Prediction of sDA values.

It is important to note here that one advantage of the Pix2Pix model is that it is trained with images that represent daylight meshes (grids), which can be easily quantified and measured post-training. Quantitative assessment methods were followed, assessing the accuracy of the prediction results of the surrogate model. The Structural Similarity Index Measure, or SSIM, was the first validation study we conducted. In this comparison, the predicted daylight simulation results were evaluated against the HoneyBee® simulation, and the results of testing (four floor plans of the CubiCasa dataset) show an average similarity value of 0.95 with an overall similarity value range of 0.91-0.99. For the second assessment method, perceptual similarity, the lowest value retrieved was 0.93 and the highest value was 0.97. Both assessments were performed using the sDA metric (Figure 6).

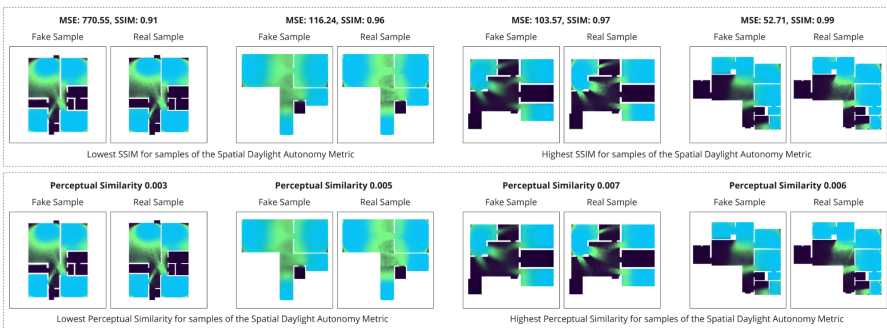
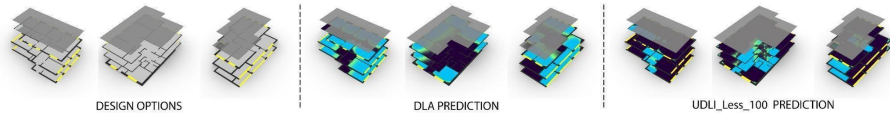


Figure 6. Four floor plans of the testing dataset (CubiCasa) with synthetic-real pairs were assessed by the SSIM (upper row) and PS measure (lower row) for the sDA metric.

#### 4. Discussions

Our system, as presented, can predict daylighting with high accuracy. More significantly, this test-case application has shown promise in automating other environmental performance goals. Automatic prediction is possible with such a surrogate model, which is useful for decision-making at the early stages of design. The prediction model is to be injected into design generation, as depicted in Figure 7. It will allow designers to explore a wide range of design choices while assessing design performance in real-time throughout the inference phase. The significance of retrieving predicted daylight analysis is its impact on morphology and associated design decisions in design development. Automation of performative aspects accelerates and improves design decision-making, allowing for a faster feedback loop between design decision and environmental evaluation.

The proposed method is 600 times faster than the typical annual daylight simulation performed by HoneyBee® and required for the floor plans under consideration. Compared to the HoneyBee® daylight simulation results that took 3 minutes for each simulation run, our surrogate model was able to offer a comparable accuracy of 90%, taking less than 0.3 seconds for each prediction. In a generative design process, when exploring 6000 design options, it would take  $3 \times 6000 = 18000$  minutes, equivalent to 12.5 days to retrieve daylight simulation results using HoneyBee® or Diva, in contrast to  $0.3 \times 6000 = 1800$  seconds, equivalent to 30 minutes using the surrogate model. Besides saving computation time, our method offers a pre-trained model that can be used by designers into generative protocols for instant feedback on daylight performance. This way, any designer can have access to environmental performance of their designs.



*Figure 7. Three building design options are shown in the far-left part, and their two copies in the central and far-right parts with automatic prediction are presented in the daylight simulation meshes for each floor plan.*

#### 5. Conclusions and Future Work

Presented here is research that contributes to the transition from simulation to prediction-based performance evaluation, using surrogate modeling. An approach was developed to provide real-time daylight performance predictions of high fidelity to enhance generative design methods. The findings suggest that deep learning approaches might be used to automate additional building performance measures. The significance of this research is to enable systematic performance-driven design space navigation by injecting trained models into a design process driven by designers' creative exploration. The ultimate objective is to enable environmentally efficient and affordable design of the built environment using data-driven approaches. This goal is aligned with SDG 11 to make cities and human settlements resilient and sustainable. In order to achieve sustainability, design processes should be performance-driven and involve environmental feedback of design options and enable identifying

environmental consequences of design decisions.

For future work, in further developing the model, we aim to use labeling the floor plans according to program activities to achieve an accurate simulation for realistic floor plans with multiple-program activities. Also, the next step is to add parameters for window heights and shading devices (and their dimensions), encoding this information into the floorplan labels. In addition, more dynamic and angled floor plans will be pursued in future applications. Another future work involves facilitating optimization by filtering the optimum layout/s in terms of their daylight performance. The framework will also be improved to achieve an articulated and searchable design space when navigating through thousands of design options. Also, additional AI models will be required to sort out successful design alternatives with higher environmental performance. Further improvements will be focused on formulating a designer-friendly interface that will be assessed by lay architects in empirical studies. To achieve computation efficiency, a cloud-based interface integrating our method will be targeted. In addition, with transfer learning, developing generalizable models will be pursued.

## References

- Attia, S. (2011). *State of the art of existing early design simulation tools for net zero energy buildings: a comparison of ten tools*.
- Caldas, L. (2001). *An evolution-based generative design system: using adaptation to shape architectural form*. Massachusetts Institute of Technology,
- Chen, J., & Stouffs, R. (2021). From Exploration to Interpretation-Adopting Deep Representation Learning Models to Latent Space Interpretation of Architectural Design Alternatives. In *the Proceedings of the 26th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2021*.
- City Intelligence Lab. (2019). Austrian Institute of Technology - AIT 2022. Retrieved from <https://cities.ait.ac.at/site/>
- Danhaive, R., & Mueller, C. T. (2021). Design subspace learning: Structural design space exploration using performance-conditioned generative modeling. *Automation in Construction*, 127, 103664. doi:<https://doi.org/10.1016/j.autcon.2021.103664>
- Forrester, A., Sobester, A., & Keane, A. (2008). *Engineering design via surrogate modelling: a practical guide*: John Wiley & Sons.
- Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1): MIT press Cambridge.
- IES. (2020). *ANSI/IES LS-1-20*. Retrieved from <https://www.ies.org/standards/definitions/>.
- Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *the International conference on machine learning*.
- Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In *the Proceedings of the IEEE conference on computer vision and pattern recognition*.
- Kalervo, A., Ylioinas, J., Häikiö, M., Karhu, A., & Kannala, J. (2019). *Cubicasa5k: A dataset and an improved multi-task model for floorplan image analysis*. Paper presented at the Scandinavian Conference on Image Analysis.
- Kim, S. H., & Boukouvala, F. (2019). Machine learning-based surrogate modeling for data-driven optimization: a comparison of subset selection for regression techniques. *Optimization Letters*, 1-22.
- Leach, N. (2021). *Architecture in the Age of Artificial Intelligence: An Introduction to AI for Architects*: Bloomsbury Visual Arts.

- Ngarambe, J., Irakoze, A., Yun, G. Y., & Kim, G. (2020). Comparative performance of machine learning algorithms in the prediction of indoor daylight illuminances. *Sustainability*, 12(11), 4471.
- Papadopoulos, S., Woon, W. L., & Azar, E. (2018). *Machine Learning as Surrogate to Building Performance Simulation: A Building Design Optimization Application*, Cham.
- Rahmani Asl, M., Das, S., Tsai, B., Molloy, I., & Hauck, A. (2017). Energy Model Machine (EMM)-Instant Building Energy Prediction using Machine Learning.
- Reinhart, C. (Producer). (2019). The Value of Daylight in Office Buildings-The 8th VELUX Daylight Symposium. Retrieved from <http://thedaylightsite.com/>
- Reinhart, C. F., & Wienold, J. (2011). The daylighting dashboard—A simulation-based design analysis for daylit spaces. *Building and environment*, 46(2), 386-396.
- Roudsari, M. S., Pak, M., & Smith, A. (2013). *Ladybug: a parametric environmental plugin for grasshopper to help designers create an environmentally-conscious design*. Paper presented at the Proceedings of the 13th international IBPSA conference held in Lyon, France Aug.
- Shaghaghian, Z., & Yan, W. (2019). Application of Deep Learning in Generating Desired Design Options: Experiments Using Synthetic Training Dataset. *arXiv preprint arXiv:2001.05849*.
- Shrivastava, A., Pfister, T., Tuzel, O., Susskind, J., Wang, W., & Webb, R. (2017). *Learning from simulated and unsupervised images through adversarial training*. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.
- Singaravel, S., Suykens, J., & Geyer, P. (2018). Deep-learning neural-network architectures and methods: Using component-based models in building-design energy prediction. *Advanced Engineering Informatics*, 38, 81-90.
- Spacemaker. (2020). Early stage planning. Re-imagined. Retrieved from <https://www.spacemakerai.com/>
- Urban Davis, J., Anderson, F., Stroetzel, M., Grossman, T., & Fitzmaurice, G. (2021). *Designing Co-Creative AI for Virtual Environments*. Paper presented at the Creativity and Cognition.
- Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4), 600-612.
- XKool. (2019). Retrieved from <https://www.xkool.ai/>
- Yousif, S., & Bolojan, D. (2021). *Deep-Performance: Incorporating Deep Learning for Automating Building Performance Simulation in Generative Systems*. Projections, the 26th Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong, China. [http://papers.cumincad.org/cgi-bin/works/paper/caadria2021\\_052](http://papers.cumincad.org/cgi-bin/works/paper/caadria2021_052).
- Zhang, R., Isola, P., Efros, A. A., Shechtman, E., & Wang, O. (2018). *The unreasonable effectiveness of deep features as a perceptual metric*. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.



# BUBBLE2FLOOR

*A pedagogical experience with deep learning for floor plan generation*

PEDRO VELOSO<sup>1</sup>, JINMO RHEE<sup>2</sup>, ARDAVAN BIDGOLI<sup>3</sup> and  
MANUEL LADRON DE GUEVARA<sup>4</sup>

<sup>1,2,3,4</sup>*Carnegie Mellon University, USA.*

<sup>1</sup>*University of Arkansas, USA.*

<sup>1</sup>*pveloso@uark.edu, 0000-0003-1597-9533*

<sup>2</sup>*jinmor@andrew.cmu.edu, 0000-0003-4710-7385*

<sup>3</sup>*abidgoli@andrew.cmu.edu, 0000-0001-5486-2413*

<sup>4</sup>*manuelr@andrew.cmu.edu, 0000-0002-4585-3213*

**Abstract.** This paper reports a pedagogical experience that incorporates deep learning to design in the context of a recently created course at the Carnegie Mellon University School of Architecture. It analyses an exercise called Bubble2Floor (B2F), where students design floor plans for a multi-story row-house complex. The pipeline for B2F includes a parametric workflow to synthesise an image dataset with pairs of apartment floor plans and corresponding bubble diagrams, a modified Pix2Pix model that maps bubble diagrams to floor plan diagrams, and a computer vision workflow to translate images to the geometric model. In this pedagogical research, we provide a series of observations on challenges faced by students and how they customised different elements of B2F, to address their personal preferences and problem constraints of the housing complex as well as the obstacles from the computational workflow. Based on these observations, we conclude by emphasising the importance of training architects to be active agents in the creation of deep learning workflows and make them accessible for socially relevant and constrained design problems, such as housing.

**Keywords.** Architectural Pedagogy; Deep Learning; Conditional GAN; Space Planning; Floor Plan; SDG 4; SDG 9.

## 1. Introduction

In the last decade, Deep Learning (DL) spearheaded the latest wave of research in Artificial Intelligence (AI) and brought a myriad of technological advancements in various fields. DL is an umbrella term that refers to the methods that “... allow computers to learn from experience and understand the world in terms of a hierarchy of concepts, with each concept defined through its relation to simpler concepts” (Goodfellow et al., 2016, p. 1). DL typically relies on the use of Deep Neural Networks,

namely, multi-layered models with small computational components that can infer the hierarchy of concepts related to the solution of a problem from a given dataset.

DL gained a lot of attention from the design technology community and the AEC industry and resulted in a significant number of publications, presented in newly established DL sections in AEC flagship conferences. In this paper, we describe our experience of teaching DL and generative design to architecture students, as part of a graduate-level course at the Carnegie Mellon University School of Architecture. After the initial modules of the course, where the students learned the fundamentals of DL frameworks, we introduced a generative design exercise named Bubble2Floor (B2F). B2F entails generating custom floor plan arrangements in a predefined row-house complex using Generative Adversarial Networks (GANs).

This paper adopts a pedagogical research method to develop a critical reflection on B2F and on the effectiveness of using DL as a generative method for architectural design education. The next sections consist of (2) a review of previous pedagogical initiatives that use GANs for architectural design, (3) a general description of B2F, (4) an analysis of the design development and challenges in B2F, and (5) the conclusion and discussion. We assume that the reader will be familiar with basic terminology associated with DL.

## 2. Review of Deep Learning in Architectural Education

A large part of the research on DL for architecture is targeted at the synthesis of design representations. Initiatives that either analyse the limits of DL in design (Joyce & Nazim, 2021) or that test these ideas in an educational setting are still in their infancy. Some of the early efforts to integrate the recent bloom of DL in design pedagogy were crystallised as a series of workshops in CAAD conferences. For instance, the line-up of workshops in Smartgeometry (Sg2018 Workshops, 2018) included Fresh Eyes, where participants used various DL techniques to classify building types and predict building performance among other inquiries into DL and design. ACADIA hosted a number of workshops focused on DL and specifically GANs. Deepdesign (ACADIA Workshops, 2020), Latent Morphologies, and The Generative Game (ACADIA Workshops, 2021) all utilised a variation of GANs as a design assisting tool. The newly established platform, Digital Futures, has been the home for several workshops on the intersection of DL and design as well. For instance, Machine Intelligence in Architecture 2.0 (Tian et al., 2021) addressed various applications of GANs in 2D, 2.5D, and 3D. Throughout these venues, GANs were the most popular generative models adopted by workshop organisers and tutors.

The growing interest in the applications of DL recently found its way into a handful of universities' curriculums. During the past years, Carnegie Mellon University (USA), University of the Arts London (UK), Goldsmith University (UK), among others, integrated creative AI and ML in their course catalogues. Some architecture schools, such as the School of Architecture at Carnegie Mellon University as well as Taubman College at University of Michigan, have initiated dedicated courses to AI, ML, DL, and design. A few other schools, however, covered this topic in their existing curriculum under the broader umbrella of computational design.

### 3. Bubble2Floor (B2F)

B2F is a design exercise that consists of generating floor plans with DL for a row-house complex of six three-story units. The first unit of the complex is used as an example to demonstrate the expected level of details in the design, while the other five units are reserved for each group of students to work on.

We designed a pipeline for B2F (Figure 1) to address the design constraints of the building layout problem based on the interaction of the designer with a well-established design representation. In this pipeline, the DL model receives bubble diagrams as input and translates them to floor plan diagrams as output. This translation is a long-standing problem, which relied on human intervention in early generative experiments (Weinzapfel & Negroponte, 1976) and on techniques based on graph embedding and triangulation (Nourian et al., 2013). Recently, this problem has been addressed with DL techniques (Nauata et al., 2020, 2021). For students to customise bubble diagrams and explore variations of floor plan designs, we divided the B2F pipeline in three parts: (3.1) data processing and synthesis, (3.2) training, and (3.3) design.

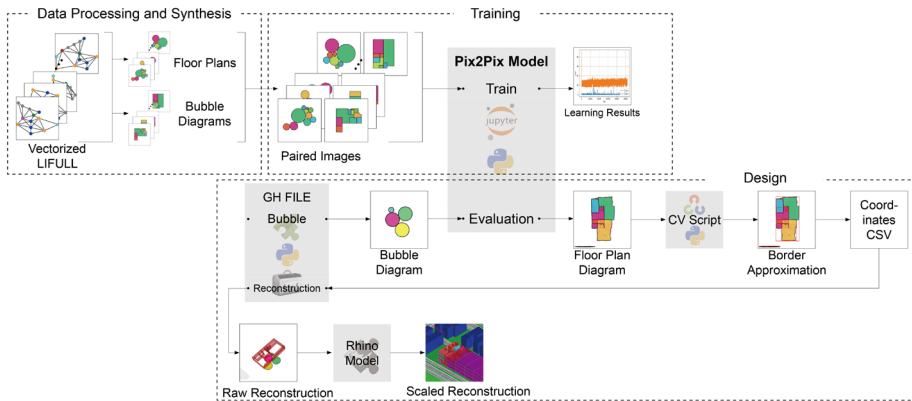


Figure 1. Computational pipeline of B2F: data processing and synthesis, training, and design.

#### 3.1. DATA PROCESSING AND SYNTHESIS

The dataset is based on vector-graphics annotations of LIFULL, an image-based dataset of millions of floor plan images from a real estate information service in Japan (Liu et al., 2017). We used the vectorized representation of the floor plans and programmatic labels (bedroom, restroom, bathroom, kitchen, living room, entrance, hall, corridor, stairs, balcony, washing room, etc.) as the base to generate pairs of floor plan and bubble diagrams (see Figure 1).

The floor plan diagram is defined by the original room layout polygons with black edges and fill colours associated with the programmatic labels. The way to generate bubble diagrams is converting the respective room polygons into discs of equivalent areas with the same edge and fill colour. Then, a constraint-based physics simulator removes overlapping between the discs. Based on the parameters and on the stochasticity of the simulation, this process can generate multiple bubble diagrams for the same floor plan. While both the bubble and floor plan diagrams contain information

related to program, area, position, and adjacencies, the floor plan diagram contains additional information about the position and room shape. Each diagram is converted into a  $256 \times 256$  PNG image file, forming a dataset of 4,987 images by pairing bubble and floor plan diagrams. The use of these two diagrams intends to display the potentials and limits of design representation and translation in early design stages.

### 3.2. THE TRAINING WORKFLOW

B2F uses a conditional GAN model named Pix2Pix (Isola et al., 2017) for image-to-image translation with paired images. Pix2Pix is a generative adversarial model that can relate conditional inputs to outputs—i.e., it uses the input images as conditional variables for generating a new image, in contrast to generating images from random noise, as in the original GAN model (Goodfellow et al., 2014). In B2F, the synthetic dataset is used to train the model to generate a floor plan diagram conditioned on a given bubble diagram. We developed a custom training workflow based on the implementation provided by (H. Kang & Jha, 2018).

Students spent one week comprehending the main concepts of conditional GANs, Pix2Pix, and the pipeline of B2F before training the model. We provided a Jupyter Notebook file to train the B2F model with the dataset through Google Colab—a cloud computing service with all required libraries and dependencies pre-installed. This notebook had the basic algorithms implemented with suggested initial settings and hyper-parameters to train the model. Also, a model pre-trained for 100 epochs was provided as the training baseline, so students could test the pipeline and refine the model with additional epochs if necessary. Some students individually trained the networks for additional epochs. Other students contributed to the development of a single model with combined training sessions. Overall, the models were trained from 200 to 500 epochs in total.

### 3.3. DESIGN WORKFLOW

We provided a post-processing pipeline to integrate B2F into an end-to-end design workflow. This pipeline incorporates three components:

- A geometric model in Rhinoceros3D (RH) including the site and the row-house building with different housing units
- A Grasshopper (GH) definition that can generate and colour the bubble diagrams according to their program
- A computer vision script based on OpenCV library and a GH definition that can translate images to geometric partitions

After the training phase, students generated new bubble diagram images, using the provided GH file, which automatically adjusts the number, size, position, and program/colour of the bubbles in real time. The trained Pix2Pix model takes the bubble diagram, translates it into corresponding feature vectors, and synthesises a floor plan diagram as an image.

From this floor plan diagram, students can explore potential configurations for the row-house building model with two different post-processing methods: (a) rectangular

approximation, and (b) coordinate clustering. In rectangular approximation, an OpenCV script tracks the borders of room partitions, approximates them into rectangular polylines, and saves their coordinates. In coordinate clustering, the script extracts point coordinates from the floor plan image. Then, a K-means clustering algorithm is used to cluster these point coordinates based on the number of clusters defined by the designer (Figure 2). This process reinforces the presence of axes and orthogonal wall partitions parameterized by the designer.

Students completed the design of the unit by adding openings and stairs to wall partitions in the RH file. Each student repeated this process for each floor to complete designing the house unit. At the end, we collected the five different units and merged them into the row house in RH (Figure 3).

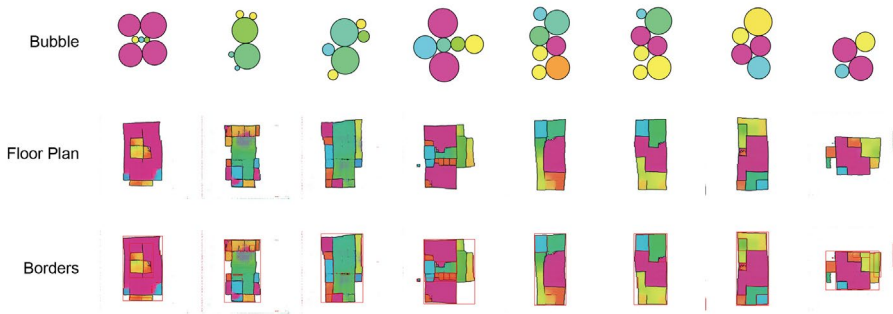


Figure 2. Floor plan generation from bubble diagrams with the rectangular approximation and coordinate clustering methods.

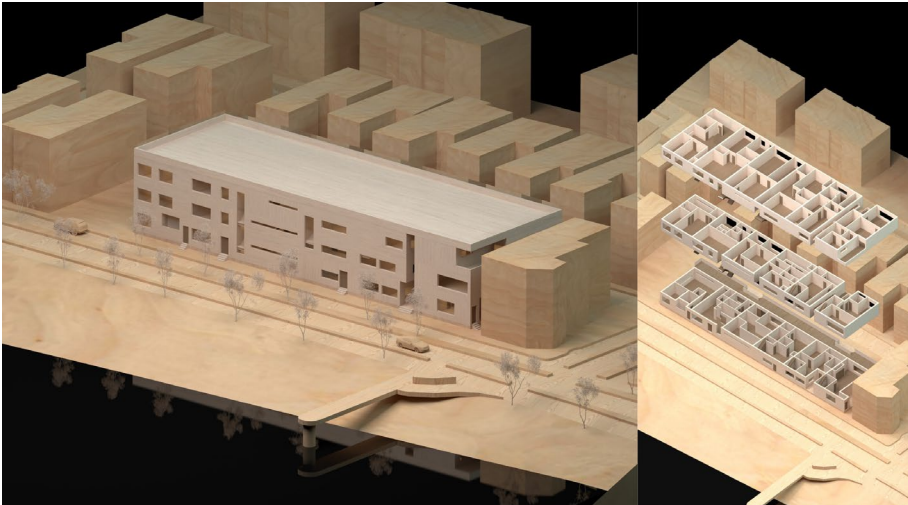


Figure 3. Axonometric models of a row house designed in the B2F exercise.

## 4. Analysis

### 4.1. DESIGN

In this section, we examine two design cases from the students' work (Figure 4) and review interesting observations and challenges that they faced during the exercise. As we conducted this study over two consecutive semesters, each case is selected from one of the two cohorts of the class.

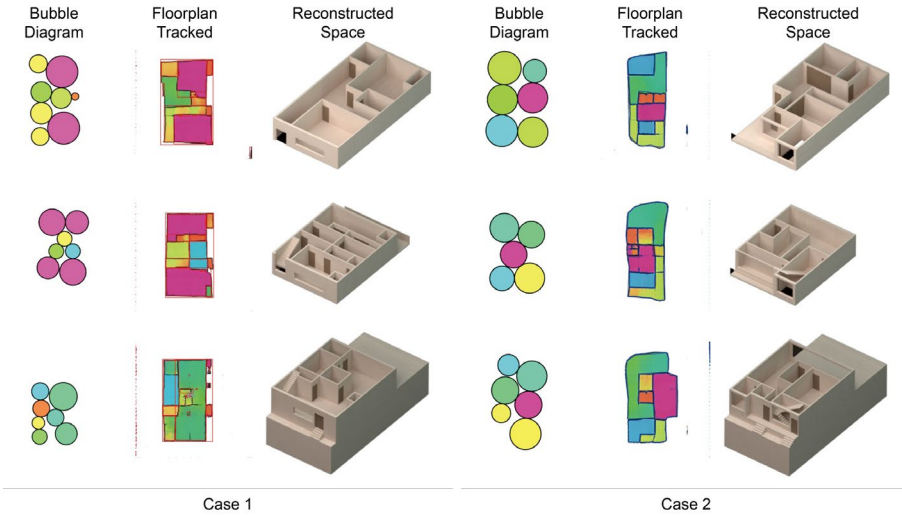


Figure 4. Two design cases from the Bubble2Floor exercise with different post-processing methods.

In case 1, students used the rectangular approximation method to reconstruct floor diagram images to 3D models. The design has an entrance facing the main street, one large living room on the first floor, four bedrooms on the second and third floors, and stairs at the same location of each floor. In case 2, students used the point clustering method for reconstruction. The design has two main entrances, one for a studio on the first floor and another for a two-bedroom apartment on the second and third floors. The apartment has two balconies facing the main street on the second and third floors.

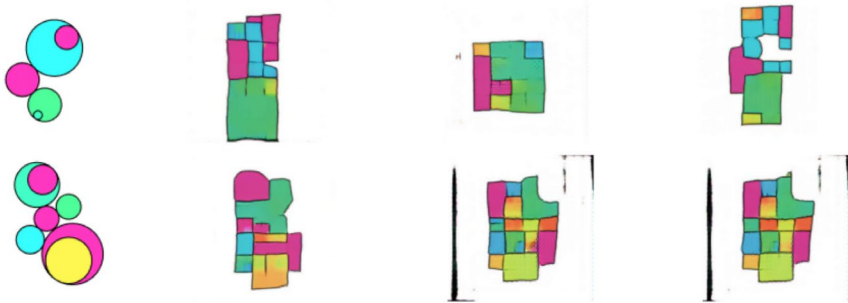
In both cases, the translation between bubble to floor diagram works according to colour codes and relationships. For example, the bubble diagram for the second floor in case 1 has two large magenta discs at the top and bottom of the image, bracing small, partitioned spaces corresponding to smaller discs with yellow, green, and light blue colours. The adjacencies between discs are preserved in the respective floor plan.

We can find some graphical differences between the cases. The floor plan diagrams in case 1 seem to have more vivid edges and sharper corners, compared to case 2, which has rounded corners. Students in case 1 and 2 trained their Pix2Pix model for 200 and 500 epochs respectively. With 500 epochs, the model seems to overfit the training data and does not generalize well to the new bubble diagrams, which results in irregular floor plan edges.

The translation from floor plans into 3D models requires manual adjustments due

to (a) the different scale between the apartments in the dataset and the target footprint of B2F, and (b) the exploration of spatial patterns that were not well-represented in the training set. Students approached this step in different ways. In case 1, students tried to keep the original information of the extracted partition from floor plan diagrams as much as they could. Thereby, they removed the unintended programmatic elements that appeared on the floor plan diagrams but kept the main organisation from the bubble diagrams. Meanwhile, in case 2, students manually edited the synthesised floor plans to reinforce the practicality of the layout.

Students also had different stances on the DL method. In case 1, they expected an almost fully automated process—i.e., if users input a bubble diagram, the pipeline should generate a functional layout. As a result, they expressed their frustration as they had to play with different steps of the pipeline. Conversely, in case 2, students acknowledged the imprecision and the limitations of the workflow as part of the exercise, so they opted for exploring emerging and unexpected architectural designs. For instance, they subverted the original bubble diagram representation by adding overlapping discs to generate courtyards in the floor plan (Figure 5).



*Figure 5. Examples of overlapping discs to create courtyards in a floor plan diagram.*

We observed that the different approaches between case 1 and 2 were aligned with the pedagogical decisions made by the instructors. In the first semester, we guided students to individually train their models and provided a Python script to apply rectangular approximation. Rectangular approximation provides a limited level of control for the designer during post-processing and limits the floor plan configurations to mosaics of axis-aligned rectangles. In contrast, in the second semester, we allowed students to collectively train the models and provided a script that relies on coordinate clustering to post-process the layout. These enabled students to explore more expressive and varied floor plan configurations from the same floor plan image synthesised by the Pix2Pix model.

#### 4.2. DATA REPRESENTATION AND SYNTHESIS

Currently, the type of representation and the dataset in the exercise were defined by the instructors before students developed their own projects. This approach was well-received by students as it enabled them to concentrate on learning and using the B2F

pipeline in the three-week span of this module. However, this choice resulted in some bottlenecks for design exploration, such as in the example of floorplan with courtyards (Figure 5). Ideally, the students should be able to create their own dataset and incorporate other forms of representation required by their design intentions, such as graphs, rectangular mosaics, etc.

#### 4.3. TECHNICAL KNOWLEDGE AND DESIGN KNOWLEDGE

A programming background was a prerequisite for this course, which enabled us to introduce fundamentals of DL and hands-on tutorials with PyTorch as DL framework. However, we minimised the discussion on in-depth concepts related to the DL models used in B2F. As a result, while the students could use the pipeline to produce interesting designs, they faced technical issues that they could not address by themselves.

During this exercise, we observed signs of mode collapse in the trained models. Mode collapse is a phenomenon where a GAN generator exploits a part of the dataset distribution, constraining its output to similar results. For instance, the model keeps generating orange rooms even if the input bubble diagram has no orange discs (Figure 6). The second problem was the generation of images with blurry edges and mixed colours, which complicates the extraction of floor plan geometry. This issue might be a result of various factors, such as the small size of the training dataset, the choice of loss function, or the hyper-parameter tuning.

Typically, the target audience of this course did not possess in-depth technical knowledge of DL, which is essential to understand how to overcome training problems. A trivial solution would be to extend the technical aspect of the module to address these methods so students could solve common DL problems and even evaluate the performance of multiple models for floor plan generation. However, this can be challenging in a course designed for architecture students. Another alternative is to embrace students from data science, engineering, and computer science to form interdisciplinary teams, which has been successfully tested in other courses at CMU, such as Art and Machine Learning (E. Kang et al., 2018).

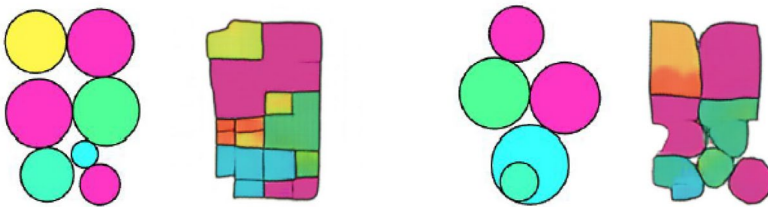


Figure 6. Two examples of mode collapse in training.

#### 4.4. INTERACTION AND INTEROPERABILITY

The use of Jupyter Notebooks allowed students to interactively run and inspect different functions and components of the pipeline, which includes the logic of the GAN model and learning algorithm. However, we observed issues in interoperability—the connection and exchange of information between the



development environment (Jupyter Notebooks) and the 3D modelling tools (Rhino 6/Grasshopper). These interoperability issues hinder the design process, as switching between these platforms during data generation, training, and inference phases is time-consuming and distracting.

#### 4.5. COMPUTING PLATFORMS AND PERFORMANCE

We used Jupyter Notebooks on Google Colab as an interactive and cloud-based development platform for B2F. On the one hand, this provided the same development environment and computation power for all students with close-to-no setup time across different sections of this course. On the other hand, this decision came at a cost; the free tier of Colab service relies on limited shared computing resources, which significantly slowed down the training and inference processes. In this setting, training our model takes over 10 hours, and the inference step takes a few seconds. This hampers students' capacity to investigate the affordances of the GAN model and, consequently, to take informed decisions and explore design alternatives. On higher-end GPUs, such as an RTX A6000, training the model with a relatively small data set for 100 epochs takes less than 20 minutes, and the inference step is executed in real time; however, this would be expensive for pedagogical purposes.

### 5. Conclusion and Discussion

The pedagogical experience with B2F establishes a deliberate position for the architect and DL. Instead of being end-users of a black-box design tool, the students are engaged in the different stages of B2F, including (a) data synthesis, (b) training the model, (c) translating the model output into proper design representations, and (d) human-AI interaction to explore design alternatives. Besides, DL is not only treated as a novel source for spatial sensibilities but also as a tool to explore solutions subject to real design, software, and hardware constraints.

By situating designers as active agents in the customisation and deployment of a pipeline for housing design, this experience revealed opportunities, frictions, and challenges for future practises based on DL. It is a step towards making DL accessible for architecture students to explore different design problems. It aims at providing the next generation of designers with the relevant skills and access to expert jobs and entrepreneurship in areas of AEC that will be drastically affected by AI. Moreover, it promotes the protagonist role of the profession in the technological development of the AEC industry and in the exploration of innovative design methods to pressing issues in design, such as the design of sustainable, safe, and affordable housing.

### Acknowledgements

We would like to express our gratitude to Prof. Ramesh Krishnamurti, Prof. Daniel Cardoso Llach and Prof. Omar Khan for their support in the ideation and development of the course. We would also like to thank Prof. Krishnamurti for reviewing the final version of this document. Finally, we would like to thank our students from the courses Learning Matters, Exploring Artificial Intelligence in Architecture and Design (Spring 2021), and Inquiries into Machine Learning & Design (Fall 2021).

## References

- ACADIA Workshops. (2020). *ACADIA 2020: Distributed Proximities*.  
<https://2020.acadia.org/workshops.html>
- ACADIA Workshops. (2021). *ACADIA 2021: Realignments: Towards Critical Computation*.  
<https://2021.acadia.org/workshops/>
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. *Proceedings of NIPS: Advances in Neural Information Processing Systems*, 2672–2680.
- Goodfellow, I., Yoshua, B., & Aaron, C. (2016). *Introduction to Deep Learning*. MIT Press.  
<http://www.deeplearningbook.org>
- Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2017). Image-to-Image Translation with Conditional Adversarial Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1125–1134).
- Joyce, S., & Nazim, I. (2021). Limits to Applied ML in Planning and Architecture: Understanding and defining extents and capabilities. Towards a New, Configurable Architecture. *Proceedings of the 39th eCAADe Conference* (pp. 243–252).
- Kang, E., Poczos, B., & Dinu, J. (2018). *Art and Machine Learning*. *ArtML2018*. Retrieved from <https://sites.google.com/site/artml2018/>
- Kang, H., & Jha, A. (2018, July 23). *Pytorch Implementation of Pix2Pix for Various Datasets*. GitHub. Retrieved from <https://github.com/znxlwm/pytorch-pix2pix>
- Liu, C., Wu, J., Kohli, P., & Furukawa, Y. (2017). Raster-To-Vector: Revisiting Floorplan Transformation. *Proceedings of the IEEE International Conference on Computer Vision* (pp. 2195–2203).
- Nauata, N., Chang, K.-H., Cheng, C.-Y., Mori, G., & Furukawa, Y. (2020). House-GAN: Relational Generative Adversarial Networks for Graph-Constrained House Layout Generation. In A. Vedaldi, H. Bischof, T. Brox, & J.-M. Frahm (Eds.), *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 162–177). Springer International Publishing. [https://doi.org/10.1007/978-3-030-58452-8\\_10](https://doi.org/10.1007/978-3-030-58452-8_10)
- Nauata, N., Hosseini, S., Chang, K.-H., Chu, H., Cheng, C.-Y., & Furukawa, Y. (2021). House-GAN++: Generative Adversarial Layout Refinement Network towards Intelligent Computational Agent for Professional Architects. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 13627–13636). <https://doi.org/10.1109/CVPR46437.2021.01342>
- Nourian, P., Rezvani, S., & Sariyildiz, S. (2013). A Syntactic Architectural Design Methodology: Integrating real-time Space Syntax analysis in a configurative architectural design process. In Y. O. Kim, H. T. Park, & K. W. Seo (Eds.), *Proceedings of the 9th International Space Syntax Symposium*.
- Sg2018 workshops. (2018). [Conference]. *Smart Geometry*.  
<https://www.smartgeometry.org/sg2018workshops>
- Tian, R., Guida, G., & Kim, D. (2021). Machine Intelligence in Architecture 2.0: The Architecture Turing Test [Education]. *Digital Futures*.  
<https://www.digitalfutures.world/workshops/74.html>
- Weinzapfel, G., & Negroponte, N. (1976). Architecture-by-yourself: An experiment with computer graphics for house design. *Proceedings of the 3rd Annual Conference on Computer Graphics and Interactive Techniques*, 10, 74–78.  
<https://doi.org/10.1145/563274.563290>

# ASYNCHRONOUS DIGITAL PARTICIPATION IN URBAN DESIGN PROCESSES

*Qualitative Data Exploration and Analysis with Natural Language Processing*

CEM ATAMAN<sup>1</sup>, BIGE TUNÇER<sup>2</sup> and SIMON PERRAULT<sup>3</sup>

<sup>1,2,3</sup> *Singapore University of Technology and Design, Singapore*

<sup>1</sup>*cem\_ataman@mymail.sutd.edu.sg, 0000-0001-8355-7119*

<sup>2</sup>*bige\_tuncer@sutd.edu.sg, 0000-0002-1344-9160*

<sup>3</sup>*simon\_perrault@sutd.edu.sg, 0000-0002-3105-9350*

**Abstract.** This paper aims to improve the usability of qualitative urban big data sources by utilizing Natural Language Processing (NLP) as a promising AI-based technique. In this research, we designed a digital participation experiment by deploying an open-source and customizable asynchronous participation tool, “Consul Project”, with 47 participants in the campus transformation process of the Singapore University of Technology and Design (SUTD). At the end of the data collection process with several debate topics and proposals, we analysed the qualitative data in entry scale, topic scale, and module scale. We investigated the impact of sentiment scores of each entry on the overall discussion and the sentiment scores of each introduction text on the ongoing discussions to trace the interaction and engagement. Furthermore, we used Latent Dirichlet Allocation (LDA) topic modelling to visualize the abstract topics that occurred in the participation experiment. The results revealed the links between different debates and proposals, which allow designers and decision makers to identify the most interacted arguments and engaging topics throughout participation processes. Eventually, this research presented the potentials of qualitative data while highlighting the necessity of adopting new methods and techniques, e.g., NLP, sentiment analysis, LDA topic modelling, to analyse and represent the collected qualitative data in asynchronous digital participation processes.

**Keywords.** Urban Design; Digital Participation; Qualitative Urban Data; Natural Language Processing (NLP); Sentiment Analysis; LDA Topic Modelling; SDG 10; SDG 11.

## 1. Introduction

Digitalization and new computational technologies have enabled citizen participation in urban design processes to become more inclusive by reaching more residents. Many digital participation tools have been implemented and applied in urban design

processes, that intend to go beyond receiving directed feedback from residents on a design proposal represented as rendered visualizations or site plans (Alter et al., 2019). Accordingly, some digital or analogue urban participation campaigns utilize map-based or 3D representations when interacting with residents (Tan, 2016; Tomarchio et al., 2019). These campaigns operate through residents participating in the urban design process by directly manipulating program or urban performance-related aspects of a curated urban proposal created by them or offered to them, usually in an engaging and collaborative digital environment, or a physical setting where verbal communication mediates the collaboration on the physical participation artefact. The aim of these processes is in line with the UN Sustainable Development Goals (SDGs): SDG 10, which promotes reducing inequality by encouraging social, economic, and political inclusion with equal opportunities and SDG 11, which enhances inclusive and sustainable urbanization in policy implementations and efficiency.

However, especially when urban designers are in an early stage of design, understanding the prominent values and goals of residents regarding the urban space is essential. This can be facilitated by highly effective means of citizen participation, such as discussions, debates, and proposals, all expressed in natural language in digital mediums. A variety of asynchronous digital participation tools with such functionality are deployed in urban practices, which allow a high number of stakeholders to participate together at each person's own convenience and own schedule in one medium to effectively discuss urban issues (Klein, 2011; Seifert & Rössel, 2019). The output of such participation processes is mostly qualitative data because of the predominant textual interaction between participants. Nevertheless, the qualitative textual data usually remains unused, which contains essential information regarding urban issues from participants' perspectives (Rathore et al., 2018). The reasons for this underutilization are, firstly, the challenges in filtering the collected data to understand the most valuable/useful information, and secondly, the lack of suitable representations and visualizations of the qualitative data in urban design processes. By considering these shortcomings in this research, we would like to facilitate the translation of collected qualitative participation data to be used by designers and decision makers in urban design processes.

This research aims to support urban designers with data analysis and informed decision-makings in the use of large-scale asynchronous digital participation tools that predominantly collect textual data. In this paper, we aim to investigate the way in which we can facilitate qualitative urban big data sources by utilizing AI-based techniques, i.e., Natural Language Processing (NLP), sentiment analysis, and LDA topic modelling to provide insights into qualitative participation data, which are mostly neglected in participatory urban design practices.

## **2. Background**

The increasing development in data processing and information technologies provides promising opportunities in urban design practices (Milz & Gervich, 2021). They encourage authorities to work on processes for residents and engage with them in important decisions and policies, which deepens the relationship between democracy and the vibrancy of civic life (Goldsmith & Crawford, 2014). In this context, digital participation tools and processes enable residents to collaborate with local governments

and provide the advantage of massive amounts of new information for effective and efficient solutions to urban problems (Ataman & Tuncer, 2022). As a general rule, these tools offer greater flexibility in time and location of data collection while promoting social learning for better collective decisions (Tekler et al., 2020). The use of mixed methods for quantitative and qualitative data collection and analysis has been particularly important in such participatory design practices (Lobe et al., 2020). Yet, especially for the qualitative data, the utilization of digital participation data is not fully covered in the urban design domain (Rathore et al., 2018).

It remains a challenge to enable the use of massive digital participation data and utilize qualitative data for informed urban design processes. Several researchers already focused on the possible ways of qualitative data analysis, e.g., mind mapping (Burgess-Allen & Owen-Smith, 2010), correspondence analysis (Habib et al., 2012), and very recently topic modelling (Lock & Pettit, 2020). Topic modelling is especially useful to organize and offer insights for stakeholders to understand large datasets consisting of unstructured text bodies. Yet, it becomes a challenge to interpret the semantic meaningfulness of topics by using such statistical measures of topical coherence (Chang et al., 2009). Therefore, the operationalization of a topic model according to a specific context and research setting is essential.

This study is unique as it deploys and combines various methods from other domains for analysing and visualizing participation data in different levels specifically for urban practices, e.g., NLP, sentiment analysis, LDA topic modelling. In the end, this paper contributes to the use of qualitative big data in asynchronous digital participation processes by providing information about:

- the interaction between and involvement of participants on different levels (i.e., module level, topic level, and entry-level) in a digital participation process,
- AI-based approaches for qualitative participation data to utilize and interpret participation outputs,
- new visualizations and representations to translate qualitative data into informed decision-makings.

Based on the above, this study concludes with several recommendations for the analysis of digital qualitative data in large-scale, individual, and asynchronous participation processes in urban design practices.

### 3. Methodology

This section presents the experiment conducted for collecting and analysing participation data in a campus transformation context. The process is explained in participation phases, i.e., pre-participation, participation, and post-participation.

#### 3.1. PRE-PARTICIPATION: TOOL SELECTION & CASE STUDY

Within the pre-participation phase, we chose and modified an open-source digital asynchronous digital participation tool, Consul Project, to implement it for a case study. The tool is written in Ruby on Rails using a PostgreSQL database for data storing. Consul Project provides several participation modules with different features and

interfaces, i.e., debates, proposals, voting, to discuss the case study (Figure 1). The case study in this experiment is the transformation project of SUTD campus. The campus is accessible by the subway, buses, and private cars and nearby several urban hubs such as a EXPO Convention & Exhibition Centre, Changi Business District, and an urban coastal park (ECP) while hosting a variety of facilities, such as a library, student housing, staff housing, sports and recreation facilities, and F&B and services to fulfil the needs of its daily users. The campus is also a part of a large-scale transformation project aiming at increasing the vibrancy of the East Coast area due to its location.

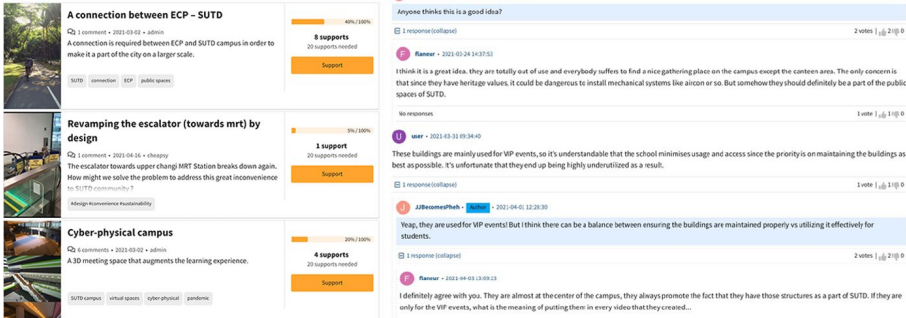


Figure 1. Consul Project Modules: Left: Proposal interface. Right: A discussion on a debate topic.

### 3.2. PARTICIPATION: PROCEDURE AND PARTICIPANTS

The 47 participants (40% female, 45% graduate and 24% architecture students) participated in the experiment voluntarily, and no incentives were used in this experiment. The selection criteria for the participants were, firstly, to be a student in the university, and secondly, to have completed their first mandatory design course to ensure a foundation in design theory and thinking. Prospective participants were sent an email that included their username, password, and user instructions, which allowed them to anonymously access the participation tool. After receiving the user account information and instructions, the participants were given two months to discuss and propose their ideas through the tool about the transformation of the university campus.

### 3.3. POST-PARTICIPATION: ANALYSIS METHODS

After the participation phase, the collected participation data was pre-processed to record timestamps of entries, parent arguments, involved participants, and support numbers. This process provided a hierarchical data structure to analyse the qualitative data. By using this pre-processed data, we investigated possible ways to visualize the interaction between participants and the evolution of debate topics and proposals. Then, we implemented AI-based NLP techniques, i.e., sentiment analysis and LDA topic modelling. We used these NLP methods as powerful computational techniques for data mining and formulating relationships among data and textual documents.

Sentiment Analysis determines the ratio of positive to negative engagements about a specific topic by analysing bodies of text, such as written discussions, comments, and reviews. It is used to obtain insights from opinions, attitudes, and emotions of individuals towards entities, e.g., topics, services, locations, or events (Liu, 2012).

Although sentiment analysis has been used in many fields such as politics, economics, business, and urban planning, the implementation of sentiment analysis in the participatory design domain is not fully discovered yet (Roberts et al., 2018). In this experiment, we used SentimentIntensityAnalyser from the NLTK library in Python, which generates sentiment scores ranging between -1.0 (negative) and 1.0 (positive) corresponding to the overall emotional leaning of the text.

LDA topic modelling, on the other hand, was deployed to classify texts in debates and proposals to a particular topic. It is an intuitive approach that calculates the similarity between source data to reveal their respective distributions of each cluster over topics (Jelodar et al., 2019). It provides the examination of multiple topics by generating a probabilistic distribution of words under a topic to extract thematic content from text-specific data (Wang & Taylor, 2018). In this experiment, the NLTK Python library was implemented to pre-process the data as it is one of the most common and powerful libraries for NLP and computational linguistics.

Before analysing the data, a basic data cleaning was conducted. As Consul Project allows its users to use words, URLs, mentions, abbreviations, etc., we cleaned the texts by removing stopwords, URL links, user mentions (@), and special characters (punctuations and numeric numbers) as they are unnecessary for further analysis. Then, we tokenized the entries and convert them to lowercase. The final data consists of 139 sentences with 1182 words.

#### 4. Results and Discussion

The analysis and interpretation of the collected data are discussed according to the methods that we described in the analysis methods for the post-participation phase. The goal is to provide effective and meaningful ways to analyse digital qualitative participation data to be used by other stakeholders.

##### 4.1. USER INVOLVEMENT & INTERACTION IN PARTICIPATION

After the data pre-processing, the user involvement in different modules is compared through the numbers of involved participants, entries, and supports. In the debate module, engagement is correlated with the number of entries and supports in each individual debate topic. In the proposal topics, on the other hand, only the support numbers represent the engagement as the participants mostly used the support feature of the module, and active participation cannot be measured through the entry numbers (Figure 2). The results confirmed the proper use of the participation modules by the participants according to their features, e.g., deliberation, feedback, supports, and likes.

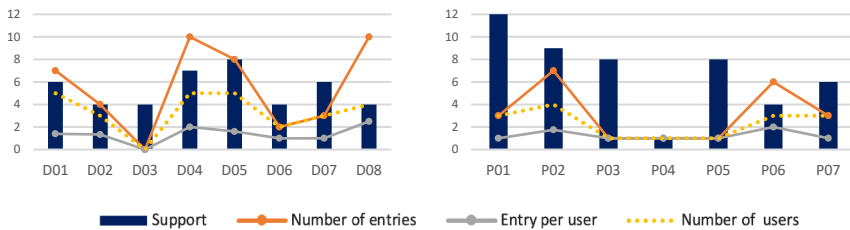


Figure 2. Number of entries and support for debate topics (Left) and proposals (Right).

After the identification of the most engaging debates and proposals, each entry in these modules is mapped according to their timestamps, their hierarchical relationship with other entries, like & dislike numbers that they received, and their sentiment scores on the entry level. The impact of each entry on the overall discussion is revealed by visualizing each discussion topic. Figure 3 presents the entry maps of Debate 4 (as an engaging discussion) and Debate 6 (as a rather inactive debate). As seen in Debate 4, the entries with positive sentiments led to an increased engagement and participation activity while the ones with negative sentiment values as in Debate 6 result in shorter discussions with limited interaction. The analysis results revealed that the initial attitude of an entry is determinant in the engagement of and the interaction between the participants. In order to identify such entries, entry maps and visualizations are essential to provide mediums to stakeholders to trace the most engaging and interacted arguments within the data set and evaluate them accordingly.

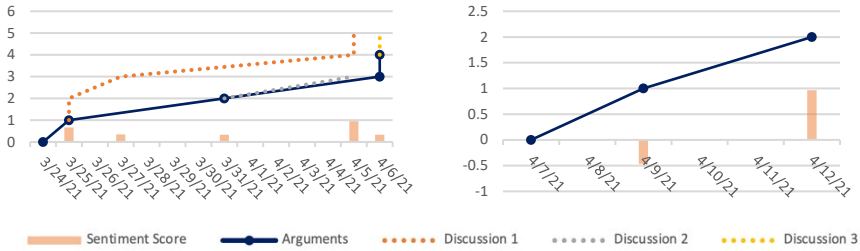


Figure 3. The analysis of Debate Topics: Arguments, discussions, and their sentiment scores (Left: Debate Topic #4, right: Debate Topic #6)

#### 4.2. THE IMPACT OF PARTICIPANTS' ATTITUDE ON PARTICIPATION: SENTIMENT ANALYSIS

In this participation experiment, debate and proposal topics contain introduction texts explaining the target issue to other participants to discuss and evaluate. Although the structure of these introductions is mostly very straightforward with texts and some images, some topics attracted more participants with longer discussions. In order to understand possible underlying reasons, we analysed the sentiment scores of introduction texts and the discussions under each debate and proposal separately. According to the results of the sentiment analysis, the introduction texts of the debate topics with positive sentiment scores (or positive attitudes) result in a higher number of entries; and therefore, more engagement (D1, D4, D5, D7, and D8 in Figure 4).

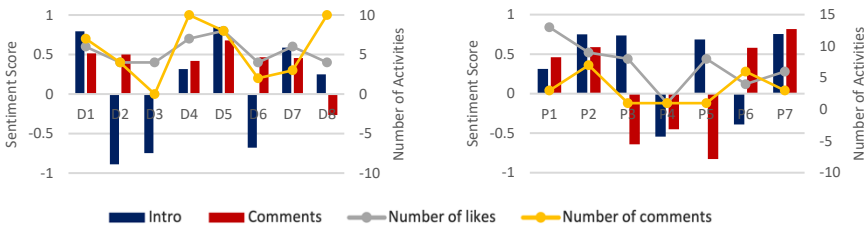


Figure 4. Sentiment analysis of textual data (Left: Debate module, right: Proposal module)



In the proposal module, on the other hand, the attitudes in the responses of the participants are more determinant. The most engaging proposal topics contain more responses with positive sentiment scores, which indicate the willingness of the participants when they like the proposal. In other cases, in which the engagement is relatively lower, negative sentiment scores are observed in the responses (P3, P4, and P5 in Figure 4). This analysis reveals the correlation between positive feedback and increased engagement in proposal topics as the support and comment numbers are in line with the attitude of the participants. As a result, the sentiment analysis of the qualitative data provided several key insights regarding digital asynchronous participation as presented below:

- Each participation module needs to be analysed separately as they contain different data types and interfaces.
- Sentiment analysis results indicate the need for triggering engagement with positive attitudes.
- Higher sentiment scores may indicate future engagement in digital participation, and therefore, these topics require further attention.

#### 4.3. VISUALIZATION OF QUALITATIVE PARTICIPATION DATA: LDA TOPIC MODELLING

In order to enable the use of qualitative data, a huge amount of textual data need to be visualized for decision-makers. In this experiment, the collected qualitative data is visualized through the LDA topic modelling algorithm, which estimates the topic prevalence distributions per entry and the term prevalence distributions per topic. The algorithm uses the corpus (i.e., the collected data consisting of the entries and replies) and the dictionary (i.e., an embedded wordlist in the NLTK library) to build the trained topics with frequent words and their weights. The weighting schema is a numerical statistic which reflects how important a term is to the discussion from the corpus of the experiment (Truica et al., 2016). In other words, the weights of the frequent words indicate their importance in a topic while analysing the data. This statistical measure primarily uses the frequency of a term in a discussion and the frequency of a term in the entire corpus.

The number of topics is the most important parameter to define. It is an ever-present concern to choose the best number of topics in topic modelling as well as in other latent variable methodologies. In this experiment, different numbers are tested in content generation by analysing the keywords of topics to see if they formulate meaningful topics. Eventually, the number 8 is adopted as the topic parameter by combining both statistical measures (weights) and manual interpretation (content). Eventually, we obtain 8 topics that resulted in the LDA topic model. Table 1 presents the LDA model built for one of these 8 topics (i.e., Topic 2) in our data set.

Table I. Most frequent words in Topic 2.

Topic # 2	0.038*"marketing" + 0.030*"budget" + 0.029*"think" + 0.028*"word" + 0.026*"spent" + 0.026*"year" + 0.019*"student" + 0.019*"faculty" + 0.017*"mil" + 0.015*"expense"
-----------	--

As seen in Table I, the most frequent words for the entries in topic 2 are "marketing" and "budget". Many of the entries also indicate the expenses of the university for marketing. Therefore, the topic can be named "financial management". However, the decision-makers need to observe all the topics and relevant terms together to interpret the collected data. Several visualization methods are deployed for topic modelling, including basic graphs for dominant words in word clouds, or more complex visualizations such as t-Distributed Stochastic Neighbour Embedding (t-SNE) clustering or intertopic distance map as the most referred visualization technique for LDA topic modelling. Figure 5 presents the intertopic distance map of the collected data (relevance metric ( $\lambda$ ) = 0.6) based on a multidimensional scaling algorithm. On the left, the area of circles is proportional to the number of words that belong to each topic across the dictionary. Meanwhile, the shorter distance between circles indicates topics that have more words in common. On the right, a bar chart presents the 30 most salient terms, and each bar indicates the total frequency of the term across the entire corpus. When a specific topic is selected as seen in Figure 5, the bar chart displays the most salient words included in that selected topic with a second (darker) bar that presents the topic-specific frequency of words that belong to the selected topic.

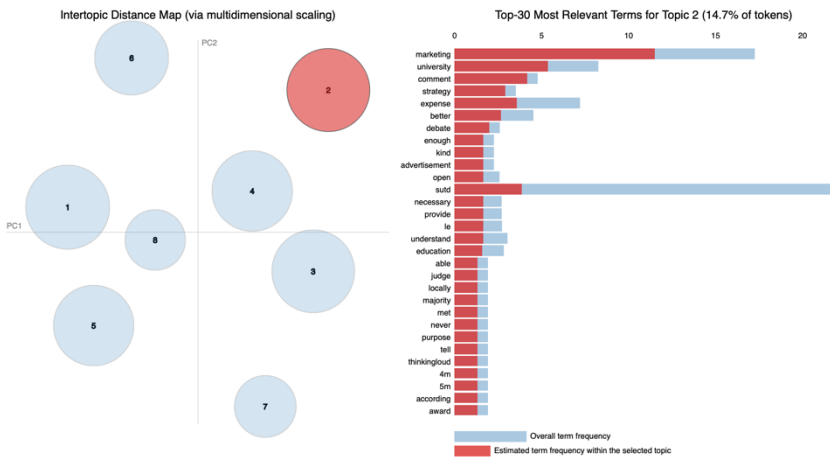


Figure 5. The intertopic distance map of the LDA topic modelling for this experiment ( $\lambda=0.6$ ).

The computational analysis and visualization methods are important to convert a large number of dimensions resulting from statistical methods, i.e., LDA topic modelling with an intertopic distance map in this case, to a reasonable number of dimensions. This study presents a valuable and novel way to filter the collected data and visualize textual information to enable decision makers to organize, understand and summarize large collections of qualitative textual participation data.

## 5. Conclusion

The consequences of the post-carbon cities require major digital shifts towards data collection in participatory urban design practices. In order to achieve a better collaboration for addressing global issues effectively, new data analysis methods are

necessary. These methods should support designers with informed decision making in the use of large-scale digital participation data, which contains a huge amount of qualitative data through discussions, debates, and proposals. In this study, we presented the results of an asynchronous digital participation experiment by analysing the textual data obtained from the participants. In order to facilitate the translation of textual data into decision-making processes, we explored several AI-based data analysis methods, i.e., NLP, sentiment analysis, and LDA topic modelling. Eventually, the long-term goal of this study is to provide systematic qualitative data processing and visualization on the entry level, module level, and topic level to support decision-makers in informed urban design practices. This study's focus on underutilized qualitative participation data might also be able to tackle the deployment of new computational methods and technologies in the analysis and visualization of qualitative data in digital asynchronous participation. This preliminary study already promises to transform the ways architects and designers utilize digital participation data in urban design contexts to increase the communication, interoperability, accountability, and quality of representations.

The next step of this study is to implement these techniques in larger data sets to further investigate the analysis of digital qualitative participation data. It will be beneficial to introduce the analysis and visualization methods to designers and decision makers to further study the use of qualitative participation data in early design phases in future works.

### Acknowledgements

This research is supported by Singapore University of Technology and Design (SGPCTRS1804) and Singapore International Graduate Award (SINGA). The authors would like to express their highest gratitude to all the participants, who participated in the experiment, for their collaboration and valuable opinions. The authors would also like to sincerely thank Doganalp Ergenc (University of Hamburg) for his contributions especially in the installation of the participation tool in pre-participation phases.

### References

- Alter, H., Whitham, R., Dawes, F., & Cooper, R. (2019). Learning by design. How engagement practitioners use tools to stretch the creative potential of their citizen participation practice. *13th International Conference of the EAD* (pp. 1387–1397). <https://doi.org/10.1080/14606925.2019.1594964>
- Ataman, C., & Tuncer, B. (2022). Urban Interventions and Participation Tools in Urban Design Processes: A Systematic Review and Thematic Analysis (1995 – 2021). *Sustainable Cities and Society*, 76, 103462. <https://doi.org/10.1016/j.scs.2021.103462>
- Burgess-Allen, J., & Owen-Smith, V. (2010). Using mind mapping techniques for rapid qualitative data analysis in public participation processes. *Health Expectations*, 13(4), 406–415. <https://doi.org/10.1111/j.1369-7625.2010.00594.x>
- Chang, J., Boyd-Graber, J., Gerrish, S., Wang, C., & Blei, D. M. (2009). Reading Tea Leaves: How Humans Interpret Topic Models. *Advances in Neural Information Processing Systems 22 - Pro- Ceedings of the 2009 Conference* (pp. 288–296).
- Goldsmith, S., & Crawford, S. (2014). *The responsive city: Engaging communities through data-smart governance*. Jossey-Bass. [https://books.google.com/books?id=NpI\\_BAAAQBAJ&pgis=1](https://books.google.com/books?id=NpI_BAAAQBAJ&pgis=1)

- Habib, F., Etesam, I., Ghoddusifar, S. H., & Mohajeri, N. (2012). Correspondence Analysis: A New Method for Analyzing Qualitative Data in Architecture. In *Digital Fabrication* (pp. 517–538). Springer Basel. [https://doi.org/10.1007/978-3-0348-0582-7\\_9](https://doi.org/10.1007/978-3-0348-0582-7_9)
- Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2019). Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications*, 78(11), 15169–15211. <https://doi.org/10.1007/s11042-018-6894-4>
- Klein, M. (2011). The MIT deliberatorium: Enabling large-scale deliberation about complex systemic problems. *ICAART 2011 - Proceedings of the 3rd International Conference on Agents and Artificial Intelligence*, 1(1). <https://doi.org/10.1109/cts.2011.5928678>
- Liu, B. (2012). Sentiment Analysis and Opinion Mining. *Synthesis Lectures on Human Language Technologies*, 5(1), 1–167. <https://doi.org/10.2200/S00416ED1V01Y201204HLT016>
- Lobe, B., Morgan, D., & Hoffman, K. A. (2020). Qualitative Data Collection in an Era of Social Distancing. *International Journal of Qualitative Methods*, 19, 160940692093787. <https://doi.org/10.1177/1609406920937875>
- Lock, O., & Pettit, C. (2020). Social media as passive geo-participation in transportation planning – how effective are topic modeling & sentiment analysis in comparison with citizen surveys? *Geo-Spatial Information Science*, 23(4), 275–292. <https://doi.org/10.1080/10095020.2020.1815596>
- Milz, D., & Gervich, C. D. (2021). Participation and the pandemic: how planners are keeping democracy alive, online. *Town Planning Review*, 92(3), 335–341. <https://doi.org/10.3828/tpr.2020.81>
- Rathore, M. M., Paul, A., Hong, W.-H., Seo, H., Awan, I., & Saeed, S. (2018). Exploiting IoT and big data analytics: Defining Smart Digital City using real-time urban data. *Sustainable Cities and Society*, 40, 600–610. <https://doi.org/10.1016/j.scs.2017.12.022>
- Roberts, H., Resch, B., Sadler, J., Chapman, L., Petutschnig, A., & Zimmer, S. (2018). Investigating the Emotional Responses of Individuals to Urban Green Space Using Twitter Data: A Critical Comparison of Three Different Methods of Sentiment Analysis. *Urban Planning*, 3(1), 21–33. <https://doi.org/10.17645/up.v3i1.1231>
- Seifert, A., & Rössel, J. (2019). Digital Participation. In D. Gu & M. E. Dupre (Eds.), *Encyclopedia of Gerontology and Population Aging* (pp. 1–5). Springer International Publishing. [https://doi.org/10.1007/978-3-319-69892-2\\_1017-1](https://doi.org/10.1007/978-3-319-69892-2_1017-1)
- Tan, E. (2016). *The Evolution of City Gaming* (pp. 271–292). [https://doi.org/10.1007/978-3-319-32653-5\\_15](https://doi.org/10.1007/978-3-319-32653-5_15)
- Tekler, Z. D., Low, R., Zhou, Y., Yuen, C., Blessing, L., & Spanos, C. (2020). Near-real-time plug load identification using low-frequency power data in office spaces: Experiments and applications. *Applied Energy*, 275, 115391. <https://doi.org/10.1016/j.apenergy.2020.115391>
- Tomarchio, L., Hasler, S., Herthogs, P., Müller, J., & Tunçer, B. (2019). Using an Online Participation Tool To Collect Relevant Data for Urban Design. *Proceedings of CAADRIA 2019 Intelligent & Informed* (pp. 747–756).
- Truica, C.-O., Radulescu, F., & Boicea, A. (2016). Comparing Different Term Weighting Schemas for Topic Modeling. *2016 18th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*, 307–310. <https://doi.org/10.1109/SYNASC.2016.055>
- Wang, Y., & Taylor, J. E. (2018). Urban Crisis Detection Technique: A Spatial and Data Driven Approach Based on Latent Dirichlet Allocation (LDA) Topic Modeling. *Construction Research Congress 2018*, 250–259. <https://doi.org/10.1061/9780784481271.025>

# A DATA-DRIVEN WORKFLOW FOR MODELLING SELF-SHAPING WOOD BILAYER

*Utilizing natural material variations with machine vision and machine learning*

ZUARDIN AKBAR<sup>1</sup>, DYLAN WOOD<sup>2</sup>, LAURA KIESEWETTER<sup>3</sup>,  
ACHIM MENGES<sup>4</sup>, THOMAS WORTMANN<sup>5</sup>

<sup>1,2,3,4,5</sup>*Institute for Computational Design and Construction (ICD),  
University of Stuttgart.*

<sup>1</sup>*zuardin.akbar@icd.uni-stuttgart.de, 0000-0002-8606-5483*

<sup>2</sup>*dylan-marx.wood@icd.uni-stuttgart.de, 0000-0003-0922-5399*

<sup>3</sup>*laura.kiesewetter@icd.uni-stuttgart.de, 0000-0001-5386-1224*

<sup>4</sup>*achim.menges@icd.uni-stuttgart.de, 0000-0001-9055-4039*

<sup>5</sup>*thomas.wortmann@icd.uni-stuttgart.de, 0000-0002-5604-1624*

**Abstract.** This paper develops a workflow to train machine learning (ML) models with a small dataset from physical samples to predict the curvatures of self-shaping wood bilayers based on local variations in the grain. In contrast to state-of-the-art predictive models, specifically 1.) a 2D Timoshenko model and 2.) a 3D numerical model with a rheological model, our method accounts for natural and unavoidable material variations. In this paper, we only focus on local grain variations as the main driver for curvatures in small-scale material samples. We extracted a feature matrix from grain images of active and passive layers as a Grey Level Co-Occurrence Matrix and used it as the input for our ML models. We also analysed the impact of grain variations on the feature matrix. We trained and tested several tree-based regression models with different features. The models achieved very accurate predictions for curvatures in each sample ( $R^2 > 0.9$ ) and extend the range of parameters that is incalculable by a Timoshenko model. This research contributes to the material-efficient design of weather-responsive shape-changing wood structures by further leveraging the use of natural material features and explainable data-driven modelling and extends the topic in ML for material behaviour-driven design among the CAADRIA community.

**Keywords.** Data-Driven Model; Machine Learning; Material Programming; Smart Material; Timber Structure; SDG 12.

## 1. Introduction

Natural materials like wood uniquely show anisotropic changes in shape in response to changes in moisture combined with impressive structural and ecological properties

(Rowell, 2005). The swelling and shrinking in wood occurs mainly perpendicular to the stiff cellulose microfibrils upon water uptake and loss, whereas the microfibrils prevent most swelling parallel to their orientation. This hygroscopic property can be utilized in "smart" weather-responsive shape-changing natural fibre composites materials such as a double-layered build-up of cross-laminated timber composite (bilayer) (Rüggeberg and Burgert, 2015) to create specific bending effects from flat to curved with various applications in architectural structures (Menges and Reichert, 2012; Wood et al., 2016; Wood et al., 2018). In bilayer, one layer serves as the actuator through shrinkage (active layer), while the other turns the motion into bending (passive layer).

A material behaviour-driven shaping method allows the construction of structurally efficient, form active, curved wood structures with minimal waste and labour, and little formwork or machining (Wood et al., 2020). However, the advantages of using wood as a self-shaping material come with the challenges of working with the imperfections of a naturally grown material and its often high levels of variation in structuring even under ideal conditions. The most influential parameters in predicting how wood hygroscopic shape changes are dependent on the arrangement of the fibrous structure and the orientation of the grains within the standardised board geometries.

In contrast to the material characteristics, methods of modelling self-shaping wood currently rely on simplified material models: either a 2D analytical model based on Timoshenko's theory (Rüggeberg and Burgert, 2015) or a combination of rheological models and a 3D non-linear numerical model (Grönquist et al., 2018). Even though these models were good at predicting the behaviour of a wood board, they always assume reduced variability on each wood board while, in reality, each wood board has a different range of properties and feature variations. These models also have a limitation on simulating only certain bilayer configurations and unaccountability for more complex designed variations in the layouts and layups of the laminates.

Besides the simplified modelling methods, the industry is capable of collecting highly specific and high-quality variable material features data from each wood board using machine vision technology (Olsson and Oscarsson, 2017; Wimmer et al., 2021), but the main focus is on grading and elimination (Ramage et al., 2017) which creates waste and material resource inefficiency (Barber et al., 2020). This acquired material features data opens the opportunity to better integrate material uncertainties as a material behaviour-driven agency in modelling and design.

Recent projects in computational design and fabrication use machine learning (ML) models to understand material uncertainties. Artificial neural networks predicted inconsistent spring back in robotic metal forming using a small dataset from physical samples (Zwierzycki et al., 2017; Rossi et al., 2019). Fragkia et al. (2021) trained a generative adversarial network to predict the geometry of self-shaping single-layer wood veneer strips but only focused on designed grain variations by the layup setup. Vasquez (2021) trained a deep convolutional neural network with paired images from serial events in a robotic concrete sheet forming to predict non-linear deformation behaviour. He et al. (2021) used a synthetic dataset to train a Gaussian mixture model convolutional operator that predicts the geometry of self-shaping textile.

We propose a data-driven workflow to predict the final shape of self-shaping wood bilayers using ML algorithms with a small dataset from physical samples to

complement the state-of-the-art predictive models. In this paper, we focus on local grain variations as the main driver of variable curvatures in small-scale wood bilayer samples made of European beech veneers. Figure 1 illustrates our main contributions in this paper as follows: 1.) an experimental method to build a dataset from wood bilayer samples, 2.) data augmentation and features generation for ML-ready data using machine-vision approaches, and 3.) the development of ML-based predictive models using various tree-based regression algorithms.

2. Methods

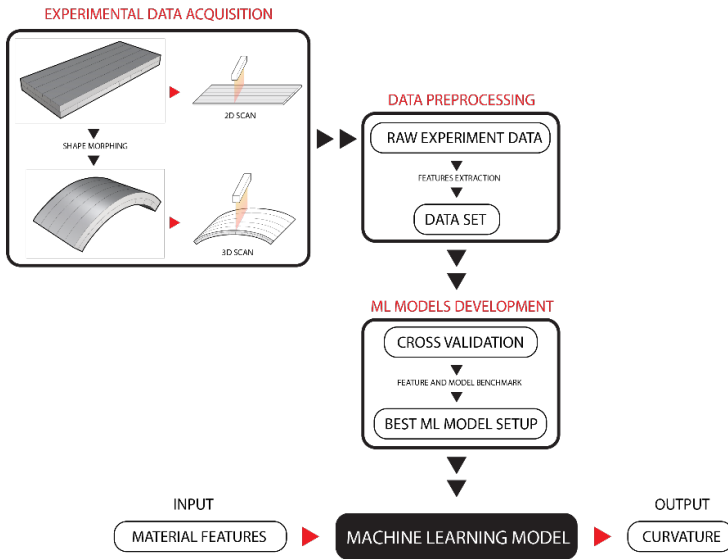


Figure 1. A flow chart of data-driven modelling for predicting self-shaping behaviour of wood bilayer. It includes bilayers sample production and data acquisition, data pre-processing, and ML models development.

2.1. SAMPLE PRODUCTION AND DATA ACQUISITION

We produced 30 small-scale wood bilayer samples using European beech veneers with visible natural variations in the longitudinal (L) axis of the sheets and variable growth ring inclinations. We equalized the active layer sheets inside a moisture-controlled chamber for adsorption at 95% RH to achieve a wood moisture content (WMC) of  $20 \pm 2\%$ . We used one thickness ratio of  $h_1 : h_2 = 1 : 1$  (passive : active) with a total thickness of 4mm. We laminated both layers perpendicularly using 1cPUR adhesive (HB S309 Purbond, Henkel & Cie. AG, Switzerland). We cut the sheets into smaller bilayer strips with a uniform size (width: 60 mm and length: 180 mm) and stored them again inside the moisture-controlled chamber for re-equalization.

We acquired data in both flat and curved states. We used a flatbed scanner to scan the surface grain of both active and passive layers right after removing the samples from the chamber. We relocated the samples to  $26 \pm 2\%$  RH and room temperature for actuation. We measured each sample's weight several times during the acclimatization

period (120 hours) to make sure that the samples are fully equalized. Finally, we 3D scanned each actuated sample using a laser scanner.

## 2.2. DATA AUGMENTATION AND PRE-PROCESSING

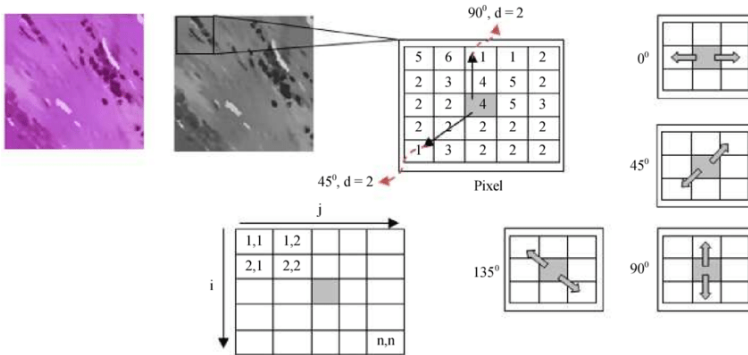


Figure 2. Illustration of GLCM with various angles ( $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ , and  $135^\circ$ ) and distance ( $d = 2$ ) orientation from the reference pixel (Haryanto et al., 2020).

To prepare ML-ready data, we did two consecutive steps to the physical samples data which are data augmentation and pre-processing. We augmented the data from 30 to 120 data points by flipping, mirroring, and rotating the grain images and 3D point clouds. We pre-processed both active and passive layer grain images to extract their features (Figure 2). We converted the grain images into greyscale images and enhanced the contrast using an exposure equalization algorithm. We extracted texture features from each image, known as Haralick features (Haralick & Shanmugam, 1973). Each feature is calculated as a Gray Level Co-occurrence Matrix (GLCM), a histogram of co-occurring grayscale values in neighbouring pixels at a given offset with a specific distance and angle over an image. In this paper, we used one distance (2 px) and three angles for each layer (active:  $0^\circ$ ,  $45^\circ$ , and  $135^\circ$  and passive:  $45^\circ$ ,  $90^\circ$ , and  $135^\circ$ ) to extract six main features: contrast, dissimilarity, homogeneity, ASM, energy, and correlation. We combined these processes in a custom python-based image analysis workflow using Scikit-image (van der Walt et al., 2014). Using Rhino and Grasshopper, we reconstructed each 3D point cloud into a surface to compute the principal curvature that is parallel to the L axis of the strip. For each feature, we built a dataset that contains six values calculated from different angles for both layers and the curvature of the bilayer strips

## 2.3. MACHINE LEARNING MODELS DEVELOPMENT

We first split the dataset into a training set (90 data points) and a test set (30 data points). We trained four tree-based regression algorithms: Random Forest (RF), Decision Tree (DT), Gradient Boosting (GB), and Extra Tree (ET) which are very good at learning from non-linear data (Lundberg et al., 2019) and understanding how the model uses the input features to predict by visualizing the decision path of the tree (Lundberg & Lee, 2017). Using six-fold cross-validation, we calculated the coefficient of



determination ( $R^2$ ), which is the degree of variability a model explains, to benchmark various combinations of input features and ML algorithms. The benchmark also helps to study the possibility of using a minimum amount of input features to reduce ML models' complexity. We used the implementation of these algorithms in Scikit-learn (Pedregosa et al., 2011).

We also calculated the curvature using the adaptation of Timoshenko's theory (Timoshenko, 1925) in Rüggeberg and Burgert (2015) which is subjected to a WMC change  $\omega - \omega_0$  where  $h_1$  and  $h_2$  denote layer thicknesses,  $\alpha_1$  and  $\alpha_2$  are the differential swelling coefficients taken from references (Hassani et al., 2015), and  $E_1$  and  $E_2$  are the stiffness in the corresponding fibre orientation of each layer, taken from Hering et al. (2012) (Equation 1). We calculated a range of curvatures using three different angles for Radial/Tangential (RT) grain ( $0^\circ$ ,  $45^\circ$ , and  $90^\circ$ ) to differentiate the material parameters of the active layer ( $E_2$ ,  $\alpha_1$  and  $\alpha_2$ ) since it was impossible to measure from the active layer.

$$\frac{1}{\rho} = \frac{6(1+m)^2}{(3(1+m)^2 + (1+mn)(m^2 + \frac{1}{mn}))} \frac{(\alpha_2 - \alpha_1)(\omega - \omega_0)}{h} = k \frac{\Delta\alpha\Delta c}{h}, m = \frac{h_1}{h_2}, n = \frac{E_1}{E_2} \quad (1)$$

Finally, we compared the prediction accuracy of the best ML model with the Timoshenko model to the curvature measured from physical samples.

### 3. Results and Discussion

Our GLCM implementation produced efficient and effective inputs for the ML algorithms. We managed to encode the local grain variations in each sample into a small number of vectors without over-simplification. Table 1 shows our analysis on the influence of grain variations on the values of six main Haralick features (contrast, dissimilarity, homogeneity, ASM, energy, and correlation) for different angles from the grain images of active and passive layers of each sample (active:  $0^\circ$ ,  $45^\circ$ , and  $315^\circ$  and passive:  $45^\circ$ ,  $90^\circ$ , and  $135^\circ$ ) after contrast equalization as shown in Figure 3. We used the data to identify the orientation of medullary ray spindles and growth rings in each strip. The data shows that the direction of medullary ray spindles is indicated by the highest values for homogeneity, ASM, energy, and correlation as well as the lowest values for contrast and dissimilarity. We also found that the deviations between the main angle variables for the GLCM indicate the orientation difference between the medullary ray spindles with the annual growth ring texture.

Figure 4 compares two samples with parallel and non-parallel relationships between the medullary ray spindles and the annual growth ring texture that are shown on the surface of the European beech veneer strips. Table 2 shows that the standard deviation in the data is smaller in non-parallel textures. This information may indicate the veneer sample's cutting orientation that can be advantageous for our research.

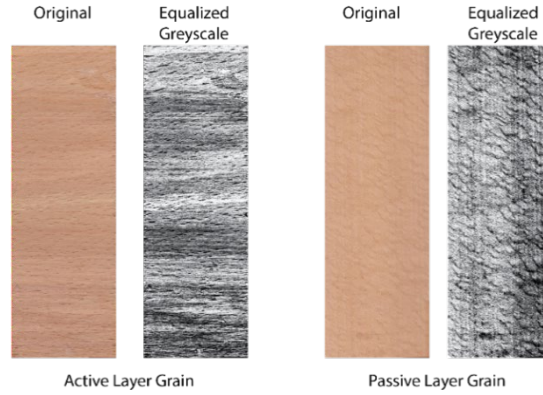


Figure 3. The active and passive layer grains from sample no.6 and the equalized greyscale images.

	Active			Passive		
	0°	45°	315°	45°	90°	135°
Contrast	<b>3354,672</b>	7543,229	7213,122	<b>5882,220</b>	6627,176	7202,329
Dissimilarity	<b>42,737</b>	67,839	66,331	<b>59,372</b>	63,448	66,708
Homogeneity	<b>0,039235</b>	0,018837	0,019352	<b>0,022476</b>	0,021648	0,019135
ASM	<b>3,3653E-05</b>	1,8177E-05	1,8521E-05	<b>2,0292E-05</b>	1,9653E-05	1,8167E-05
Energy	<b>0,005801</b>	0,004263	0,004303	<b>0,004504</b>	0,004433	0,004262
Correlation	<b>0,690379</b>	0,303462	0,333927	<b>0,455012</b>	0,387353	0,332715

Table 1. The Haralick features were calculated from sample no.6 in Figure 3. The bold values indicate the dominant angle in the grain.



Figure 4. (Left) Spindles formed by the medullary rays run mostly parallel to the annual growth ring. (Right) Spindles formed by the medullary rays running mostly at an angle of 315° to the annual growth ring.

Left			Right		
0°	45°	315°	0°	45°	315°
<b>5178,516</b>	8084,819	8464,274	8776,126	8683,538	<b>8148,223</b>
Std. Deviation: <b>1797,534</b>			Std. Deviation: <b>443,994</b>		

Table 2. The contrast from different angles and the data deviation from images in Figure 4.

Our benchmark using six-fold cross-validation shows that two algorithms (DT and ET) give very good  $R^2$  for the training set. ET with energy features as the inputs performs as the best model with the highest  $R^2$  of 0.902 (Figure 5). The results indicate that choosing an appropriate model and input features is important to capture the natural variations in wood grains that affect the curvatures. Figure 6 shows the distribution of predicted curvatures from the best model (ET with energy features as inputs) versus real curvatures measured from physical samples and curvatures calculated using a Timoshenko model. Using only visual input for the L grain directions, the data-driven model can predict the curvature of bilayers with a thickness ratio  $h_1 : h_2$  of 1 : 1 with higher accuracy than the Timoshenko model and such expands the range of thickness ratios that can be accurately predicted as the Timoshenko model usually over-predicts curvatures in this range (Grönquist et al., 2018).

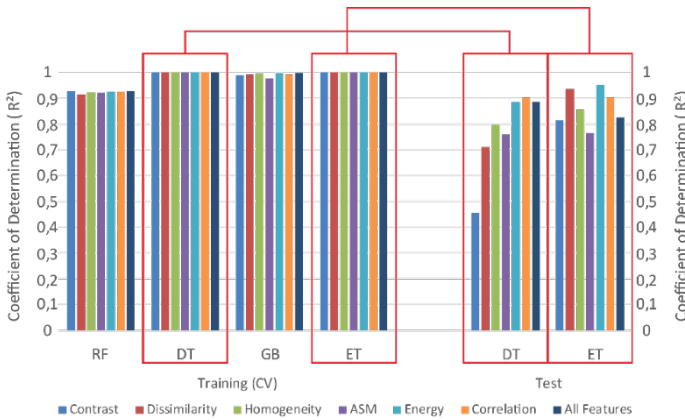


Figure 5. Coefficient of determination ( $R^2$ ) of various ML algorithms using a different set of features for both training set (evaluated using -fold cross-validation) and the test set  $R^2$  of the best models out of the cross-validation. Algorithms: Random Forest (RF), Decision Tree (DT), Gradient Boost (GB), and Extra Tree (ET).

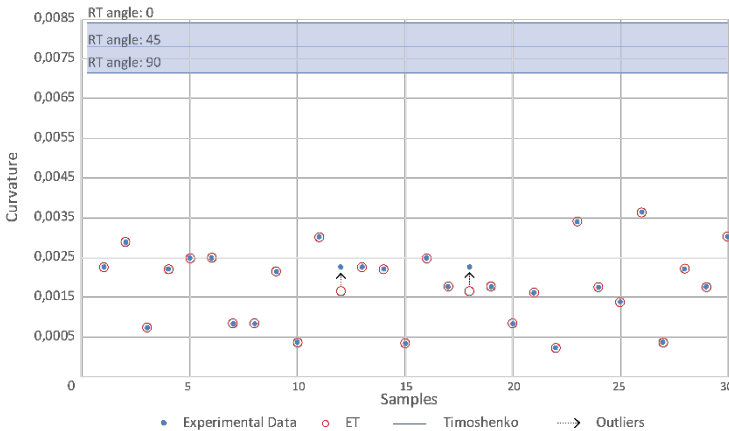


Figure 6. Distribution of predicted curvatures from the best model (ET with energy features as inputs) vs. the real curvatures vs. Timoshenko in three different RT angles.

#### 4. Limitations

We recognize some limitations in our current research that we would like to develop in the future. There is a limitation in material properties and scale as material behaviour changes when the scale changes. The effect of features might also be different across different scales. To work on this issue, we plan to test several techniques such as hybrid modelling that uses a data-driven model based on experimental data (high fidelity) to correct parameters in simulated models (low fidelity) or transfer learning.

Our current models are now limited to a narrow range of curvatures due to the constant bilayer configuration. We want to extend the curvature range in the dataset by adding more variations in thickness and aspect ratios. We also aim to develop models that can read, process, and return spatial information. We will test and implement various analysis techniques that can represent local variations in grain as 2D or 3D vector fields by processing images or computerized tomography scan data as inputs and return 3D information as the prediction. Such models will be more sufficient to be implemented as design tools for free-form wood structures, yet require bigger data set as the complexity of the models also expands. In this case, we will compare our future models with 3D numerical models to evaluate their performance.

#### 5. Conclusion

In conclusion, our paper reports a proof-of-concept for employing ML algorithms, a small dataset from physical samples, and machine vision techniques to model the hygroscopic behaviour of self-shaping wood bilayers in a way that incorporates the naturally inhomogeneous characteristics of wood. Despite focusing only on one wood bilayer setup, our models can work on a thickness ratio setup that was impossible for a Timoshenko model and establish a reliable relationship between the local grain variations and the curvature of self-shaping wood bilayers. The modelling of such relationships is difficult with state-of-the-art models, yet it is essential to produce more realistic predictions.

Using machine-vision-based features generation techniques such as a GLCM and tree-based regression models opens a new research horizon in the development of explainable Artificial Intelligence (AI)-based predictive models for designing self-shaping wood bilayer structures. Even though it is not yet further explored in this paper, tree-based regression models can provide decision paths that can help to explain how specific combinations of variable grains in wood boards produce certain curvatures.

We also envision technology-driven innovations in wood construction that promote more material resourcefulness and responsible material consumption by minimizing material loss due to the wood sorting process that only considers idealized wood variations. We plan to combine the techniques in this paper with combinatorial optimization algorithms to build a tool for designing a self-shaping wood structure that not only accounts for material feature variations but also material resource availability.

Finally, our research should be integral with the United Nation's Sustainable Development Goal 12 to ensure sustainable consumption and production patterns and complement previous research in ML for material behaviour-driven design within the CAADRIA community, specifically in data generation technique and ML model explorations.

## Acknowledgements

We acknowledge the support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2120/1 – 390831618.

## References

- Fragkia, V., Foged, I. W., & Pasold, A. (2021). Predictive Information Modeling: Machine Learning Strategies for Material Uncertainty. *Technology|Architecture + Design*, 5(2), 163–176. <https://doi.org/10.1080/24751448.2021.1967057>
- Grönquist, P., Wittel, F. K., & Rüggeberg, M. (2018). Modeling and design of thin bending wooden bilayers. *PLOS ONE*, 13(10), e0205607. <https://doi.org/10.1371/journal.pone.0205607>
- Grönquist, P., Wood, D., Hassani, M. M., Wittel, F. K., Menges, A., & Rüggeberg, M. (2019). Analysis of hygroscopic self-shaping wood at large scale for curved mass timber structures. *Science Advances*, 5(9), eaax1311. <https://doi.org/10.1126/sciadv.aax1311>
- Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 3(6), 610–621. <https://doi.org/10.1109/TSMC.1973.4309314>
- Haryanto, T., Pratama, A., Suhartanto, H., Murni, A., Kusmardi, K., & Pidanic, J. (2020). Multipatch-GLCM for Texture Feature Extraction on Classification of the Colon Histopathology Images using Deep Neural Network with GPU Acceleration. *Journal of Computer Science*, 16(3), 280–294. <https://doi.org/10.3844/jcssp.2020.280.294>
- Hassani, M. M., Wittel, F. K., Hering, S., & Herrmann, H. J. (2015). Rheological model for wood. *Computer Methods in Applied Mechanics and Engineering*, 283, 1032–1060. <https://doi.org/10.1016/j.cma.2014.10.031>
- He, X., Grassi, G., & Paoletti, I. (2021). Geometric Deep Learning: Prediction of Shape-Shifting Textiles. In *The 12th Annual Symposium on Simulation for Architecture and Urban Design (SimAUD): Human+*, SimAUD 2021. The Symposium on Simulation for Architecture and Urban Design (SimAUD).
- Hering, S., Keunecke, D., & Niemz, P. (2012). Moisture-dependent orthotropic elasticity of beech wood. *Wood Science and Technology*, 46(5), 927–938. <https://doi.org/10.1007/s00226-011-0449-4>
- Lundberg, S., & Lee, S.-I. (2017, November 24). A Unified Approach to Interpreting Model Predictions. In *The 31st Conference on Neural Information Processing Systems 2017*. The Conference on Neural Information Processing Systems (NeurIPS).
- Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S.-I. (2019). Explainable AI for Trees: From Local Explanations to Global Understanding. *ArXiv:1905.04610 [Cs, Stat]*. Retrieved January 10, 2022, from <http://arxiv.org/abs/1905.04610>
- Menges, A., & Reichert, S. (2012). Material Capacity: Embedded Responsiveness. *Architectural Design*, 82(2), 52–59. <https://doi.org/10.1002/ad.1379>
- Olsson, A., & Oscarsson, J. (2017). Strength grading on the basis of high resolution laser scanning and dynamic excitation: A full scale investigation of performance. *European Journal of Wood and Wood Products*, 75(1), 17–31. <https://doi.org/10.1007/s00107-016-1102-6>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12(85), 2825–2830.

- Ramage, M. H., Burrige, H., Busse-Wicher, M., Fereday, G., Reynolds, T., Shah, D. U., Wu, G., Yu, L., Fleming, P., Densley-Tingley, D., Allwood, J., Dupree, P., Linden, P. F., & Scherman, O. (2017). The wood from the trees: The use of timber in construction. *Renewable and Sustainable Energy Reviews*, 68, 333–359. <https://doi.org/10.1016/j.rser.2016.09.107>
- Rossi, G., & Nicholas, P. (2018). Re/Learning the Wheel. Methods to Utilize Neural Networks as Design Tools for Doubly Curved Metal Surfaces. In *ACADIA 2018: Recalibration. On Imprecision and Infidelity*, (pp. 146–155). The Association for Computer Aided Design in Architecture (ACADIA).
- Rowell, R. M. (Ed.). (2012). *Handbook of Wood Chemistry and Wood Composites (2nd ed.)*. CRC Press. <https://doi.org/10.1201/b12487>
- Rüggeberg, M., & Burgert, I. (2015). Bio-Inspired Wooden Actuators for Large Scale Applications. *PLOS ONE*, 10(4), e0120718. <https://doi.org/10.1371/journal.pone.0120718>
- van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., & Yu, T. (2014). scikit-image: Image processing in Python. *The PeerJ Bioinformatics Software Tools Collection*, 2, e453. <https://doi.org/10.7717/peerj.453>
- Vasques, A. N. (2021). Utilizing U-Net shaped networks as simulation tools in form-finding processes for fabric. In *The 12th Annual Symposium on Simulation for Architecture and Urban Design (SimAUD): Human+, SimAUD 2021. The Symposium on Simulation for Architecture and Urban Design (SimAUD)*.
- Wimmer, G., Schraml, R., Hofbauer, H., Petutschnigg, A., & Uhl, A. (2021). Two-Stage CNN-Based Wood Log Recognition. In O. Gervasi, B. Murgante, S. Misra, C. Garau, I. Blečić, D. Taniar, B. O. Apduhan, A. M. A. C. Rocha, E. Tarantino, & C. M. Torre (Eds.), *Computational Science and Its Applications – ICCSA 2021* (Vol. 12955, pp. 115–125). Springer International Publishing.
- Wood, D., Brütting, J., & Menges, A. (2018). Self-Forming Curved Timber Plates: Initial Design Modeling for Shape-Changing Material Buildups. In *The Annual Symposium of the International Association for Shell and Spatial Structures, IASS 2018*. The International Association for Shell and Spatial Structures (IASS).
- Wood, D. M., Correa, D., Krieg, O. D., & Menges, A. (2016). Material computation—4D timber construction: Towards building-scale hygroscopic actuated, self-constructing timber surfaces. *International Journal of Architectural Computing*, 14(1), 49–62. <https://doi.org/10.1177/1478077115625522>
- Wood, D., Grönquist, P., Bechert, S., Aldinger, L., Riggerbach, D., Lehmann, K., Rüggeberg, M., Burgert, I., Knippers, J., Menges, A. (2020). From Machine Control to Material Programming: Self-Shaping Wood Manufacturing of A High Performance Curved Clt Structure – Urbach Tower. In Burry, J., Sabin, J., Sheil, B., & Skavara, M. (Eds.), *Fabricate 2020* (pp. 50–57). UCL Press
- Wood, D., Vailati, C., Menges, A., & Rüggeberg, M. (2018). Hygroscopically actuated wood elements for weather responsive and self-forming building parts – Facilitating upscaling and complex shape changes. *Construction and Building Materials*, 165, 782–791. <https://doi.org/10.1016/j.conbuildmat.2017.12.134>
- Zwierzycki, M., Nicholas, P., & Ramsgaard Thomsen, M. (2018). Localised and Learnt Applications of Machine Learning for Robotic Incremental Sheet Forming. In K. De Rycke, C. Gengnagel, O. Baverel, J. Burry, C. Mueller, M. M. Nguyen, P. Rahm, & M. R. Thomsen (Eds.), *Humanizing Digital Reality* (pp. 373–382). Springer Singapore. [https://doi.org/10.1007/978-981-10-6611-5\\_32](https://doi.org/10.1007/978-981-10-6611-5_32)

# QUANTIFYING THE INTANGIBLE

*A tool for retrospective protocol studies of sketching during the early conceptual design of architecture*

JESSICA BIELSKI<sup>1</sup>, CHRISTOPH LANGENHAN<sup>2</sup>, CHRISTOPH ZIEGLER<sup>3</sup>, VIKTOR EISENSTADT<sup>4</sup>, ANDREAS DENGEL<sup>5</sup> and KLAUS-DIETER ALTHOFF<sup>6</sup>

<sup>1,2,3</sup>*Technical University of Munich.*

<sup>4,5,6</sup>*DFKI/University of Hildesheim.*

<sup>1</sup>*jessica.bielski@tum.de, 0000-0003-4936-1993*

<sup>2</sup>*langenhan@tum.de, 0000-0002-6922-2707*

<sup>3</sup>*c.ziegler@tum.de, 0000-0001-8095-416X*

<sup>4</sup>*viktor.eisenstadt@dfki.de, 0000-0001-6567-0943*

<sup>5</sup>*andreas.dengel@dfki.de, 0000-0002-6100-8255*

<sup>6</sup>*klaus-dieter.althoff@dfki.de, 0000-0002-7330-6540*

**Abstract.** Sketching is a craft supporting the development of ideas and design intentions, as well as an effective tool for communication during the early architectural design stages by making them tangible. Even though sketch-based interaction is a promising approach for Computer-Aided Architectural Design (CAAD) systems, it remains a challenge for computers to recognise information in a sketch. Design protocol studies conducted to deconstruct the sketch and sketching process collect solely qualitative data so far. However, the 'metis' projects aim to create an intelligent design assistant, using an artificial neural network (ANN), in the manner of Negroponte's Architecture Machine. By assimilating to the user's idiosyncrasies, the system suggests further design steps to the architect to improve the design decision making process for economic growth, qualitative self-education through the dialogue and reducing stress. For training such ANN quantitative data is needed. In order to produce quantifiable results from such a study, we propose our open-source web-tool 'Sketch Protocol Analyser'. By correlating different parameters (i.e. video, transcript and sketch built) through the same labels and their timestamps, we create quantitative data for further use.

**Keywords.** Design Protocol Studies; Sketching; Data Collection; Architectural Design Process; ANN; SDG 3; SDG 4; SDG 8; SDG 9.

## 1. Introduction

Sketching is a craft supporting the development of ideas and design intentions

(Lawson, 2005) as well as an effective tool for communication during the early architectural design stages by making them tangible. Therefore, sketch-based interaction is a promising method for an intuitive interaction with a Computer-Aided Architectural Design (CAAD) system that naturally integrates into the design process (Leclercq and Juchmes, 2002). Early attempts to enable sketch-based interaction with information systems date back more than half a century (Sutherland, 1964). The advancing digitisation of the construction industry and the associated growing data volumes increased the use of state-of-the-art approaches of artificial intelligence (AI), such as Case-based Reasoning (CBR) and deep learning (DL). Nevertheless, up to this day the formalisation of building on the computer usually focuses on fully designed buildings, while sketches of the early conceptual design are hardly considered. Thus, it remains a challenge for computers to recognise information in a sketch that is relevant for any given application (Johnson et al., 2009), implying a limited database of sketches for training the latest technologies like artificial neural networks (ANNs). Similar to the computer, researchers cannot fully fathom the content of sketches, as the hand drawings need interpretation and every designer or architect uses their own language.

Using retrospective protocol studies, hand drawings, their underlying intentions and ideas can be formalised to differentiate variants and individual design steps on a microscopic level. It is necessary for an AI to be able to understand this vague and imprecise character of hand drawings (Johnson et al., 2009), however, the requirements for understanding need to be formalised. While previous protocol studies for architectural designing through hand drawings produce qualitative results (Suwa and Tversky, 1997), the methods of the 'metis' projects require the collection and processing of quantitative data to train an ANN. With hand drawings and the correlating design decisions, a case base is created to train an ANN which suggests further design steps to architects similar to text auto-completion on a smartphone. The ANN supports the user by introducing personalised predictions and thus, reacts to the idiosyncrasies of the user (Negroponte, 1973). Thus, the architect can focus on the creative aspects while increasing economic growth through speeding up the process and increasing architectural quality, and supporting mental health through reducing work-related stress. Further, it supports the architect in the life-long learning of creating their own 'Guiding Design Principles' (Lawson, 2004, pp. 112, 113), as the ANN learns from its specific user and the architect widens their cognitive horizon through the suggestions.

In this paper we explore if and what kind of elements can be found within sketches of early conceptual design and further, if back-referencing these elements to the current design activity, derived from design theory (Laseau, 2000; Lawson 2004; Lawson 2005), is enabled. Thus, we propose a workflow for the data collection of retrospective protocol studies of hand drawings. This includes an approach for pre-processing the raw data of the study sessions and a novel tool created for obtaining quantifiable data.

## 2. Related Work

Hand drawings allow designers to visualise, review, revise (Buxton, 2007) and communicate ideas. Laseau (2000) describes sketches as a container, hub and incubator of vast and different kinds of information by providing interactive feedback for the architect during the design process, as well as a way for communication, individually, within the team and with the public. The internal discourses - also observed by



Goldschmidt (1994) - is also used for advancing and improving the design through reflection, as a 'reflective practitioner' (Schön and Wiggins, 1992). The cognitive scientist Goel (1995) focuses on the imprecise character of hand drawings, identifying their ambiguity as deliberately leaving room for interpretation and thus, a medium for self-inspiration. These mental activities behind designing are clearly evident through the progression throughout the sketch. Laseau (2000) and Lawson (2004, 2005) divide this progress into different design activities of an iterative process without a predetermined order.

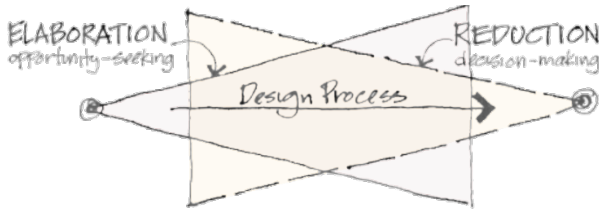


Figure 1. Koeberg and Bagnall's concept of the design process (Adapted from Laseau, 2000, p. 115)

Laseau (2000, p. 115) describes the general design process consisting of two main stages: the 'Elaboration', during which the designer maximises the data of possible solutions 'seeking opportunity', whereas the 'Reduction' aims to narrow down the notable references and applicable solutions for the specific design problem until the best possible solution is found through 'decision-making' (see Figure 1). These references - or also referred to as architectural precedents - for 'seeking opportunity' and 'decision-making' are used for both phases. They are used as a source of inspiration, design conditions and explicit information during the 'Elaboration', as well as a tool for evaluation during the 'decision-making', both phases being accommodated by reference building as a medium of communication (Richter, 2010, p. 154).

Lawson (2005, pp. 48, 49) elaborates the design process further by denying a chronological component. Rather he describes it as a combination of different building blocks fitting perfectly into each other (see Figure 2). According to the author, the design process is divided into 'analysis', 'synthesis' and 'evaluation' being the connection between the corner stones of the 'Problem' and the 'Solution'. Thus,

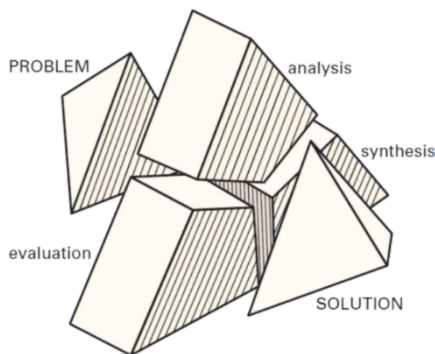


Figure 2. The design process seen as a negotiation between problem and solution through the three activities of analysis, synthesis and evaluation (Adapted from Lawson, 2005, p. 49)

Lawson (2005) concludes that every solution based on a re-iterative design process is uniquely created for the current design problem. The three blocks of the repetitive process - 'analysis', 'synthesis' and 'evaluation' - can similarly be found in Laseau's (2000, p. 13) detailed description of the design process, who defines five categories: 'analysis', 'exploration', 'discovery' and 'verification' with constant 'communication' throughout all of them. Both authors assign their respective categories to certain mental activities during the designing (Laseau, 2000; Lawson, 2005). Vice versa, these mental activities which are also referred to as the 'intention behind' (Lawson, 2004, pp. 17, 18), are identified through the currently sketched actions (e.g. modification, deletion) and elements (e.g. lines, symbols). Their combination, referred to as 'design step' within this paper and as 'design event' by Lawson (Ibid.), is an indicator for the ongoing design phase. Lawson (2005) specifically looks into sketching based on 'Graphic Thinking' (Laseau, 2000) and its external and internal dialogue. This project builds on the premise that these mental activities of the design process are generally comprised and showcased through protocol studies (Lawson, 2004, p. 15).

Design protocol studies are generally divided into 'Thinking aloud' studies and retrospective (reporting) studies. Both protocol study types deal with the problem of an artificial setting (e.g., without client and stakeholder communication, design briefings, and the general unfamiliar workspace), within a limited work environment (i.e. time limitation, limited access to other sources of knowledge) (Lawson, 2004, p. 16). However, Suwa and Tversky (1997), as well as Lawson (2004) identify a high amount of distortion of the natural workflow during the design process for the first type. '[T]alking aloud may adversely interfere with participants' perceptions during their sketching activities' (Suwa and Tversky, 1997, p. 386), because 'it introduces a degree of introspection and conscious attention to process that is abnormal [producing verbalization as] ... a director of thinking that takes place as a consequence' (Lawson, 2004, p. 16) rather than true design process thoughts. In order to minimize these negative effects, retrospective reporting is employed. In order to mitigate 'selective recall' during the report of the design process, Suwa and Tversky (1997, pp. 397) offer the participants the recorded video of their sketch. Through the visual cues 'about the exact sequence of sketching, including, the timing, hesitations, returns and redrawings' (Suwa and Tversky, 1997, p. 398) the participant is guided again through their own design process. Finally, these retrospective protocol studies allow for hand drawings, their underlying intentions and ideas to be formalised to differentiate variants and individual design steps on a microscopic level (Lawson, 2004, pp. 15-17). However, protocol studies are interpreted for further use resulting only qualitative results, similarly to Suwa and Tversky (1997). Even though every design process is truly unique - because of the problems, tools and designers themselves - Lawson (2004, p. 16) sees the potential and possibility for reproducible results of protocol studies in the manner of the natural science paradigm without the attempt to characterise designing.

### 3. Problem Statement

For an AI-based method, it is necessary to be able to understand the vague and imprecise character of hand drawings (Johnson et al., 2009), however, the requirements for understanding need to be formalised. While previous protocol studies for architectural designing through hand drawings produce qualitative results (Suwa and

Tversky, 1997), the methods of the 'metis' projects require the collection and processing of quantitative data to train an artificial neural network (ANN). With hand drawings and the correlating design decisions as a basis, a case base of quantitative data needs to be created to train an ANN which suggests further design steps to architects similar to text auto-completion on a smartphone.

4. Approach

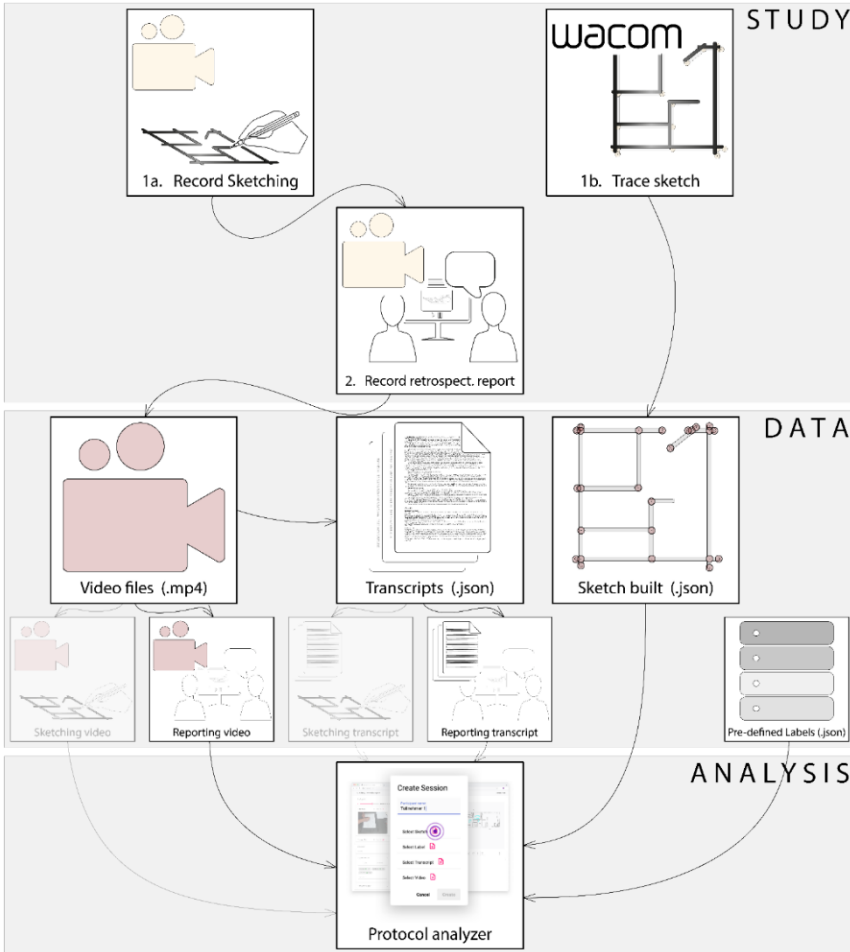


Figure 3. Data collection, data output and data analysis (greyed out: Data added after testing)

In order to obtain quantitative data, the collected data has to be quantifiable. Looking into the process of sketching, we observed two parameters that can be processed: the sketch itself and the verbalization of a retrospective study. Further, those two are related through time correlation. Thus, we propose the novel approach of our open-source web-tool called the 'Sketch Protocol Analyser'. Within this tool we introduce another correlation method of sketch and verbalization: custom labels.

Within our process (see Figure 3), we propose three main categories: the study, the data and the analysis. During the study, the data is collected in two main stages: the first round of the interview during which the sketching is recorded using a pivoting 4k camera and the sketching is traced by drawing on a sheet of paper on top of a WACOM tablet (see Figure 4). This trace of the sketch includes polygon points for creating lines, as well as other parameters like time and pressure. During the second round of the study, the study participants are presented with the sketch video on a computer screen. While they are retrospectively reporting on their thoughts and internal dialogue, another video is being recorded introducing the time correlation. All of this data is being saved in its respective file formats using the session label created at the beginning of the study session.

Thus, we collected video files (.mp4) through the recordings and a sketch built (.json) through the WACOM tablet and accompanying app for 'Real Time Ink'. The video of the retrospective reporting is further asynchronously processed into transcripts using Google Speech-to-Text services. This transcript includes speaker tags (i.e., interviewer and interviewee) and time stamps for each transcribed word, which is reformatted befitting the .json file format.

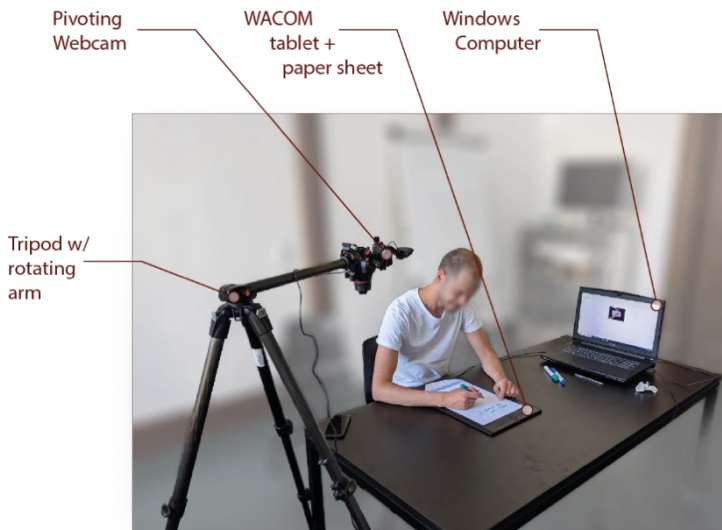


Figure 4. Set-up of the study (from left to right): Tripod with a pivoting 4k camera, paper on top of a WACOM tablet and WACOM fine-liner pen; Computer for recording and saving data

Finally, the video recording of the retrospective report, its transcript, the sketch built, and the custom labels are used as an input to the web-tool 'Sketch Protocol Analyser' for creating an analysis session with a personalised name (see Figure 3).

The study set-up (see Figure 4) and workflow were tested with two volunteering working architects. After studying the prepared design task, they were given 15 minutes to design without any specifications on style, detail, scale, or desirable final state of the sketch. Even though one of the participants was silent during the sketching, the other one verbalised the thoughts during drawing, however not in a 'Thinking Aloud' manner. It was rather short remarks during the reflection on the previously

taken design activities. Further, the desire for pausing the video during the reporting became apparent if a certain design element needed longer explanation.

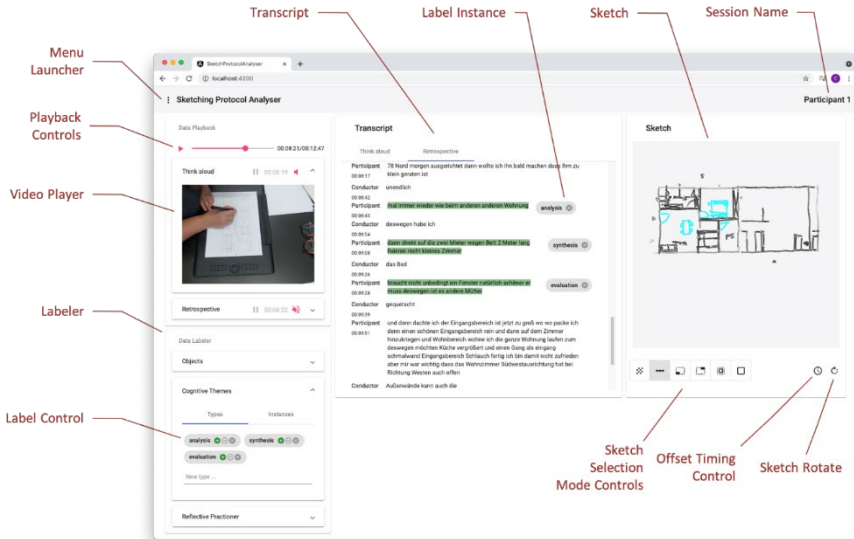


Figure 5. Interface of 'Sketching Protocol Analyser 1.2' (Adapted from Ziegler, 2021)

Figure 5 shows the graphic user interface of a sketch session of the 'Sketch Protocol Analyser 1.2', including the video recording of the sketching and its transcript, for the researcher to analyse and produce quantifiable results. On the left-hand side, the user can switch between the two videos or watch both at the same time. As soon as the videos are started the transcripts and the sketch is built. Using the buttons on the individual labels of the 'Data Labeler', the labels can be added to or removed from marked parts of the transcripts and sketches. As it can be seen in Figure 5, the created categories are: 'Objects' (e.g. 'wall'), 'Reflective Practitioner' (e.g. 'practice'), and 'Cognitive themes' (e.g. 'analysis'). For marking parts of the sketch multiple selection options are included, such as selection rectangles by dragging diagonally, individual polygon lines and semi-automatic line selection. All created files are automatically saved. The labelled components are saved as linked instances with IDs within the transcript files and the sketch built, which are referenced for each label and label category in the 'Label' .json file.

For our purposes we adapted the design phases of Lawson (2005): 'analysis', 'synthesis' and 'evaluation' within the label category 'Cognitive themes' (see Figure 5). Actions and their respective elements (e.g., 'outlining parcel') within the transcript and sketch are selected, which leads to the identification of the design intention behind (e.g. 'requesting/requested information', i.e. visualisation for seeing ratio/dimensioning/..., for 'outlining parcel'). Thus, the label of the according design phase (e.g. 'analysis' for 'requesting/requested information') is used to highlight these parts within the transcript and drawing. We aim to train the ANNs of the 'metis' projects, using this labelled data, to suggest actions and elements (also referred to as 'further design steps') depending on

the recognised design phase.

## 5. Discussion and Future Work

Through our web-tool ‘Sketch Protocol Analyser’ we were able to produce quantitative data from design protocol studies. The design process is recorded and labelled in order to identify actions and elements as design steps (e.g. ‘outlining parcel’) and to derive design phases (e.g. ‘analysis’) from indicated design intentions (e.g. ‘requesting/requested information’). This data produced from the ‘Sketch Protocol Analyser’ is suitable for further use within our deep learning (DL) pipeline for training the neural network, as its model is specifically created for using these data formats. This ANN will be trained to suggest further design steps, a combination of actions and elements, depending on the recognised design phase.

The modifications of the web-tool, after testing the set-up and analysis process, seem to have made a difference in usability and integration of all possible observations. Thus, all the data collected within the sketching process can be taken into account during the analysis.

However, for this workflow the supervisor of the protocol study still needs to interpret the data and manually assign the labels to the text within the transcript and sketch. Especially identifying the intention (e.g. ‘requesting/requested information’) behind taken design steps (e.g. ‘outlining parcel’) – even when taking the report by the participant into account – and the derived design phases (i.e. ‘analysis’, ‘synthesis’, ‘evaluation’) cannot be considered fully established and may also vary from researcher to researcher due to interpretation. Further, all design and sketch protocol studies are set in an artificial environment, as previously mentioned. They neglect briefings by the client, limited access to other sources of knowledge, and time and resource limits while creating in an unfamiliar work environment. Retrospective reporting is prone to distortions introduced by the ‘post-hoc rationalisation of their processes. ... [while restricting the time] to allow for reflection and time to pass’ (Lawson, 2004, p. 16). The reflection of the architect as a ‘reflective practitioner’ (Schön and Wiggins, 1992) happens both during the process and after a prolonged time period.

As a next step a large-scale study has been scheduled for December 2021. The analyser web-tool is used for producing quantifiable data customised for our ANN models, which are currently being tested. Due to the global pandemic the study has not been completed yet and more appointments have been set for 2022, after employing safety adjustments for the study sessions (e.g. interviewer and interviewee in two separate rooms) to comply with the according restrictions and rules. First tests with our artificial neural network, using the labelled data of conducted sessions, show promising results of predicting the first consecutive design phase. We plan to present our complete study results at the International Conference on Computational Creativity (ICCC) 2022, as well as aim to establish our labelling approach of design phases for elements and actions as ‘design steps’ at the eCAADe 2022 in Leuven.

## 6. Conclusion

Our system enables us to produce quantitative data from design protocol studies, while previous studies resulted only in qualitative data. By adding customised labels, the user

creates further parameters to be quantified, additionally to behavioural parameters e.g., timing, speed and pressure of the sketching process. Furthermore, the customisation of the labels and its open-source model gives other researchers the opportunity to transfer, translate and create specific labels for their needs.

This quantifiable data from sketch protocol studies allows us to further advance to train ANNs, suggesting further design steps, within the 'metis' projects. The intelligent design assistant supports architects by speeding up their process, allowing them to focus on their creative work, while maintaining or even increasing architectural quality.

Nevertheless, we aim for a better understanding of the role and process of designing and sketching to improve the interaction between architects and their technological support systems of the early design stages. We hope to inspire other researchers to further investigate the design process and its implications to improve the work environment for architects for increasing work quantity and meet the demands on the architectural quality, facilitating growth and quality in education, while considering the importance of reducing stress for improving architects' mental health.

### Acknowledgements

We express our deepest gratitude for the two testers of the study set-up for offering their valuable time and feedback, despite their full schedule. Further, we want to thank the DFG for funding the 'metis' projects.

### References

- Buxton, B. (2010). *Sketching user experiences: Getting the design right and the right design*. Morgan Kaufmann.
- Goel, V. (1995). *Sketches of thought*. MIT Press.
- Goldschmidt, G. (1994). On visual design thinking: the vis kids of architecture. *Design Studies*, 15(2), 158–174.
- Johnson, G., Gross, M. D., Hong, J. & Do, E. Y.-L. (2009). Computational support for sketching in design: a review. *Foundations and Trends in Human-Computer Interaction*, 2(1), 1–93.
- Laseau, P. (2000). *Graphic Thinking for Architects and Designers*. Wiley, USA.
- Lawson, B. (2004). *What Designers know*, 1st edition. Elsevier, Oxford.
- Lawson, B. (2005). *How Designers Think*, 4th edition. Elsevier, Oxford.
- Leclercq, P. & Juchmes, R. (2002). The absent interface in design engineering. Artificial Intelligence for Engineering Design, Analysis and Manufacturing. *AI EDAM*, 16(3), 219–227.
- Negroponte, N. (1973). *The architecture machine: Toward a more human environment*. The MIT Press.
- Richter, K. (2019). *Augmenting Designers' Memory: Case Based Reasoning in der Architektur*. (Doctoral Dissertation, Bauhaus-Universität Weimar).
- Schön, D., and Wiggins, G. (1992). Kinds of seeing and their functions in designing. *Design Studies*, 13(2), 135–156.
- Sutherland, I. E. (1964). Sketchpad a man-machine graphical communication system, *Simulation*, 2(5), 329-346.
- Suwa, M and Tversky, B 1997, 'What do architects and students perceive in their design sketches? A protocol analysis', *Design Studies*, 18(4), 385-403.
- Ziegler, C. (2021, September 21). *Sketch Protocol Analyser*. GitHub. Retrieved November 18, 2021, from <https://github.com/KSD-research-group/sketch-protocol-analyser>.





# **Urban Analytics, Urban Modelling and Smart Cities**



# EVOLUTIONARY DESIGN OF RESIDENTIAL PRECINCTS

## *A Skeletal Modelling Approach for Generating Building Layout Configurations*

LIKAI WANG<sup>1</sup>, PATRICK JANSSEN<sup>2</sup> and KIAN WEE CHEN<sup>3</sup>

<sup>1,2</sup>*National University of Singapore.*

<sup>3</sup>*Princeton University.*

<sup>1</sup>*wang.likai@outlook.com, 0000-0003-4054-649X*

<sup>2</sup>*patrick@janssen.name, 0000-0002-2013-7122*

<sup>3</sup>*chenkianwee@gmail.com, 0000-0003-4397-7324*

**Abstract.** This paper presents a ‘skeletal’ parametric schema to generate residential building layout configurations for performance-based design optimization. The schema generates residential building layout configurations using a set of ‘skeletal’ lines that are created based on various design elements and coincident with factors such as walkways, spacing, and setback requirements. As such, the schema is able to generate diverse and legitimate design alternatives. With the proposed parametric schema, a case-study optimization is carried out for a Singapore Housing Development Board (HDB) project. The case study considers a set of performance criteria and produces results with higher practical referential value. The case study demonstrates that the optimization with the parametric schema can improve the overall quality of the design and provide designers with various design options.

**Keywords.** Parametric Modeling; Building Layout; Performance-based Design; Algorithmic Design; Design Optimization; SDG 11.

## 1. Introduction

In Singapore, 80 percent of the population lives in public housing precincts, constructed by the Housing Development Board (HDB) (Cheong, 2019). The design and construction of these public housing precincts can have a great societal and environmental impact on sustainable city development. In this regard, novel and efficient design methods and techniques are being explored for improving the overall performance of HDB precinct designs with respect to various liveability criteria. Over the past decades, a number of studies have been conducted to integrate techniques such as parametric design and computational optimization into the design for HDB precincts (Chen et al., 2017; Chen & Norford, 2017; Janssen & Kaushik, 2013; von Richthofen et al., 2018). These studies demonstrate the potential of computational optimization and performance-based design in assisting architects in their design of HDB precincts. Despite this, there are still certain barriers making it difficult to apply computational optimization to HDB precinct designs.

Considering real-world design scenarios, a decisive factor in the utility of evolutionary optimization workflows is their ability to incorporate a set of constraints relating to various design rules for the precinct typology (Wang et al., 2018a, 2018b). The challenge is therefore to develop generative algorithms that are capable of generating designs with both adequate variability and feasibility. For variability, design variants need to differ significantly. For feasibility, design variants need to adhere to the constraints. This challenge also explains why it is difficult to use existing design optimization workflows for HDB projects in practice.

In these existing generative and optimization workflows, the parametric schema for design generation is often either over-constrained or under-constrained. The over-constrained case results in a lack of designs that vary significantly. Using this approach, buildings are typically placed over a predefined grid, thus resulting in a barracks-like building layout configuration (Chen & Norford, 2017; von Richthofen et al., 2018). In contrast, the under-constrained case results in a lack of designs that are deemed feasible, as buildings are typically placed arbitrarily within a predefined boundary (Chen et al., 2017; Janssen & Kaushik, 2013). Both cases can ultimately defeat the purpose of applying the optimization algorithm in the first place.

In response, this study presents a novel parametric schema for middle-scale HDB precinct design generation, in which a skeletal modeling approach is adopted. To overcome the lack of design variability and feasibility, we use the “skeletal” line as the organizational device for generating building clusters in this study. These skeletal lines allow design constraints to be more easily handled. In addition, by varying the position and direction of these skeletal lines, solutions with a wide variety of building block arrangements and orientations can be created. Hence, the parametric schema can achieve both variability and feasibility. With this parametric schema, this study establishes an optimization workflow for evolving high-performing precinct designs for HDB development projects.

Following this section, the skeletal modeling approach is first elaborated. A case study is presented for the generation and optimization of designs for an existing HDB project. The result of the case study demonstrates that the parametric schema is able to generate practical solutions that are in alignment with a wide range of constraints applicable to HDB projects. Meanwhile, the variations in position and direction of the skeletal lines also allow the optimization to identify design solutions that outperform existing designs created without the support of an optimization system.

## **2. Method**

### **2.1. SELECTED PROJECTS**

To extract the design logic, we investigate several recent and ongoing housing development projects and select two representative projects for further analysis (Figure 1). These two projects, Tampines GreenGlen and Punggol Northshore Edge, consist of 649 units and 388 units respectively. As displayed in Figure 1, both projects have a parking block building and several tower clusters, each comprising of two to four residential building towers. The towers in each tower cluster are connected by walkways on the ground or air bridges on certain floor levels. In each tower, there can

be two to four units (apartments).

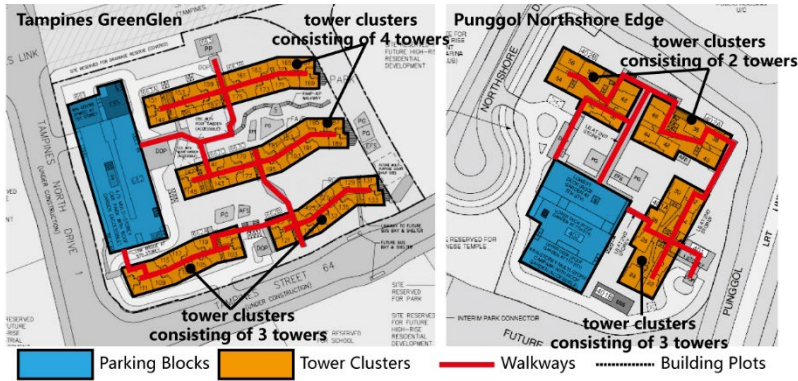


Figure 1. Selected projects.

With a clear spatial structure of precinct configurations, these two projects are ideal for the initial implementation of the proposed parametric schema. Furthermore, their similarity to other projects will allow the parametric schema to be future upgraded at a later date to cover more complex precinct configurations. Based on these two projects, we first identify the underlying dependencies of the design elements constituting the precinct configuration and then schematically represent the dependencies as a series of generative steps. Finally, the generative steps are encoded as a multi-stage algorithm and implemented in one integrated parametric model.

## 2.2. UNDERLYING DEPENDENCIES OF DESIGN ELEMENTS

Based on the two selected projects, a hierarchical dependency among different elements that constitute the precinct design can be identified (Figure 2). First, the parking block plays a critical role in determining the overall precinct configuration in both projects. The parking block functionally connects the exterior city road network to the interior circulation system of the precinct, which, as a result, defines the main entrance of the precinct. Hence, in these two projects, the parking block is always placed along the site boundary where there is an existing road. The block is positioned either parallel or perpendicular to the boundary.

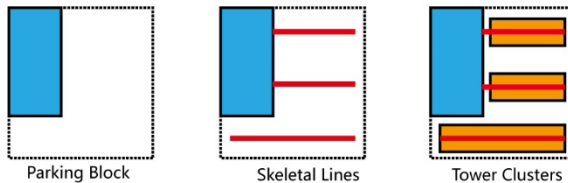


Figure 2. The dependency relationship between different design elements

The second-level element defining the precinct configuration is a set of organizational lines that determine the overall spatial structure within the site. These lines connect the parking block and tower clusters and coincide with walkways and air

bridges. These lines also serve as the connection between different adjacent tower clusters as well as the internal link among the buildings in one cluster. Thus, in the proposed parametric schema, these ‘skeletal lines’ are used as the primary organizational device for generating the building layout.

The direction and length of skeletal lines play a decisive role in determining the orientation and density of the residential towers. For the direction, all the skeletal lines are roughly parallel or perpendicular to the edge of the parking block, while the skeletal lines also serve as the axis to create tower clusters. The length of each line defines the number of towers in the cluster (Figure 3). Moreover, skeletal lines can include small turnings, which can help to accommodate more towers or prevent the end tower from being too close or crossing the site boundary of the site.



Figure 3. The skeletal lines serve as the axis of tower clusters

The last level of elements is tower clusters. In each tower cluster, there can be two to four towers, depending on the length of the skeletal line. In addition, each tower can have two to four apartments, resulting in square towers, L-shaped towers, or slab towers. As shown in Figure 3, the different tower shapes allow for configurations that allow towers to be positioned closer to one another. The solid-void pattern is stable and in relation to the number of towers in the cluster (Figure 4). Moreover, vertical circulations (elevators and fire escape stairs) are shared by the towers in the same cluster. Constrained by the fire evacuation requirements, the number of towers in each cluster is limited to a maximum of four towers.

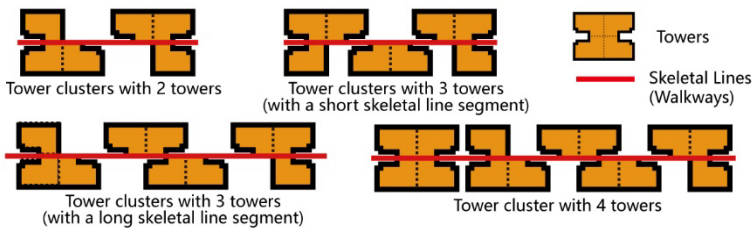


Figure 4. Tower floor plan patterns with respect to different numbers of towers in the cluster

To summarize the dependent relationship, even though the parking block is the first-level design element, the second-level element of skeletal lines play a more critical role in defining the spatial structure of the precinct.

2.3. GENERATIVE STEPS

Based on the hierarchical dependencies among the design elements, six steps are defined, as shown in Figure 5. The first step generates a rectangle-like parking block at one of the corner points along the site boundary. The second step generates a set of skeletal axes orientated parallel or perpendicular to the two inner edges of the parking block. The third step generates a set of skeletal lines along each skeletal axis. Each skeletal line hosts one cluster of towers. The fourth step adjusts the positions and orientations of the skeletal lines in response to the site boundary. The fifth step subdivides the skeletal lines into smaller segments, where each segment will host one tower. The last step generates tower clusters along each skeletal line. The floor plan of each tower is created according to the pre-defined solid-void pattern.

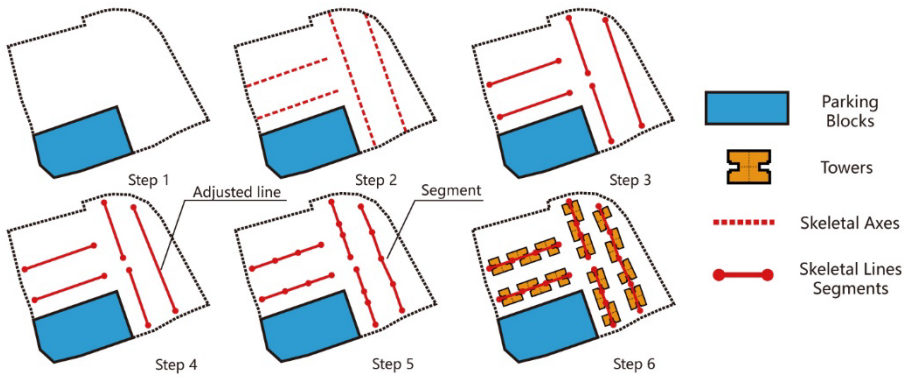


Figure 5. Six generative steps

2.4. PARAMETRIC SCHEMA

Based on the generative steps, we encode these steps into a parametric schema with four sub-algorithms. The four algorithms generate 1) the parking block, 2) the skeletal axes, 3) the skeletal lines, and 4) the tower clusters. These four algorithms are executed as a sequence of steps. The output of each algorithm will serve as the input for the next algorithm.

The first algorithm defines the position of the parking block (Figure 5 -Step 1). It uses a position parameter to select one point on the site boundary. This point is then used to generate a rectangle for the parking block (Figure 6). The rectangle is roughly parallel to the occupied segment of the boundary. Additionally, the width and length of the rectangle are self-adjusted to ensure that the area of parking is sufficient for the predefined parking capacity.

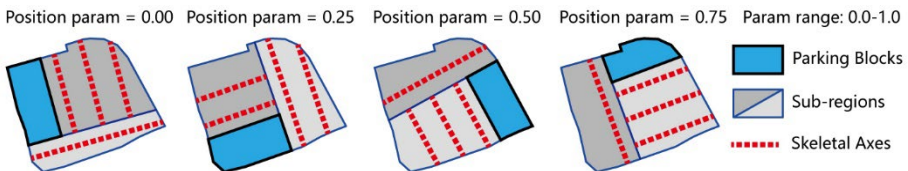


Figure 6. Different parking block positions and the corresponding subdivision of sub-regions

The third algorithm corresponds to the generative steps 3 and 4 in Figure 5. For axes capable of hosting just one tower cluster, a skeletal line with a single segment will be directly created from the axis. For axes capable of hosting more than one tower cluster, the skeletal line is divided into multiple skeletal lines (Figure 5 - step 3). In addition, the length of each skeletal line will be adjusted to satisfy setback and spacing requirements to the site boundary or adjacent skeletal lines. For those skeletal lines close to the boundary, the line orientation will be adjusted to make these lines tend to be more parallel to the boundary (Figure 5 - step 4). Moreover, when a line is too close to the boundary, this line will be moved backward to satisfy a setback requirement.

The last algorithm generates tower clusters. The algorithm first calculates the number of towers in each cluster according to the length of the input skeletal line and then subdivides this line into smaller segments. In most cases, these segments are co-linear if the original skeletal line is parallel or perpendicular to the parking block. For those skeletal lines that have been rotated with respect to the site boundary, each segment will be self-adjusted, and the straight skeletal line will become a segmented line with several turnings (Figure 5 - step 5). After the segments are adjusted, the floor plans of the towers are generated based on each segment (Figure 5 - step 6). The floor plan of the apartment configuration of each tower is determined according to its position in the cluster and the pre-defined solid-void pattern as shown in Figure 4.

Figure 7 illustrates a randomly generated set of precinct configurations using the parametric schema. On the one hand, these designs display high design variability in terms of the building layout configuration, which overcomes the barrack-like layout using over-constrained schemas. On the other hand, the proposed schema effectively eliminates invalid designs, thus facilitating the optimization to focus on searching for high-performing solutions. In contrast, when using under-constrained parametric schemas that arbitrarily place multiple buildings within the boundary, the design solution space can be dominated by invalid designs, including overlapping buildings or layouts that do not meet the setback or spacing requirements. As a result, the optimization can be overwhelmed by the task of identifying legitimate solutions among a vast number of invalid designs (Wang et al., 2018a, 2018b).

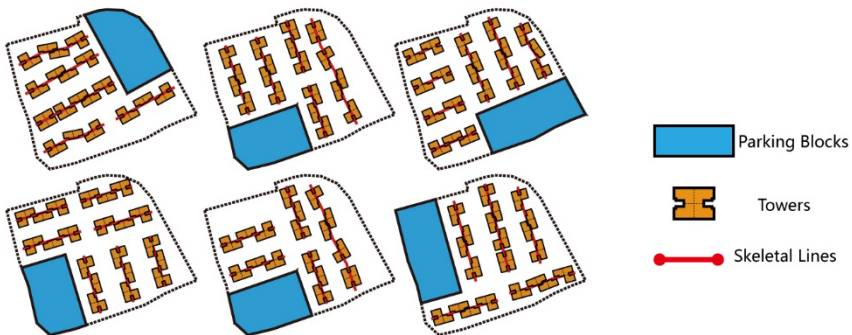


Figure 7. Random sampling building configurations

### 3. Case Study

The proposed schema is applied within a case-study optimization, where precinct



configurations are evolved for a housing typology based on the Tampines GreenGlen project. Design solutions are evaluated against three performance metrics: 1) solar irradiation, 2) unobstructed views, and 3) accessibility. It should be stressed that the main focus of this paper is on design generation rather than on design evaluation and optimization. Hence, the purpose of the case study is to demonstrate that the proposed parametric schema can produce designs maintaining the organizational features of the existing residential precinct typology. Hence, the evaluation metrics are simplified.

### 3.1. EVALUATION METRICS AND CONSTRAINT

The first design objective is to minimize annual solar radiation (ASR) received by the facades of all residential buildings. As Singapore has a tropical climate, lower ASRs mean less solar exposure, which can enhance passive cooling. Second, the accessibility and unobstructed views are included in the optimization to improve residents' convenience, health, and well-being. Finally, for residential precinct design, land-use efficiency is also important, and designs normally need to be close to the plot ratio defined by the urban design code. For the evaluation, ASR and unobstructed views are simulated using the Ladybug Grasshopper plugin (Roudsari & Pak, 2013), while accessibility and land-use efficiency are assessed by two *ad hoc* algorithms.

To evaluate the accessibility, we develop an algorithm to create a circulation network that connects all the skeletal lines back to the parking block, which constitutes the entrance to the site (Figure 8). The algorithm generates various additional connection lines between the disconnected skeletal lines. With all skeletal lines connected, the accessibility metric is evaluated by adding up the distance from each tower cluster to the parking block. The optimization objective is to minimize the sum of these distances. In future research, we envisage that the circulation network can include other destinations apart from the parking block, such as bus stops or footbridges, but in the case study, we simplify the system with one destination.

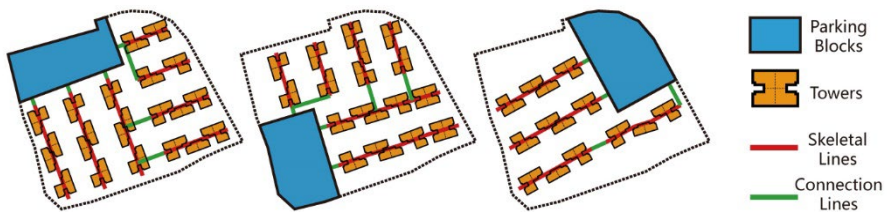


Figure 8. Examples of the generation of connection lines between skeletal lines or the parking block

For land-use efficiency, another algorithm is developed to allow the number of floors to be varied for each tower cluster, which further differentiates the design solution by varying building heights. At the same time, the number of floors is constrained to ensure the generated solutions have the gross floor area (GFA) close to the target area defined by the plot ratio. Figure 9 shows a set of randomly generated designs subject to this requirement. In this case study, we set a plot ratio of 2.0 and a height limit of 50 meters (or 16 floors with a 3-meter floor height). As shown, the inclusion of building heights can further differentiate the mutual shading effect and the unobstructed views among different design variants

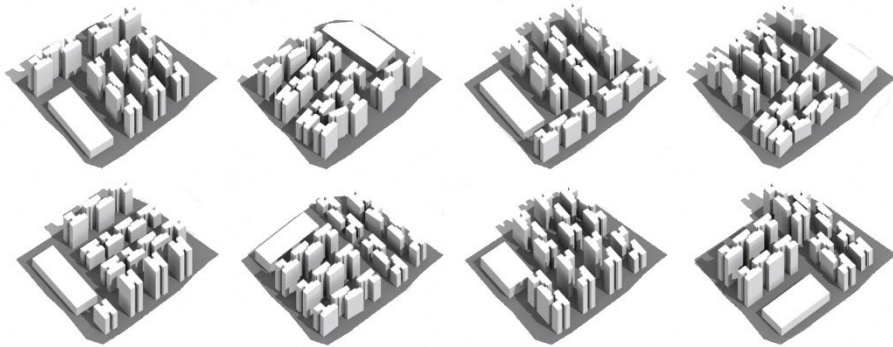


Figure 9. Randomly generated design solutions

### 3.2. DESIGN OPTIMIZATION

For evaluating designs, three evaluation metrics are formulated. While it would be possible to apply a multi-objective Pareto optimization approach, previous research has shown that such approaches often undermine the efficiency and effectiveness of the optimization (Wortmann & Fischer, 2020). In this research, the decision was therefore taken to combined these evaluation metrics into a single weighted objective function.

A hybrid algorithm, called Steady-State Island Evolutionary Algorithm (SSIEA) (Wang et al., 2020) is used to evolve the design population. In comparison with other optimization algorithms in Grasshopper, such as Galapagos, EvoMass is embedded with an island-based model that divides the population into multiple subpopulations, and, thereby, the search of each subpopulation is guided to focus on a different region in the design space. By focusing the optimization on different regions, SSIEA is capable of providing optimization results with a variety of solutions and avoids the optimization ending up with a family of homogeneous designs.

## 4. Result

The SSIEA was used to run the optimization process. In SSIEA, five islands (subpopulation) were defined, each with a population of 30 design variants. A total of 2100 designs were evolved. In the final design population, we select four high-performing design options, each with a distinct building layout configuration. Figure 10 shows these selected options from the optimization result as well as the existing configuration for comparison.

By comparing each of the options based on the three metrics of evaluation, different strengths can be identified. The first two options show similar performance, while the parking block is located at two entirely different positions. For the third option, it outperforms the first two in terms of unobstructed views but comes with a downside of higher solar exposure, as indicated by ASR. The fourth option is inferior to the first two options. In comparison to the third option, although it has lower solar exposure, the corresponding accessibility and unobstructed views are inferior. Lastly, all four options have a GFA that is closely similar to that of the existing one. With different advantages and disadvantages, the four options offer different alternative trade-off solutions

associated with the three objectives, which provides designers with room for choice in early-stage decision-making.



Figure 10. Optimization result.

To further validate the proposed parametric schema and the utility of the design optimization, we also created an additional design based on the existing building layout configuration (the last one in Figure 10). In comparison, the existing design has higher solar exposure than the evolved designs, and its accessibility is also worse than most designs except for option 4. In addition, option 3 outperforms the existing design in all aspects, while it is noticeable that the layouts of option 3 and the existing design are similar. In contrast, the existing design offers a better view than options 1, 2, and 4 and surpasses option 4 in terms of accessibility. The comparison highlights that the existing design already incorporates a building layout with relatively competitive performance. Nevertheless, the optimization approach is still able to discover alternative high-performing design options.

## 5. Discussion and Conclusion

As the study demonstrates, the skeletal parametric schema provides a viable method to generate building layout configurations for residential precinct projects. The wide variation of skeletal line configurations offers sufficient variability and diversity, while the organizational arrangement of buildings also maintains a desirable spatial order. In comparison to other existing parametric schemas for building layout configuration generation, the proposed parametric schema avoids generating invalid design variants without having to resort to over-constraining the layout configurations.

From the practical perspective, the study highlights that the skeletal parametric schema can generate designs that are typologically consistent with the existing projects. In other words, the proposed parametric schema is able to capture the design knowledge derived from existing projects. Hence, it enables the computer to generate human-like designs in an automated way. The optimization algorithm then enables a rapid assessment of the performative potential of the selected housing typology. Such optimization-based design exploration can free the designers from the tedious trial-and-error design processes of manually seeking desirable design solutions, which, in turn, also reduces human errors and bias. As such, the designers will be able to devote more time and effort to design ideation and conceptualizing innovative solutions to challenge

the computed optima.

To conclude, the paper presents a skeletal parametric schema based on a Singapore public housing project. The parametric schema can be incorporated into design optimization processes to provide rapid feedback and meaningful design solutions. By offering practical examples, the optimization, on the one hand, can be directly applied in future projects to help reduce the time and effort spent on routine design tasks. On the other hand, the optimization results can become a benchmark for decision-making. It enables designers and other stakeholders to evaluate design proposals in comparison to evolved solutions based on existing housing typologies. Future research directions include testing the parametric schema with other public housing projects and further developing strategies for refining performance metrics to guide the optimization process to produce more balanced designs. In addition, more detailed design elements, such as apartment types and circulation cores will be further embedded into the schema, which can allow for a more thorough design evaluation.

## References

- Chen, K. W., Janssen, P., & Norford, L. K. (2017). Automatic Parameterisation of Semantic 3D City Models for Urban Design Optimisation. *Future Trajectories of Computation in Design: 17th International Conference, CAAD Futures 2017*, 51–65.
- Chen, K. W., & Norford, L. (2017). Evaluating urban forms for comparison studies in the massing design stage. *Sustainability (Switzerland)*, 9(6). <https://doi.org/10.3390/su9060987>
- Cheong, K. H. (2019). Creating Liveable Density Through a Synthesis of Planning, Design and Greenery. In T. Schröpfer & S. Menz (Eds.), *Dense and Green Building Typologies. SpringerBriefs in Architectural Design and Technology* (pp. 7–12). Springer Singapore. [https://doi.org/10.1007/978-981-13-0713-3\\_3](https://doi.org/10.1007/978-981-13-0713-3_3)
- Janssen, P., & Kaushik, V. (2013). Evolutionary Design of Housing: A template for development and evaluation procedures. *Proceedings of the 47th International Conference of the Architectural Science Association*, 197–206.
- Roudsari, M. S., & Pak, M. (2013). Ladybug: A parametric environmental plugin for grasshopper to help designers create an environmentally-conscious design. *Proceedings of BS 2013: 13th Conference of the International Building Performance Simulation Association*, 3128–3135.
- von Richthofen, A., Knecht, K., Miao, Y., & König, R. (2018). The ‘Urban Elements’ method for teaching parametric urban design to professionals. *Frontiers of Architectural Research*, 7(4), 573–587. <https://doi.org/10.1016/j.foar.2018.08.002>
- Wang, L., Janssen, P., & Ji, G. (2018a). Utility of Evolutionary Design in Architectural Form Finding: An Investigation into Constraint Handling Strategies. In John S. Gero (Ed.), *Design Computing and Cognition '18* (pp. 177–194). Springer. [https://doi.org/10.1007/978-3-030-05363-5\\_10](https://doi.org/10.1007/978-3-030-05363-5_10)
- Wang, L., Janssen, P., & Ji, G. (2020). SSIEA: a hybrid evolutionary algorithm for supporting conceptual architectural design. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 34(4), 458–476. <https://doi.org/10.1017/S0890060420000281>
- Wang, L., Janssen, P., & Ji, G. (2018b). Efficiency versus effectiveness: A study on constraint handling for architectural evolutionary design. *Learning, Adapting and Prototyping - Proceedings of the 23rd CAADRIA Conference* (pp. 163–172).
- Wortmann, T., & Fischer, T. (2020). Does architectural design optimization require multiple objectives? A critical analysis. *RE: Anthropocene, Design in the Age of Humans - Proceedings of the 25th CAADRIA Conference* (pp. 365–374).

# POI DATA VERSUS LAND USE DATA, WHICH IS MORE EFFECTIVE IN MODELING THEFT CRIMES?

JIAJIA FENG<sup>1</sup>, YUEBING LIANG<sup>2</sup>, QI HAO<sup>3</sup>, KE XU<sup>4</sup> and  
WAISHAN QIU<sup>5</sup>

<sup>1,3,4</sup>*Department of Architecture, Tsinghua University, Beijing, China.*

<sup>2</sup>*Department of Urban Planning and Design, The University of Hong Kong, Hong Kong*

<sup>5</sup>*Department of City and Regional Planning, Cornell University, Ithaca, NY, USA*

<sup>1</sup>*jjiajiafsz@126.com, 0000-0003-0931-2767*

<sup>2</sup>*yuebingliang@connect.hku.hk, 0000-0003-2089-4606*

<sup>3</sup>*hao-q14@tsinghua.org.cn, 0000-0002-6063-0018*

<sup>4</sup>*xk\_thu@163.com, 0000-0003-1286-9777*

<sup>5</sup>*wq43@cornell.edu, 0000-0001-6461-7243*

**Abstract.** Alleviating crime and improving urban safety is important for sustainable development of society. Prior studies have used either land use data or point-of-interests (POI) data to represent urban functions and investigate their associations with urban crime. However, inconsistent and even contrary results were yielded between land use and POI data. There is no agreement on which is more effective. To fill this gap, we systematically compare land use and POI data regarding their strength as well as the divergence and coherence in profiling urban functions for crime studies. Three categories of urban function features, namely the density, fraction, and diversity, are extracted from POI and land use data, respectively. Their global and local strength are compared using ordinary least square (OLS) regression and geographically weighted regression (GWR), with a case study of Beijing, China. The OLS results indicate that POI data generally outperforms land use data. The GWR models reveal that POI Density is superior to other indicators, especially in areas with concentrated commercial or public service facilities. Additionally, Land Use Fraction performs better for large-scale functional areas like green space and transportation hubs. This study provides important reference for city planners in selecting urban function indicators and modelling crimes.

**Keywords.** POI; Land Use; Urban Functions; Theft crime; Predictive Power; SDG 16.

## 1. Introduction

In 2015, the United Nations General Assembly (2015) set up seventeen Sustainable Development Goals, one of which is to “promote peaceful and inclusive societies for

sustainable development". To achieve this goal, the United Nations aims at reducing all kinds of urban crimes, including violence, illicit financial flows and theft, by 2030. Therefore, investigating how to control urban crime is extremely important for realizing a more sustainable future for global citizens.

Under the tendency of finer-grained planning, neighbourhood-level crime study has become a heated topic. In this urban level, neighbourhood functions have long been revealed to significantly correlate with crime occurrence. Understanding the impact of neighbourhood functions on crime is crucial for urban planners to avoid potential high crime areas. In solving the problem, an increasing number of researches have emerged in the wake of high-resolution urban data. Land use data, which represents urban land types as polygons, is most commonly used to represent neighbourhood functions in urban crime studies. Most crime studies use land use fractions, i.e., the percentage of land use area to the neighbourhood, as the indicator of neighbourhood functions. For instance, Stucky and Ottensmann (2009) found that robberies are more common in neighbourhoods with larger fraction of commercial area. Sohn (2016) discovered that residential areas were negatively related to crime density in the neighbourhood level. Recent studies have also examined how land use mix affects crime patterns. For example, a Herfindahl index of land uses was constructed by Wo (2019) for the Los Angeles neighbourhoods to capture mixed land use and assess the effect of mixed land use on crime. De Nadai et al. (2020) used the average entropy among land uses as the indicator of land use mix and found that its effects on urban crime varies from one city to another. Most recently, big data reflecting city dynamics have become widely available. Within this regard, POI data, which abstracts physical facilities as points, has been proposed to profile neighbourhood functions. To link POI data to urban crime, three indicators have been constructed from POI data, including POI density (i.e., POI count within a certain area), POI fraction (i.e., the proportion of POI count to the total POI count) and POI mix (i.e., the mixture of POI in a neighbourhood). Bendler et al. (2014) first integrated POI density in crime analysis for San Francisco, and found the densities of POIs like nightlife, food and drink to be strong predictors of theft crime. Redfern et al. (2020) calculated the POI kernel densities for crime prediction in ten UK cities, and reported that the additional POI predictors contributed much to the model accuracy. Wang et al. (2016) used both POI fraction and POI raw count to characterize the neighbourhood functions and confirmed the efficiency of POI indicators, especially the raw count, in crime inference problems in Chicago. These studies show the advantage of POI data in reflecting more accurate and diverse urban facilities.

Overall, most existing studies have used either land use or POI data to investigate the correlations between neighbourhood functions and urban crime. In terms of indicators, three types of indicators have been introduced in the literature: Density, Fraction and Mix. Note that Fraction and Mix can be constructed from both POI and land use data while only POI data has Density attributes. Nevertheless, there have not been systematic comparisons among different indicators from land use or POI data on their strength, divergence and coherence to profile neighbourhood functions. There exist disagreements of land use and POI data's impact on crime in previous research. For example, using land use data, Sohn (2016) argued that residential land use was negatively related to crime, while the research of Rumi et al. (2018) showed that there exist no significant correlations between residential POI and various crimes. The use

of different types of indicators brings about further contradictions. For instance, Wang et al. (2016) reported that the crime rate was negatively correlated with the raw count of shop POI, while the correlation was statistically insignificant with its fraction. Comparing the performance of land use and POI data, with regard to the construction methods of indicators, can help avoid possible theoretical contradictions regarding the effect of neighbourhood functions on crime, and provide valuable guidance on the selection of indicators for future studies. To achieve the goal, this study provides a comparative study of different indicators based on land use and POI data in association with neighbourhood functions and urban crime in the neighbourhood level.

### 2. Study Area and Data

This study was conducted for theft crimes in Beijing, China. We focus on the area within the Fifth Ring Road of Beijing, where the majority of urban crimes occurred. The theft crime is chosen as a representative because of its high occurrence, great threats to the citizens' possession, and sensitivity to the urban contexts. The theft crime data used in this study was collected from the China Judgements Online. In this study, we use 4182 theft crimes from 2014 to 2016 that fall inside our study area for research.

The land use data was provided by Beijing Municipal Commission of Planning and Natural Resources. It consists of land parcels covering the study area which are labelled with a land use category. The POI data was crawled from Gaode Map, which provides the geographical coordinates, category and name of each POI. To facilitate comparative analysis with POI data, we reclassified land use data into 10 new categories and POI data into 14 new categories. Each POI category belongs to one land use type and one land use type corresponds to 1-3 POI categories.

Apart from land use and POI data, several supplementary data sources are used to construct control variables: 1) permanent residential population provided by the Sixth National Census of China; 2) road junctions collected from OpenStreetMap; 3) building footprints extracted from Baidu Map; 4) social media data crawled from Microblog; 5) night light intensity data extracted from remote sensing images.

### 3. Methodology

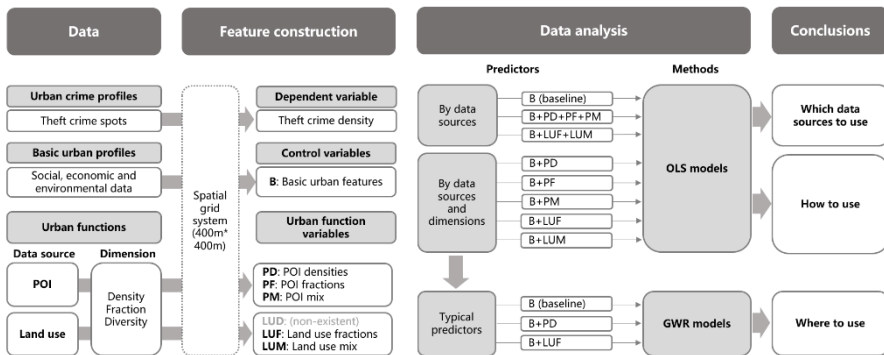


Figure 1. Flowchart for the whole study

In this section, we first introduce the variables and models used in our study, and then elaborate on the experiment framework we developed for the comparative analysis. A flowchart for the whole study is shown in Figure 1.

### 3.1. VARIABLES

**Theft Crime Density:** We divide the study area into regular-shaped grid cells and each crime lot is assigned to the grid cell that contains it. The crime intensity of each cell was then measured as the total number of crimes that occurred in the grid. Specifically, a raster grid with a resolution of 400m by 400m was used, which resulted in 4267 grid cells in total. The spatial distribution of theft crime intensity is shown in Figure 2.

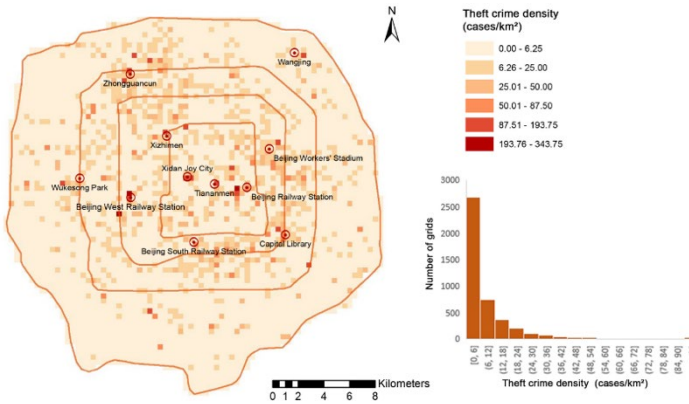


Figure 2. Spatial and statistical distribution of theft crime intensity

**Neighbourhood Function Variables:** Two types of indicators were constructed from land use data to profile neighbourhood functions: 1) Land Use Fraction (LUF): the proportion of land use area per category to the area of each grid; 2) Land Use Mix (LUM): a Herfindahl index of the ten specific land use categories for each grid following Wo (2019):

$$LUM = 1 - \sum_{j=1}^N LUF_j^2 \quad (1)$$

where  $LUF_j$  represents the area fraction of land use category  $j$  out of  $N$  categories. In terms of POI data, three types of indicators are constructed: 1) POI Density (PD): the normalized count of POIs per category in each grid; 2) POI Fraction (PF): the proportion of POI count per category to the total POI count in each grid (Chi et al. 2016); 3) POI Mix (PM): a Herfindahl index of the POI categories for each grid.

**Control Variables:** Apart from neighbourhood function variables, we also include five types of control variables, including population density, microblog count, night light intensity, road junction count and floor area ratio of each grid.

### 3.2. MODELS

**OLS regression:** OLS regression is one of the most widely used regression approaches in urban studies. In this study, the OLS regression takes theft crime density as the dependent variable, and the independent variables are our proposed indicators of



neighbourhood functions using land use or POI data, along with other basic urban variables as control variables.

Geographically Weighted Regression (GWR): One possible limitation of OLS regression is their poor consideration of spatial non-stationarity. To tackle this issue, this study adopts a GWR model which allows the estimated parameters to vary over the spatial domain. It also takes spatial autocorrelation into account and thus can mitigate the bias caused by spatial autocorrelation.

### 3.3. EXPERIMENT FRAMEWORK

To compare the performance of POI and land use data as indicators of urban functions in urban crime studies, we develop a two-step experiment framework. First, we use the OLS regression model to compare the global prediction performance of different indicators constructed from POI or land use data. A baseline model is constructed using only control variables. Besides, we also build two models incorporating all the land use features and all the POI features respectively, and five more specific models with LUF, LUM, PD, PUF and PM as predictors respectively. Second, we devise a GWR model to uncover the local prediction performance for urban crime using POI or land use data. Specifically, GWR models with variables from POI and land use data are conducted, and the spatial variance of the model performance can be illustrated with local  $R^2$ . We further distinguish the applicability of land use and POI features in different neighbourhoods using correlation analysis and case studies.

## 4. Results

### 4.1. RESULTS FROM LR MODELS

While OLS Model 1 acts as a baseline, land use variables and POI variables are applied in OLS Model 2 and OLS Model 3 respectively. As shown in Table 1, OLS Model 3 has the best accuracy of all, with an adjusted  $R^2$  of 0.234. Therefore, POI data generally provide more powerful predictors for theft crime prediction than land use data do.

Table 1. Performance of the OLS models

	OLS Model							
	M1	M2	M3	M2-2	M2-3	M3-1	M3-2	M3-3
$R^2$	0.120	0.139	0.240	0.139	0.121	0.236	0.139	0.122
Adjusted $R^2$	0.119	0.136	0.234	0.136	0.120	0.232	0.135	0.121
AIC	3.47e+4	3.46e+4	3.41e+4	3.46e+4	3.47e+4	3.41e+4	3.46e+4	3.47e+4

Predictors for the OLS models: M1: B; M2: B+LUF+LUM; M3: B+PD+PF+PM; M2-2: B+LUF; M2-3: B+LUM; M3-1: B+PD; M3-2: B+PF; M3-3: M+PM.

Regarding the p-values for urban functional variables, 9 out of 14 categories among POI density variables show significant impact on theft crime density, while only 4 of the 10 land use fraction categories are significant predictors, which reveals that POI densities might be generally more effective than land use fractions in theft crime prediction. POI mix proves to be significant as well.

Table 2. The coefficients of significant urban functional features with the top-3 positive or negative impacts in OLS M2 and OLS M3

	Feature	Coefficient (top-3 positive)	Feature	Coefficient (top-3 negative)
OLS M2	LUF(Medical)	15.086	LUF(Residential)	-2.080
	LUF(Commercial)	9.003		
	LUF(Financial)	6.166		
OLS M3	PD(Shop and catering)	84.795	PD(Government)	-17.885
	PD(Nightlife)	27.024	PD(Enterprise)	-13.518
	PD(Station)	20.996	PF(Station)	-4.200

In OLS M2, less than three of the significant urban function features ( $p < 0.1$ ) have negative coefficients.

Table 2 shows the coefficients of significant urban functional features with the top-3 positive or negative impacts in OLS Model 2 and OLS Model 3. Among all the significant predictors, the POI densities of governments and enterprises, the POI fractions of stations and the land use fraction of residence are negatively related to crime density, while the rest predictors have positive effects.

Some influencing patterns of urban functions found from OLS Model 3 are coherent to those from OLS Model 2. For instance, POI densities of typical commercial facilities are all significant predictors with positive coefficients, which is in accordance with the effects of commercial land in OLS Model 2.

OLS Model 3 also reveals more detailed impacts from other urban functions that are not shown in OLS Model 2. Enterprise and bank are two typical POI categories in financial land, but their densities have adverse impacts on theft crime. Moreover, for the same urban function, its density feature and fractional feature might have opposite impacts. For example, the POI density of stations has a positive coefficient, while the POI fraction of stations has a negative coefficient. As for diversity features, OLS Model 3 shows that POI mix has a significant promoting effect on theft crime, while land use mix in OLS Model 2 is insignificant. Compared with land use data, POI data might provide a higher-resolution depiction of urban functions for a land parcel and thus contribute to theft crime prediction.

To construct a well-performed and efficient model with the least independent variables necessary, systematic experiments are carried out to explore which dimensions of information from land use or POI data are most effective. According to their results shown in Table 1, POI densities have the strongest predictive power; POI fractions and land use fractions have similar and moderate predictive power; POI mix and land use mix are almost ineffective.

#### 4.2. RESULTS FROM GWR MODELS

Based on the results of the OLS models, the GWR experiments only focus on the former three kinds of effective theft crime predictors: land use fractions, POI fractions and POI densities. We construct 4 GWR models, namely GWR Model 1 (B), GWR Model 2-2 (B+LUF), GWR Model 3-1 (B+PD), and GWR Model 3-2 (B+PF). The

adjusted  $R^2$  and AIC values testifies that in most cases, GWR model has stronger explanatory power over OLS model due to its consideration of spatial heterogeneity (Table 3).

Table 3. Results of geographically weighted model for thief crimes

	GWR Model 1 (B)	GWR Model 2-2 (B + LUF)	GWR Model 3-1 (B + PD)	GWR Model 3-2 (B + PF)
$R^2$	0.178	0.267	0.385	0.284
Adjusted $R^2$	0.150	0.194	0.314	0.186
AIC	34618.96	34498.87	33835.37	34587.64

On the local scale, by contrast, POI densities do not always have an advantage. Figure 3 shows the spatial distributions of local  $R^2$  of the GWR models, each of which has distinctive spatial patterns. (For POI data, only GWR Model 3-1 is retained as the better-performing model as compared with GWR Model 3-2). Compared with the baseline (GWR Model 1), GWR Model 2-2 obtains prominent increases in local  $R^2$  in the nearby area of the western 3<sup>rd</sup> Ring Road, which implies the strong predictive power of land use fractions. Correspondingly, the result of GWR Model 3-1 indicates that POI density features have notable predictive power near central Beijing, especially in the northern areas of the 2<sup>nd</sup> Ring Road.

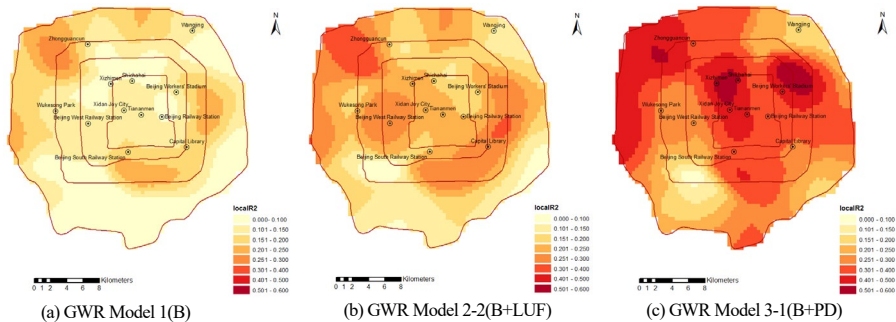


Figure 3. Spatial distribution of local  $R^2$  from the GWR models

Since the predictive power of land use fractions and that of POI density features show distinctive patterns, it is necessary to make a direct comparison. Therefore, a Power ratio indicator is calculated for each grid via the following equation:

$$Power\ ratio = \frac{local\ R^2\ in\ GWR\ Model\ 3 - 1}{local\ R^2\ in\ GWR\ Model\ 2 - 2} \quad (2)$$

The higher Power ratio is, the more advantage POI density features have over land use features. As shown in Figure 4, most grids have Power ratio larger than 1. The patterns of the power ratio agree with the intuitive judgement from local  $R^2$  as discussed above. In this study, grids with power ratios larger than 2 are defined as prominent pro-

POI grids, and those with power ratios smaller than 1 are defined as prominent pro-land-use grids.

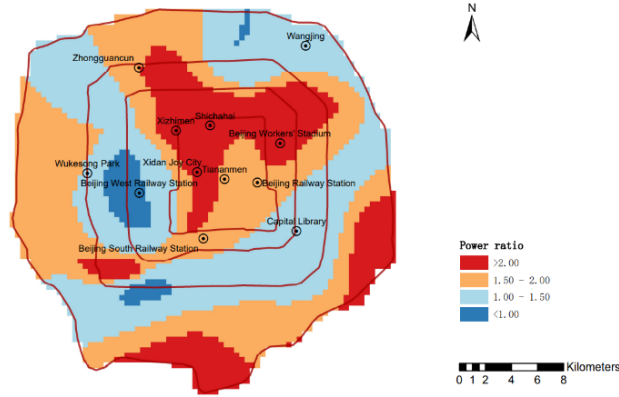


Figure 4. Spatial distribution of the power ratio

In total, 4 clusters of pro-POI grids and 2 clusters of pro-land-use grids are recognized. The northern area of the 2<sup>nd</sup> Ring Road has remarkably large power ratios, where POI density features have prominent advantages over land use fractions. By contrast, a cluster of grids near the western 3<sup>rd</sup> Ring Road have power ratios smaller than 1, indicating that GWR Model 2-2 obtains better local R<sup>2</sup> than GWR Model 3-1 does with less independent variables.

Table 4. Pearson Correlation between urban functional variables and the power ratio

Land use feature	Correlation	POI density feature	Correlation
LUF(Residential)	.015	PD(Residence)	.019
LUF(Administration)	.022	PD(Government)	.154**
		PD(Police station)	.083**
LUF(Culture)	.043**	PD(Culture)	.053**
LUF(Education)	.097**	PD(Education)	.095**
LUF(Medical)	.021	PD(Medical)	.089**
LUF(Commercial)	.064**	PD(Hotel)	.076**
		PD(Shop and catering)	.063**
		PD(Nightlife)	.128**
LUF(Financial)	.056**	PD(Enterprise)	.054**
		PD(Bank)	.056**
LUF(Municipal)	-.050**	PD(Public facility)	.151**
LUF(Green space)	-.137**	PD(Park)	.064**
LUF(Transportation)	-.094**	PD(Station)	.015

\*\* P < .01. \* P < .05.

To further explore factors that decide whether land use features or POI density features perform better locally, the Pearson correlation coefficients are calculated between urban functional variables and the power ratio (Table 4). A significant positive coefficient indicates that the more intense the urban function is, the more superior POI density features are to land use fractions in theft crime prediction, and vice versa.

For all the POI categories except residence and station, the density features are positively correlated to the power ratio at the  $p < 0.01$  level. The top five most influential POI categories are government, public facility, nightlife, education, and medical. The results indicate that in places with high POI prosperity, POI density features generally have stronger predictive power over land use features.

The Pearson correlations between land use features and the power ratio are more complicated. For cultural land, educational land, commercial land and financial land, their land use fractions are positively correlated to the power ratio at the  $p < 0.01$  level, where POI density features tend to have stronger predictive power. Meanwhile, the fractions of municipal land, green space and transportation land are negatively correlated to the effect ratio at the  $p < 0.01$  level. In these lands, POI density features tend to be inferior to land use fractions.

## 5. Conclusion

This study has three major findings:

(1) Among two data sources of urban functions, land use and POI, the latter one could provide generally more effective predictors in urban crime prediction.

(2) Among three dimensions of urban functional features, i.e., density, fraction and diversity, density features have the strongest correlations with urban crime occurrences, and these features could only be constructed from POI data.

(3) On most occasions, POI densities are optimal urban crime predictors in functional areas with small or medium-sized facilities, especially administration and public services, and commercial services. However, land use fractions provide better depictions for large-scale infrastructures like transportation hubs, green space and municipal services, as well as function zones with buildings of divergent scales like residential.

As alleviating crime becomes an important aspect of sustainable development in the social dimension, our findings provide several notable implications in the neighbourhood level in terms of urban crime prediction and planning policy making towards safer cities.

First, our research suggests that POI densities are effective predictors in theft crime prediction. While urban functional organizations have been proven to be related to crime occurrences, it might be difficult to acquire up-to-date and high-resolution land use data for many cities. Fortunately, POI data are now highly accessible, which could become a potential substitute to land use data.

Second, the superiority of POI densities over all fractional and diversity features indicates that the intensity of urban functions has notable impacts on theft crime. Previous planning policies are usually coarse organizations of different land use types, while our study suggests that the facility densities in a function zone should be also taken into account. For example, among commercial lands, places where nightlife spots

gather tend to trigger more theft crimes, so limiting the density and business hours of nightlife spots might be of great help to reduce theft crime rate.

Third, our study also points out that POI densities and land use fractions are fit for different urban functions when studying their associations with theft crimes. To promote the security of cities through crime prediction or function zone adjustment, the best solution might be to focus on the POI densities of those functions undertaken by small or medium-sized facilities, and the land use fractions of large-scale infrastructures or function zones with heterogeneous building forms.

Last but not least, our findings regarding the power of POI data and land use data in urban modelling might be applied in other topics. The principles we discuss on how to choose between the two data sources and further construct proper features to depict different urban functions might be also useful for modelling other phenomena associated with urban functional organizations, such as urban vitality.

This study also has some limitations that need to be addressed in future research. Land use data and POI data were generated in different years (2014 and 2016). Although we use theft crime data from 2014 to 2016 accordingly, the inconsistency of data collection time might still influence the results. Another problem is that the predictive powers of POI data and land use data may vary at different spatial granularity (400m by 400 m in our case). Further research should make clear how to choose urban functional indicators when the spatial granularity changes. Besides, this study is conducted specifically for theft crime. More researches are needed to reveal if the superior predictive power of POI data remains true for other types of urban crimes.

## References

- Bendler, J., Ratku, A., & Neumann, D. (2014). Crime Mapping through Geo-Spatial Social Media Activity. *ICIS 2014 Proceedings*.  
<https://aisel.aisnet.org/icis2014/proceedings/ConferenceTheme/12>
- De Nadai, M., Xu, Y., Letouzé, E., González, M. C., & Lepri, B. (2020). Socio-economic, built environment, and mobility conditions associated with crime: A study of multiple cities. *Scientific Reports*, 10(1), 13871. <https://doi.org/10.1038/s41598-020-70808-2>
- Redfern, J., Sidorov, K., Rosin, P. L., Corcoran, P., Moore, S. C., & Marshall, D. (2020). Association of violence with urban points of interest. *PLOS ONE*, 15(9), e0239840. <https://doi.org/10.1371/journal.pone.0239840>
- Rumi, S. K., Deng, K., & Salim, F. D. (2018). Crime event prediction with dynamic features. *EPJ Data Science*, 7(1), 43. <https://doi.org/10.1140/epjds/s13688-018-0171-7>.
- Sohn, D. W. (2016). Residential crimes and neighbourhood-built environment: Assessing the effectiveness of crime prevention through environmental design (CPTED). *Cities*, 52, 86-93. <https://doi.org/10.1016/j.cities.2015.11.023>.
- Stucky, T. D., & Ottensmann, J. R. (2009). Land use and violent crime. *Criminology*, 47(4), 1223-1264. <https://doi.org/10.1111/j.1745-9125.2009.00174.x>
- United Nations General Assembly (2015, October 21). *Transforming our world: the 2030 Agenda for Sustainable Development*. Retrieved December 2, 2021, from [https://www.un.org/ga/search/view\\_doc.asp?symbol=A/RES/70/1](https://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1)
- Wo, J. C. (2019). Mixed land use and neighborhood crime. *Social science research*, 78, 170-186. <https://doi.org/10.1016/j.ssresearch.2018.12.010>
- Wang, H., Kifer, D., Graif, C., & Li, Z. (2016, August). Crime rate inference with big data. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 635-644).

# REALTIME URBAN INSIGHTS FOR BOTTOM-UP 15-MINUTE CITY DESIGN

CESAR CHENG<sup>1</sup>, YUKE LI<sup>2</sup>, RUTVIK DESHPANDE<sup>3</sup>, RISHAN ANTONIO<sup>4</sup>, TEJAS CHAVAN<sup>5</sup>, MACIEJ NISZTUK<sup>6</sup>, RAMANATHAN SUBRAMANIAN<sup>7</sup>, CAMIEL WEIJENBERG<sup>8</sup> and SAYJEL VIJAY PATEL<sup>9</sup>

<sup>1,2,3,4,5,6,7,8,9</sup>*Digital Blue Foam.*

<sup>1,2,3,4,5,6,7,8,9</sup>{cesar|yuke|ru-

vik|rishan|tejas|maciej|ram|camiel|sayjel}@digitalbluefoam.com

**Abstract.** This paper introduces a real-time neighbour scoring system, using data collected from various web-based APIs, to facilitate “15-minute city” designs. The system extends on the current state of the art in three ways; first, it incorporates a multi-source urban API, to automate the extraction of location-based information from online sources; second, it provides a quantitative method to calculate and index “15-minute city” performance; and third, it provides a web-based application, to allow real-time feedback of neighbourhood design performance complementing the design refinements at a building and tenancy level. In addition to discussing its theoretical basis, and technical implementation, this paper provides a case study to demonstrate how the neighbourhood scoring system is incorporated into the design of a hypothetical mixed-use urban development.

**Keywords.** Industry Innovation and Infrastructure; Sustainable Cities and Communities; Urban Walkability; Urban Accessibility; 15-minute City; Spatial Analysis; SDG 9; SDG 11.

## 1. Introduction

The "15-minute city" is driving a paradigm shift in how people conceive neighbourhood design sensibilities. This design approach, popularised by Carlos Moreno's 2020 TED Talk (Moreno, 2020), arranges a variety of amenities and services within a 15-minute walking distance of all residents to accommodate their needs. While cities like Madrid, Milan, Ottawa, and Seattle are embracing the 15-minute city concept as part of a top-down urban planning strategy (Yeung, 2021), there are still a few hurdles to overcome within the design community. First and foremost, no one seems to agree on what a 15-minute city is; a review of 15-minute city literature (Blečić et al., 2020; Duany et al., 2021; Mehaffy et al., 2015; Pozoukidou et al., 2021) has failed to establish an objective approach for determining whether or not a location is a 15-minute city. Additionally, determining pedestrian accessibility to various locations around a

city is challenging. Until now, only skilled specialists with access to advanced spatial analysis tools, large urban data sets, and GIS software such as ArcGIS or QGIS have been able to do so.

This article outlines a system for dealing with these issues. Multiple GIS data sources are combined in a single, user-friendly online web-based tool. As a result, the user can analyse data more quickly and with less technical knowledge. A new neighbourhood score system was introduced to help benchmark existing neighbourhood designs. The scoring method is based on 15-minute city assumptions and is implemented in an online design tool as illustrated in Figure 1. The tool allows for a quick assessment of a neighbourhood's quality and how it will evolve as new building designs are introduced. The method and its effectiveness are presented in a case study scenario where new mixed-use areas are introduced to the neighbourhood.



*Figure 1. This article introduces Digital Blue Foam's "Urban Insights" tool to provide real-time, descriptive spatial metrics at the neighbourhood-scale*

## 2. Related Work

Recently, researchers in spatial data analysis, network analysis, and geospatial information systems, have become increasingly interested in the relationship between walkability and land use. For example, the 'Space Syntax' methodology, developed at the University College of London (UCL) uses graph theory to measure various aspects of network structure and connectivity (Karimi, 2012). Space Syntax offers stand-alone software packages and plug-ins for GIS software to perform network analysis for pedestrian accessibility. Another example is recent from the MIT-Harvard CityForm Lab. Like Space Syntax, they have developed a methodology to perform urban network analysis. The MIT Team released plugins for Rhino3D and GIS to calculate walkability and accessibility to urban services and amenities (Sevtsuk et al., 2012). In both cases, the proposed tools are limited to users with advanced knowledge in spatial analysis and



network theory, as well as knowledge of the associated software platforms. PedCatch (White et al., 2016), is a web-based interactive platform that leverages satellite and other imagery sources, as well as crowdsourced mapping data, to investigate urban pedestrian accessibility. By integrating Open Data from governments with OpenStreetMaps to rank and map the walkability of over 700,000 streets and footpaths, Walkonomics (Quercia et al., 2015) provides urban pedestrians with enhanced wayfinding to any location, including the most picturesque tree-lined streets and parks. Pedestrians First (Chestnut, 2018) is a collection of interactive tools that evaluates walkability in cities around the globe, with a focus on infants, toddlers, and their caregivers. Additionally, there are network analysis libraries that are available for R and Python languages such as NetworkX (Glanz et al., 2012), Momepy (Fleischmann et al., 2019), OSMnx (Boeing, 2017) among others. These libraries are also limited to expert users with coding skills and domain knowledge.

Available to a wider public, some online-based data providers have offered API services that provide easy-to-interpret information of urban accessibility. Examples of these are WalkScore API and HereAPI 15-minute city project. WalkScore awards a location with a score based on a calculation that measures pedestrian friendliness by analysing data sources such as Open Street Maps, US Census data, and Google. HereAPI 15-minute city project is a web-based service that uses geospatial data and isochrones to display the number of locations within a given travel distance with a particular travel mode. This approach reduces the complexity of the calculation, facilitates access to location data and presents results more intuitively for non-experts. The latter approach was followed in this article. The work builds upon the information these services provide and complements it with additional metrics, a user interface, and a scoring system. The user interface facilitates access to urban data and the interpretation of results. The scoring system renders a complete informative analysis useful in the early stages of project development in urban interventions.

## 2.1. RESEARCH QUESTION

Before the 15-minute city can be fully adopted by the design and planning community from the bottom-up, the following questions should be addressed:

- What methods can be used to extract and consolidate different 15-minute city data sets into useful inputs?
- Can a single, universal 15-minute city metrics be developed? How to qualify and compare designs using unified metrics which doesn't oversimplify the design problems?
- With what methods can a broad audience of designers, planners, and policy-makers be engaged in the 15-minute city discourse?

## 3. Methodology

The methodology focused on developing a set of indexes to qualify neighbourhoods. These indexes are used to score a locality based on available contextual data in different areas of the world. To develop the quality score of a neighbourhood under the conceptual framework of the "15-minute city", it is critical to consider both time and

spatial dimensions. For this, the isochrones analysis is employed to define the area of search for urban features. An isochrone is typically used to depict travel time within a given road network for a particular travel mode (walking, cycling, driving). The isochrone calculates a set of endpoints that lie at an equal time distance from a starting point.

The isochrone was employed to define a walking or cycling shed within a 5-15 minute commute time. Then relevant urban features that fall within that distance can be searched. These urban features are essential services that need to be accessible within a given distance. The 15-minute City concept defines the following services: residential (houses, apartments), office, health (hospital and clinics), institutions (preschool, schools, universities, college), public facilities (library, community centre), commerce (shops, malls), markets (markets and supermarkets), restaurants (café, restaurant, bars), transportation & utility (train station, bus stop, airport).

### 3.1. NEIGHBORHOOD QUALITY SCORE

To calculate the neighbourhood score, the system filters urban features by category. Each category is given a score based on the number of features as well as their inclusion within either a 5 or 15 minutes isochrone from a reference point. If the urban feature is found within a 5-minute transit distance, it scores higher than the 15-minute transit distance. The final score is the aggregation of these values. Equation (1) outlines the formula used in the calculation.

$$S = \left( \frac{\left( \sum_{i=1}^n f(a_i) \right) - L}{H - L} \right) \cdot 100$$

$$A = \begin{bmatrix} \text{Commercial} \\ \text{Residential} \\ \text{Hospital} \\ \text{Industrial} \\ \text{Office} \\ \text{Public} \\ \text{Park} \\ \text{Restaurant} \\ \text{Transport} \\ \text{Parking} \end{bmatrix} ; f(x) = \begin{cases} 100 & \text{if, } x \text{ exists within 5 mins of walking} \\ 50 & \text{if, } x \text{ exists within 15 mins of cycling} \\ 0 & \text{if, } x \text{ does not exist} \end{cases} \tag{1}$$

*L* = The lowest score; Absence of all programs in 15 mins of cycling  
*H* = The highest score; Presence of all programs in 5 mins of walking

### 3.2. URBAN DIVERSITY INDEX

Another important indicator of neighbourhood performance is urban diversity. Measuring the distribution of urban features produces a useful indicator to understand the quality of a neighbourhood (Balletto et al., 2021). The right balance between various urban services within a short distance makes for more vibrant and active communities. For the calculation of this index, Simpson’s diversity index was taken as

$$1 - \sum_{i=1}^k \frac{n_i(n_i - 1)}{n(n - 1)} \tag{2}$$

a reference, as it allows to determine the diversity of elements within a group of features. The Simpson diversity, often used in biology to measure biodiversity, is a quantitative measure that reflects differences in type within a population (or data-set) with a finite number. In the case of the index, urban features found within the isochrone shed area are used. Equation (2) outlines the formula we used to calculate Urban Diversity.

### 3.3. DESCRIPTIVE SPATIAL METRICS

To provide a general description of urban morphology, descriptive statistics are used to identify the mean, minimum and maximum values for building footprint area and building height as presented in Figure 2. These metrics serve as a proxy to understand local urban regulation of the study area as it describes the floor and ceiling values of building form found in the area. The results are presented in charts that help visualize the pattern of distribution for height and foot-print area within the isochrone.

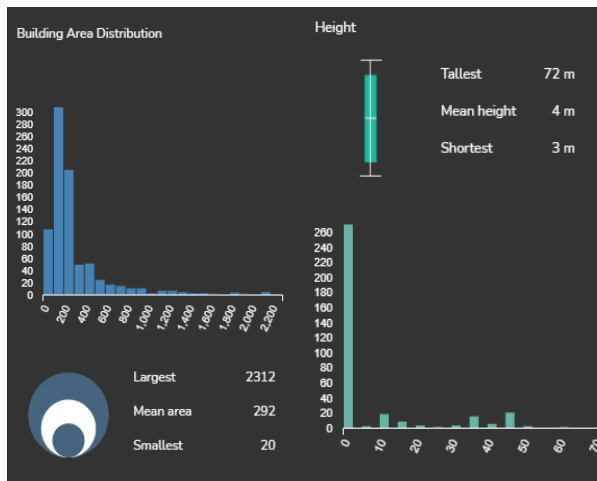


Figure 2. Building footprint area distribution (blue), building height distribution (green)

## 4. Implementation

Using the described method DBF has built a platform that allows to quickly estimate the quality of neighbourhoods to inform decision-making in urban development visible in Figure 3. The platform interface allows users to navigate a map environment where they can search specific locations, drop a pin to perform the study of the area and obtain a score for the area. Figure 4 illustrates the result of the analysis for different neighbourhoods in the world. As data availability varies significantly from city to city, the score accuracy is relative to the quality of the data found for a given area. When data is insufficient, or the online sources do not cover the current urban context, it is possible to upload up-to-date information from a field survey. Figure 6 shows a before and after comparison taking into account the impact of additional urban functions from the design intervention scenario. The platform allows users to generate multiple design scenarios with a variety of program distributions and illustrate the impact of such

interventions in the neighbourhood making it easy to visualize the improvement achieved and compare the scenarios against each other.



Figure 3. Digital Blue Foam's 'Urban Insights' User Interface design tool for "15-minute city" planning and development



Figure 4. Analysis for different neighbourhoods in various locations

## 5. Case Study

Figure 6 presents the neighbourhood score system in the context of a major mixed-use urban development, comprising several components. This urban development proposes a scenario where additional areas for commercial, office and residential space are introduced to the neighbourhood to improve the availability of services and increase the diversity of land use. In this example, the scenario was modelled using the DBF design platform. Users can create building configurations and produce massing studies with associated programs, the output of the design is exported as GeoJson format and brought into the DBF Urban Insights Platform. The model is consolidated and geoc-

located into the larger urban context so it can be part of the larger map environment that displays the entire city. The system can re-run urban performance analysis taking into account the proposed scenario. The system evaluates the distribution of the program within 15 minutes for two scenarios, the existing condition and the proposed scenario taking into account the new functions that have been introduced to the neighbourhood. The user is also given suggestions on how to improve the design based on missing programs as shown in Figure 5. Based on this feedback the user refined the design by increasing the number of commercial spaces, allocating area to education and adding more residential units to increase density. The existing condition shows one important urban feature, a hospital. However, large areas are dedicated to parking space, and it lacks density in residential areas.

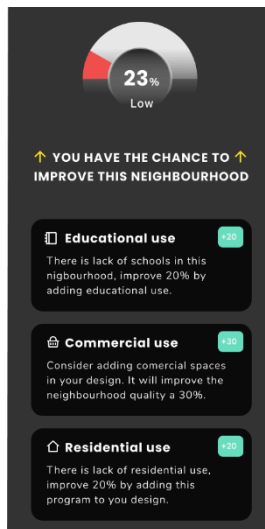


Figure 5. Urban Insights Identifies missing programs and makes suggestions to improve the neighbourhood

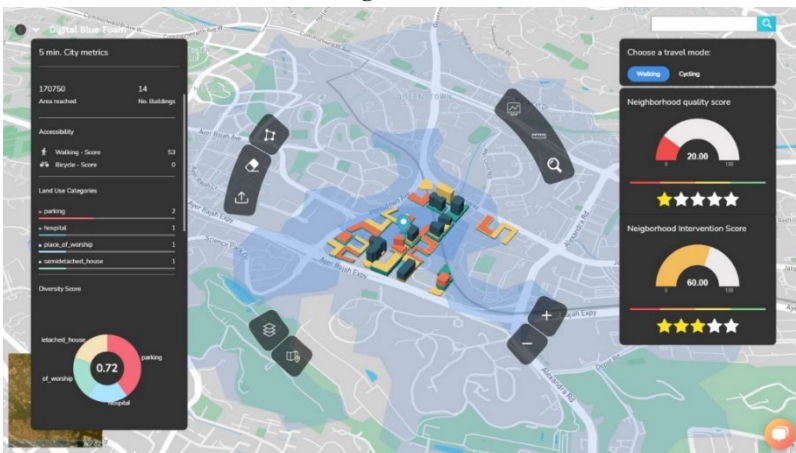


Figure 6. Comparing neighbourhood scores before and after design intervention

## 6. Discussion

Focusing on the development of a 15-minute city design framework, a novel neighbourhood scoring system implemented in an online design tool was introduced. Based on specific geolocation, the system can retrieve contextual data-based, perform calculations, and generate visualizations to describe the spatial quality of a given neighbourhood. The primary contribution of this work is the development of tools, frameworks, and metrics, to help designers, planners and developers identify missing elements and potential opportunities to improve the quality of neighbourhood designs. The novelty of the system is the holistic approach for the integration of multiple data sources into an intuitive, platform-independent web system that can be accessed by a broader audience. Unlike traditional software packages such as ArcGIS or Qgis, the system is a design tool available to all stakeholders involved in the planning and design process of neighbourhoods. Typically, geospatial analysis is performed by experts who present results to other agents involved, keeping them away from the analysis process and leading to a fragmented and linear planning process. The system abstracts the complexity of analytical processes, presents easy to interpret results, and allows for real-time interactive feedback for multiple agents to engage in the planning process serving as a more effective tool to communicate information regardless of the technical knowledge of the user.

## 7. Conclusion

This article describes new frameworks, metrics, tools, and methodology to transform translating the “15-minute city” from just an idea into a practical and repeatable urban development strategy. This article makes three primary contributions to advance 15-minute City design:

- Consolidates various isolated data sources into a single platform to make them part of the pre-design process so users accessing the platform can directly make use of relevant information regarding walkability and accessibility and implement it to their projects;
- Simplifies the process of performing pedestrian accessibility analysis and interpretation of results using a user-friendly interface for users with no advanced technical knowledge on network theory and spatial statistics;
- Proposes a method to set a standard to compare the urban quality of neighbourhoods based on a scoring system.

With the developed system, “15-minute city” ceases to be an idea, slowly becoming a part of reality. The architects and planners receive a design tool allowing the implementation of this concept into the design process. Thanks to this, the value of the urban fabric has a chance to gain a new and sustainable quality.

### 7.1. POTENTIAL IMPACT

This work represents a first step towards the development of real-time, location-based, tools to enhance neighbourhood design and planning. Ultimately, by empowering all stakeholders, including residents, designers, planners, developers, and policymakers,

with tools to perform neighbourhood quality analysis, it is possible to profoundly impact collective-decision making and build consensus. Together, these stakeholders will be able to identify missing elements in urban plans or point out potential opportunities to create more walkable, accessible and diverse neighbourhoods. The tool can be extended by incorporating surveys or user feedback to indicate potential sites for lacking amenities in a precinct, culminating in an inclusive and informed design process.

## 7.2. FUTURE WORKS

Future development of the Neighbourhood Quality Scoring system will focus primarily on the integration of additional location information in the data-set. Currently, the system searches for the presence or absence of a given urban feature within a 5 or 15-minute isochrone shed. The system puts a higher weight on those features found in closer proximity to the origin point. To further develop this scoring method additional data properties can be introduced to assign more relevant weights to the analysis. Using a more complete data-set with information about demographics or population density, it would be possible to assign specific weights to particular urban features.

For example, in areas with a higher presence of the elderly population, the presence of hospitals, community centres and parks would have a higher weight. Similarly, in areas with a higher presence of households with children, pre-schools and schools would have a higher weight. This would make the score reflect better the specific needs of the inhabitants of a particular neighbourhood.

The current distance function measures the origin of the isochrone to the closest urban feature and is currently limited to testing whether the point lies within 5 or 15 minutes distance. To achieve higher precision and allow for a more accurate comparison between the scores of different neighbourhoods a distance decay function can be introduced and a larger search area beyond the shed of 15 minutes can be included. A distance decay function would allow the system to capture differences in accessibility that cannot be perceived within the 15-minute boundary but still contribute to the overall accessibility of a given neighbourhood.

With more complete data and more measuring precision, one can begin to reveal location information that has remained hidden or inaccessible by the larger public. To overcome the current lack of publicly available data in some areas, future versions of the tool will look into local data providers that could supply the missing data on request. Currently, the system allows users to upload to DBF web platform user-generated or custom geospatial data that may be available to them to make it part of their design workflow. The proposed system can set the basis for more robust sensing systems that could incorporate Internet of Things (IoT) technology to present real-time live feedback of the state of a neighbourhood and the city. As cities become more complex, it is important to begin to understand the multiple processes occurring at other levels beyond the physical realm. Today, digital devices can capture relevant information such as environmental noise, air quality and pollution or information regarding human activities such as commercial deliveries, Internet activity, wireless transactions. For this information to be meaningful and accessible to a larger public it is necessary to develop interfaces to communicate with potential users, therefore extending the capabilities of the scoring system we are proposing could serve to unlock new opportunities for the

use of data to improve life in cities by empowering the residents of these cities with access to more information.

Another dataset that can be integrated into the tool, to extract and assess information from visual data, is the Google Street View Imagery (SVI). In this approach, the image segmentation model can be used. Being a tool from a subset of computer vision, it can mask and extract ratios of elements that are perceived from the street view perspective, such as the tree and vegetation coverage, sky view factor, presence of vehicular and pedestrian traffic, urban furniture and other street equipment.

## References

- Balletto, G., Ladu, M., Milesi, A., & Borruso, G. (2021). A methodological approach on disused public properties in the 15-minute city perspective. *Sustainability*, 13(2), 593. <https://doi.org/10.3390/su13020593>
- Blečić, I., Congiu, T., Fancello, G., & Trunfio, G. A. (2020). Planning and design support tools for walkability: A guide for urban analysts. *Sustainability*, 12(11), 4405. <https://doi.org/10.3390/su12114405>
- Boeing, G. (2017). OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, 65, 126–139. <https://doi.org/10.1016/j.compenvurbsys.2017.05.004>
- Chestnut, J. (2018, March 28). Pedestrians first: Tools for a Walkable City. Institute for Transportation and Development Policy. Retrieved February 3, 2022, from <https://www.itdp.org/publication/walkability-tool/>
- Duany, A., Steuteville, R. (2021, February 8). *Defining the 15-minute city* [Text]. CNU. <https://www.cnu.org/publicsquare/2021/02/08/defining-15-minute-city>
- Fleischmann, M. (2019). Momepy: Urban morphology measuring toolkit. *Journal of Open Source Software*, 4(43), 1807. <https://doi.org/10.21105/joss.01807>
- Glanz, T., Nam, Y., & Tang, Z. (2012). Sustainable urban design and walkable neighborhoods. *IntechOpen*. <https://doi.org/10.5772/27285>
- Karimi, K. (2012). A configurational approach to analytical urban design: ‘Space syntax’ methodology. *URBAN DESIGN International*, 17(4), 297–318. <https://doi.org/10.1057/udi.2012.19>
- Mehaffy, M. W., Porta, S., & Romice, O. (2015). The “neighborhood unit” on trial: A case study in the impacts of urban morphology. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability*, 8(2), 199–217. <https://doi.org/10.1080/17549175.2014.908786>
- Moreno, C. (2020, October). *The 15-minute city* [Video]. TED Conferences, [https://www.ted.com/talks/carlos\\_moreno\\_the\\_15\\_minute\\_city](https://www.ted.com/talks/carlos_moreno_the_15_minute_city)
- Pozoukidou, G., & Chatziyiannaki, Z. (2021). 15-minute city: Decomposing the new urban planning eutopia. *Sustainability*, 13(2), 928. <https://doi.org/10.3390/su13020928>
- Quercia, D., Aiello, L. M., Schifanella, R., & Davies, A. (2015). The digital life of walkable streets. *Proceedings of the 24th International Conference on World Wide Web* (pp. 875–884). <https://doi.org/10.1145/2736277.2741631>
- Sevtsuk, A., & Mekonnen, M. (2012). Urban network analysis toolbox. *International Journal of Geomatics and Spatial Analysis*, 22(2), 287–305.
- White, M., Kimm, G., & Langenheim, N. (2017). Pedestrian access modelling with tree shade – won’t someone think of the children. *Procedia Engineering*, 198, 139–151. <https://doi.org/10.1016/j.proeng.2017.07.078>
- Yeung, P. (2021). *How “15-minute cities” will change the way we socialise*. Retrieved November 29, 2021, from <https://www.bbc.com/worklife/article/20201214-how-15-minute-cities-will-change-the-way-we-socialise>



## PRESENCE STICKERS

*A seamlessly integrated smart living system at a solitary elderly home*

CHOR-KHENG LIM<sup>1</sup>

<sup>1</sup>*Department of Art and Design, YuanZe University*

<sup>1</sup>*kheng@saturn.yzu.edu.tw, 0000-0003-0990-5191*

**Abstract.** This research develops Presence Stickers, the capacitive sensing module that can be easily affixed to the existing space element surfaces (such as a wall, door, stairs), and daily objects or home furniture (such as a chair, cabinet, table, sofa, etc.). These 30\*30 cm Presence Stickers can actively sense people's physical behaviors and body movements in spaces. From the preliminary analysis of observing the 80-year-old elderly subject's daily activities, the movement trajectory of the 'Move-Stop' pattern is found. There will be a Touch (T) and Touchless (TL) relationship between the body and the space elements or objects. Furthermore, the touchless, or non-contact, intimate relationship can also be divided into two types: 1. The body that 'Passes by' (P) the spatial elements or objects, and 2. The body that 'Stays' (S) in front of the object and performs activities. These three types of the intimate relationship between bodies and objects, i.e., T, TL-P, and TL-S, were used as the main sensing conditions to develop the Presence Stickers sensor module. We affixed 8 Presence Stickers on 9 objects in six spaces and finally, the life pattern can be analyzed and the sensors provide the customized intelligent application function for the elderly.

**Keywords.** Intuitive Interaction Design; Capacitive Sensor; Daily Object; Touchless; Body Movement; Smart Home; SDG 3.

### 1. Introduction

In the field of Human-Computer Interaction design, user interfaces that interact with natural human abilities such as gestures, touch, or voice are often referred to as Natural User Interfaces (NUI). In tactile interactive design, Tangible User Interface, TUI emphasizes the interaction between the physical environment and digital information, through the more natural grasping, operation, assembly, and other tactile relationships with objects as an interactive interface design, which can more freely and intuitively operate between the virtual and the physical (Ishii & Ullmer, 1997). The Organic User Interface, OUI proposed by Holman et al. in 2008, emphasizes flexible, tangible interface, and interactive interface that includes sensing and actuation capabilities. It combines daily life environment and behavior patterns to provide more intuitive

interaction (Holman & Vertegaal, 2008). For the future home based on the design of Ubiquitous Computing, Nabil (2017) proposed Organic User Interfaces Interiors, by applying Organic User Interface Design to interior spaces design. 'OUI Interiors' mainly turns the daily objects into interactive artefacts. In this sense, interior spaces, surfaces (walls, floors, tables, and ceilings), interior objects such as furniture, and decorative accessories can become computationally-driven interactive artefacts, potentially changing their physical appearances, i.e., shape, colour, pattern or texture.

In addition to the Tangible Interaction of direct contact with physical objects, there is also a non-contact distance relationship between human behavior and physical objects. Greenberg (2011) used the concept of 'Proxemic' proposed by Hall (1969) as a basis and extended it to the distance relationship between humans and physical objects (digital and non-digital devices). In the Ubicomp interaction, the intimate relationship between the human body and physical objects is emphasized and the concept of 'Proxemic Interaction' is proposed. Greenberg (2011) points out that there are five dimensions of proxemics for UbiComp: Distance, Orientation, Movement, Identity, and location. Non-contact or touchless interactive interfaces have flourished in recent years and are mostly used in public places. To avoid physical contact, touchless interfaces are even more critical in the era of the COVID-19 pandemic. In fact, this research aims to explore how we should make better use of the touchless interface provided by the products of technology in daily life, instead of just relying on existing input technology. This is because there is an intimate relationship between the human body and space (or objects). How to activate an existing space or objects to detect the presence of human beings, even if there is no contact.

From the perspective of Ubiquitous Computing's smart home design, the subject of this study is to explore how to interact with spaces or objects through intuitive body behaviors and to obtain digital information. And the result will be used or as a user interface design for smart home products. At the same time, the research scope includes the intimate relationship between physical behavior and daily-life objects in spaces. Through intuitive body movements and activities in daily life, this research analyzes the intimate relationship including contact and non-contact, between people and objects.

This study is based on the UbiComp Proxemics dimensional study proposed by Greenberg (2011) and explores how to integrate the capabilities of sensors and actuators into the spatial surfaces (such as walls, floors, doors, and windows) and everyday objects (such as furniture, decorations) that are closely related to the body; how to perceive body movements and behaviors in the space and provide feedback for intelligent life assistance. This research aims to introduce smart auxiliary functions through simple sensing components or modules implanted in space surfaces or everyday objects without changing life behaviors or replacing household objects.

Regarding the sensing technologies, the most commonly used currently includes pressure sensors or capacitive sensors; the non-contact sensing ones include infrared sensors, capacitive sensors, and image recognition. An innovative sensor design, Touché, which is developed by the Disney Research team (Sato et al. in 2012) uses capacitive sensing technology to sense human body touch and perceive subtle body movements. Moreover, the Touché capacitive touch sensing can reflect on objects of different scales (pens, doorknobs, mobile phones, tables), all with the same sensitivity.

And different body gestures can also be distinguished (Sato, Poupyrev, & Harrison, 2012). In addition to capacitive sensing that can sense the touch of the human body, Capacitive proximity sensors can sense the non-contact body gestures at a certain distance. Compared to other sensing technologies, capacitive sensing has the advantages of low cost, low power consumption, and simple realization, so it is also used as an indoor human localization sensing application (Arshad et al., 2017).

This research will take advantage of capacitive sensing, which can sense from proximity to contact, and develop a sensing system that interacts with natural behaviors. The goal of this research is to design a smart living system that can be used by solitary elderly people through touch/touchless interface design. It is hoped that the sensing system can be seamlessly integrated into the home of elderly people and improve life assistance. This research develops Presence Stickers that can be easily affixed to the existing space elements (wall, floor, door, stairs) or home furniture (chair, cabinet, bed, table, sofa, etc). These 30\*30 cm Presence Stickers can actively sense people's physical behaviors and body movements in spaces. The signals acquired from the Presence Stickers in spaces will then integrate into the controller system through the MQTT technology and logically define the behaviors classification. Users can define their own feedback, such as switching home appliances, lights, electric fans, air-conditioning, and other daily applications. Figure 1 shows the design framework of Presence Stickers.

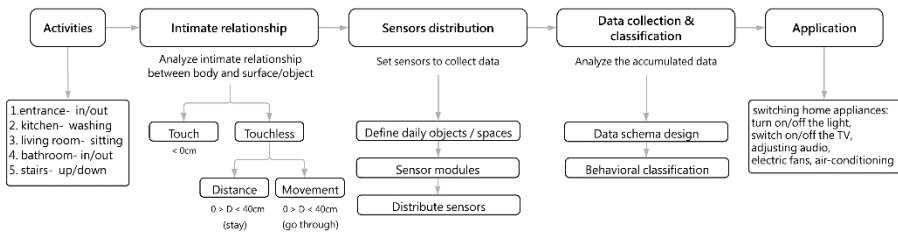


Figure 1. The design framework of Presence Stickers

## 2. Activities and intimate relationship between body and object

This research derives the design guidelines for the sensor module of Presence Stickers from the review of literature research on interaction theory Ubicomp Proxemics and Capacitive Proximity sensing technology. The notion of Ubicomp Proxemics proposed by Greenberg (2011) points out that the close relationship between the human body and physical objects has five different dimensions, including Distance, Orientation, Movement, Identity, and Location. However, because the purpose of this research is to perceive the existence and activity behavior of people in space through objects, the first step in the design framework (Figure 1) is to observe and interview the elderly’s home life pattern to understand the elderly’s daily behaviors and activities and find out relevant everyday objects and relationship attributes, and then use the appropriate Ubicomp Proxemics dimension to analyze the intimate relationship between the body and objects.

In this study, a solitary 80-year-old male older adult was interviewed, and we

recorded his daily life behaviors and activities at home spaces within 24 hours. Table 1 shows his life behaviors record, including the time, activities, spaces, relevant objects, and interaction behaviors. From the preliminary analysis of observing the activities of this elderly at home in one day (Table 1), it is obvious that he performed different activities in different spaces (living room, kitchen dining room, room, bathroom, outdoor). In response to changes in activities at different times, he first walked through the aisle space to the next space. It can be clearly seen that he had a 'Move' and 'Stop' activity trajectory when switching between different activities. Due to this phenomenon, this research only focuses on the two dimensions of the Ubicomp Proxemics concept: Distance and Movement, to explore the distance and movement relationship between body and spaces/objects when the human body moves in space.

Based on the movement trajectory of the Move-Stop pattern, the second step in the design framework (Figure 1) of this research is to analyze the intimate relationship between the human body and the space surface or object. There will be a Touch (T) and a Touchless (TL) relationship between the body and the space elements or objects. The touch or contact relationship relates to the behaviors that must touch the object, such as sitting on a chair; Furthermore, a touchless or non-contact intimate relationship can be divided into two types: 1. The body that 'Passes by' (P) the spatial elements or objects, and 2. The body that 'Stays' (S) in front of the object and performs activities.

- The first type, the 'Pass by' (P) relationship between the body and the object, is a short-time (<1 minute) and directional movement. There will be a distance of less than 40 cm between the body the spatial elements or objects, such as the movement of going up and down the stairs, the relationship between the body and handrails or walls.
- The second type, the 'Stay' (S) relationship between the body and the object is longer (>1 minute), and the body will remain at a certain distance (<20cm) for continuous activities, such as the behaviors of washing hands and teeth, it will maintain a intimate relationship of <10cm between the body and the surface of the cabinet under the sink.

Based on the analysis of the intimate relationship mentioned, this study encodes the elderly subject's daily life behaviors and activities according to three interaction modes: Touch (T), 'Pass by' mode of Touchless interaction (TL-P), and 'Stays' mode of Touchless interaction (TL-S). From the coding data, Table 1 shows that the interaction between the body of the elderly subject and the object is mostly direct contact behavior, and it is mainly in the 'Stop' activity trajectory regarding the Move-Stop pattern; In addition, non-contact relationships mainly occur in 'Move' activity trajectory. It includes the body behaviors, such as passing by the walls between the aisle space or passing through the door when entering different rooms. Furthermore, the intimate non-contact relationship between the body and the surface of the cabinet was observed when the elderly subject was using the sink.

The three dimensions of the intimate relationship between the body and objects in the home space of the elderly in this study include: Distance, Movement, and Touch. Figure 2 shows the relative behaviors of the elderly subject in these three intimate relationship dimensions which will be examined by the Presence Stickers.

Table 1. 24 hours daily life behaviors and activities of the elderly subject

Time	Activities (Move-Stop)	Space	Object	Interaction (T) (TL) (S) (P)
07:15	waking up	bedroom	bed, quilt, pillow, switch	(T) lying on the bed, (T) tucking in the quilt, (T) sleeping on the pillow, (T) sitting on the bed, (T) turning off the light, aircon *
07:20	brushing the teeth, washing the face, toilet	bathroom	switch, door, toilet, toilet paper, sink, cabinet	(T) turning on the light * (TL-P) opening/closing the door, (TL-S) standing near/in front of the sink cabinet, (T) sitting on the toilet (T) holding toilet paper, flushing
	Walking to	bedroom	door	(TL-P) opening/closing the door
	Walking to	corridor	wall	(TL-P) walking
08:00	breakfast, washing dishes	kitchen, dining room	table, chair, tableware (plate, fork, bowl, spoon, chopsticks), sink, kitchen counter	(T) sitting on the chair, (T) touching the table, (T) holding the tableware (TL-S) standing near/in front of the sink cabinet/counter (T) turning the faucet
	Walking to	corridor	wall	(TL-P) walking
09:00	watering, gardening	garden	main door, water pipe, tools, broom	(TL-P) opening/closing the door, (T) holding the pipe, tools
11:00	playing on a cell phone (Facebook, games)	garden or living room	main door, chair/sofa, cushion, cell phone	(TL-P) opening/closing the door, (T) sitting on the chair/sofa (T) holding the cushion, (T) holding a cell phone
	Walking to	corridor	wall	(TL-P) walking
12:00	taking a bath	bathroom	door, faucet, toiletries, cabinet	(TL-P) opening/closing the door, (T) turning the faucet, (T) holding the toiletries, (TL-S) standing in front of the cabinet, taking out the cloth
	Walking to	corridor	wall	(TL-P) walking
12:30	playing on a cell phone (games)	bedroom	door, bed, cell phone	(TL-P) opening/closing the door, (T) sitting on the bed, (T) holding a cell phone
	Walking to	corridor	wall	(TL-P) walking
13:00	lunch, washing dishes, doing chores	kitchen, dining room	table, chair, tableware (plate, fork, bowl, spoon, chopsticks), sink, kitchen counter	(T) sitting on the chair, (T) touching the table, holding the tableware (TL-S) standing near/in front of the sink cabinet/counter (T) turning the faucet
	Walking to	corridor	wall	(TL-P) walking
14:00	napping	bedroom	door, bed, quilt, pillow	(TL-P) opening/closing the door, (T) lying on the bed
16:00	waking up, going to the toilet	bathroom	door, toilet, sink	(TL-P) opening/closing the door, (T) sitting on the toilet, (T) holding toilet paper, flushing (TL-S) standing near/in front of the sink cabinet
	Walking to	Bedroom	door	(TL-P) opening/closing the door
	Walking to	corridor	wall	(TL-P) walking
16:10	playing on a cell phone (Facebook, games)	living room	chair/sofa, cushion, cell phone	(T) sitting on the chair/sofa (T) holding the cushion, (T) holding a cell phone
16:30	exercising (Treadmill)	living room	Treadmill, towel	(T) running on the treadmill (T) wiping sweat
17:00	doing chores (sweeping the floor)	second floor	stairs, bloom	(TL-P) climbing up/down the stairs (T) holding the bloom
	Walking to	corridor	wall	(TL-P) walking
18:00	taking a bath, resting/doing chores	bathroom	door, faucet, toiletries, cabinet	(TL-P) opening/closing the door, (T) turning the faucet, (T) holding the toiletries, (TL-S) standing in front of the cabinet, taking out the cloth
	Walking to	corridor	wall	(TL-P) walking
19:00	dinner	kitchen,	table, chair,	(T) sitting on the chair,

		dining room	tableware (plate, fork, bowl, spoon, chopsticks), sink, kitchen counter	(T) touching the table, (T) holding the tableware, (TL-S) standing near/in front of the sink cabinet/counter (T) turning the faucet
	Walking to	corridor	wall	(TL-P) walking
20:00	watching TV (news)	living room	chair/sofa, cushion, remote control	(T) sitting on the chair/sofa (T) switching on the TV* (T) holding the cushion, (T) holding the remote control
	Walking to	corridor	wall	(TL-P) walking
21:00	playing on a cell phone (Facebook, games) reading	bedroom or study room (2 <sup>nd</sup> floor)	door, switch, bed	(TL-P) opening/closing the door, (T) turning on the light, aircon* (T) sitting on the bed (T) sitting on the chair
22:00	sleeping	bedroom	quilt, pillow, switch	(T) lying on the bed, (T) tucking in the quilt, (T) sleeping on the pillow, (T) turning off the light
03:00	toilet	bathroom	door, toilet, toilet paper, sink, cabinet	(TL-P) opening/closing the door, (T) turning on the light (T) sitting on the toilet, (T) holding toilet paper, flushing (TL-S) standing near/in front of the sink cabinet
03:10	sleeping	bedroom	bed, quilt, pillow, switch	(T) lying on the bed, (T) tucking in the quilt, (T) sleeping on the pillow, (T) turning off the light
05:00	toilet	bathroom	door, toilet, toilet paper, sink, cabinet	(TL-P) opening/closing the door, (T) turning on the light (T) sitting on the toilet, (T) holding toilet paper, flushing (TL-S) standing near/in front of the sink cabinet (T) turning off the light
05:10	sleeping	bedroom	bed, quilt, pillow, switch	(T) lying on the bed, (T) tucking in the quilt, (T) sleeping on the pillow, (T) turning off the light

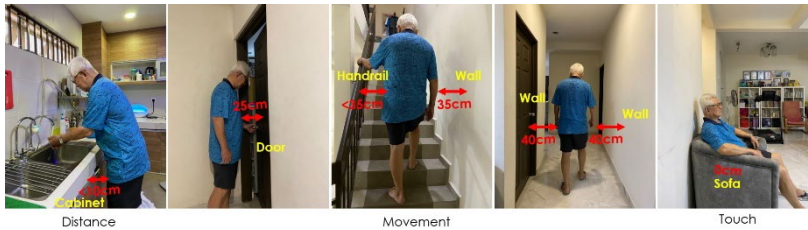


Figure 2. The behaviors of the elderly subject in three dimensions of intimate relationship

### 3. Presence Stickers sensing module

Based on the above analysis, the three types of the intimate relationship between bodies and objects (T, TL-P, and TL-S) will be used as the main sensing conditions to develop the Presence Stickers sensor module. Then this study carried out the experiment of the design and development of Presence Stickers. At the same time, the sensing module would be tested regarding the activity behaviors of the elderly subject in the three relational dimensions: Distance, Movement, and Touch (Table 2).

From the research literature, the use of capacitive sensing as a sensing interaction application between the human body and objects was originally proposed in 1995 by the MIT Media lab research team Zimmerman and others. Zimmerman et al. (1995)

pointed out that The Human-Computer Interaction design can allow objects to sense the human body as an intuitive and natural interaction by using the electric field sensing method that costs less, consumes less power, and can be simply implemented. Braun et al. (2015) also analyzed the research and application of capacitive sensing and believed that the use of capacitive proximity sensing in a smart environment has its advantages, and proposed a design guide for this application. In order to have both contact and non-contact proximity sensing functions, this study adopted capacitive proximity sensing as the sensory design of the Presence Stickers module.

The Presence Stickers module is equipped with a breadboard with Arduino Nano, coupled with a 30\*30 cm aluminum foil as the electrode sensing plate. ADCTouch library is used as the main computation and control of capacitive sensing. To integrate the data of multiple sensors placed in the different spaces and objects, the module is coupled with the WEMOs Wi-Fi module to transmit the accumulated sensor data to the cloud. This module is powered by a lithium battery. To stabilize the current of the module, a voltage regulator module is added. Since capacitive sensing will be affected by environmental factors such as temperature and humidity, object materials, human conditions, etc., this research had added a Reset module that can reset the sensors through Wi-Fi connection, can be used as a remote adjustment in the research experiment. Figure 3 shows the information framework of the Presence Stickers sensing module.

As mentioned, this study uses the ADCTouch library as the program algorithm for capacitive sensing. This library can only be applied to AVR's microcontrollers. ADCTouch is a library that allows users to create a capacitive sensor without any external hardware. This library makes use of the AVR's internal wiring to get decent resolution with just a single pin. As the design of Presence Stickers in this study needs to be put on or stuck on any objects or surfaces, the module design needs to be very simple and light, so the simple circuit design of ADCTouch meets the needs of the design and has been adopted.

This research aims to easily stick multiple Presence Stickers on different spaces and objects, to sense human behaviors in the space, to collect data for a long time for the computation of behaviors classification, and finally to provide the customized intelligent application function for the solitary elderly at home. Therefore, Wi-Fi is used as the wireless network information integration method. With the WEMOs Wi-Fi module and the lightweight MQTT publish-subscribe network protocol, all the sensors, Presence Stickers, have become IoT devices.

Table 2. Sensing conditions and the relative activity behaviors

UbiComp Proxemics	Types of intimate relationship	Relationship between body and daily objects	Activities/ behaviors	interior space
Touch	Touch (T)	sitting on the chair	playing on a cell phone (Facebook, games)	living room
Movement	Touchless - Pass by (TL-P) d < 40cm; t < 1minute	facing the door, walking/passing through the wall/handrail	entering/leaving the rooms climbing up/down the stairs	all rooms, aisle/corridor stairs
Distance	Touchless- Stay (TL-S) d < 20cm; t > 1minute	standing near/ in front of the sink cabinet /counter	washing hands/ dishes	kitchen/ bathroom

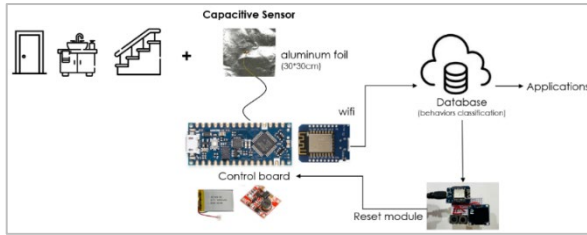


Figure 3. The information framework of the Presence Stickers

#### 4. Experimental setup and sensors distribution

Figure 4 shows the house floor plan and the sensor distribution. The capacitance sensing values of the Presence Stickers put on the objects (chair, door, wall, handrail) were tested based on the behaviors of the three intimacy behaviors between bodies and objects: T, TL-P, and TL-S, which were analyzed in this study. From the testing results of the sensing values, as shown in Figure 5, it is known that under three different sensing conditions of intimacy, different value patterns will appear. In the testing, the same behavior and action will be tested repeatedly 4 to 6 times, and the sensing value patterns obtained under the same action test will remain similar, which means that the sensing function is stable and reliable.

After completing the sensor test, the experiment of sensor contribution and data collection was conducted in five places in the elderly's home, including the entrance, living room, kitchen, bedroom, bathroom, and staircase. Seven Presence Stickers were affixed to the surfaces of different objects or space elements, including entrance's door, toilet's door, stair handrail 1 (lower step), stair handrail 2 (higher step), sink cabinet, bedroom door, and corridor wall. After the sensors were set up, we conducted a one-week test and collected the accumulated sensing data of the elderly subject's behaviors and activities at home.

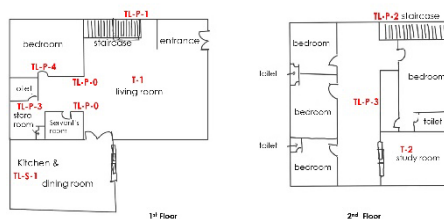


Figure 4. Sketch of the subject's house floor plan and sensors distribution

#### 5. Behaviors classifications from accumulated data

The cumulative performance data of all the sensors collected from the experiment are first distinguished by "space name ID", and then classified by three types of intimacy (T-#, TL-P-#, TL-S-#), and at the same time recorded with the Date and Time. Based on the record and analysis of the life pattern of the elderly subject as shown in Table 1, the collected accumulated sensing data will be logically defined according to the time and the relationship when different sensors are triggered to determine the daily life



behavior of the elderly subject. Figure 6 shows the sensor data schema and behavior classification for two periods of the accumulated data on different days:

---

*Date: 2021.12.8 (Figure 6)*

*The sensor data schema:* Corridor (TL-P-0) → Living (T-1) → Corridor (TL-P-0) → Bedroom (TL-P-4) → Bathroom (TL-P-3) → Bathroom (TL-P-3) → Bedroom (TL-P-4) → Corridor (TL-P-0) → Living (T-1)

*The behavior classification:* Walking (Pass by) → Sitting → Walking (Pass by) → Entering Bedroom → Entering Bathroom → Leaving Bathroom → Leaving Bedroom → Walking (Pass by) → Sitting

*The actual behavior and activities:* After three o'clock in the afternoon, the subject got up from a nap, went to the living room, sat in a chair and played with the phone, and then did some housework. After that, he went to the bathroom in the bedroom to take a shower, and then walked to the living room, and sat on a chair for a rest.

---

*Date: 2021.12.10 (Figure 6)*

*The sensor data schema:* Kitchen (TL-S-1) → Stair1 (TL-P-1) → Stair2 (TL-P-2) → Corridor (TL-P-3) → Study (T-2)

*The behavior classification:* Washing (Standing near the sink) → Climbing upstairs (lower step) → Climbing upstairs (higher step) → Walking (Pass by) → Sitting (Reading)

*The actual behavior and activities:* After seven o'clock in the evening, the elderly man stood in front of the sink to wash the dishes after he had had dinner.

---

When the behavior classification can be successfully defined completely, it can be connected to the Presence Stickers system through the IoT switch device to provide life assistance, such as switching lights, air conditioning, and other smart applications. Since this research has just completed the sensing module and behavior definition, it was just tested in the laboratory and it worked well, but the application of the feedback mechanism in the subject's home has not yet been tested.

## 6. Conclusion

The Presence Stickers sensing module in this study is made with low-cost capacitive sensing components. The feedback settings of the module can be customized, which is suitable for the home life of different elderly people. Meanwhile, more importantly, users can reach the smart living mechanism without changing any items in the home, so that the elderly can enjoy the convenience of technology-assisted life in the original familiar environment. Moreover, the collected sensing data and behavior classification can also be used as a record of the elderly's life patterns, and can also be provided as a reference for smart medical treatment.

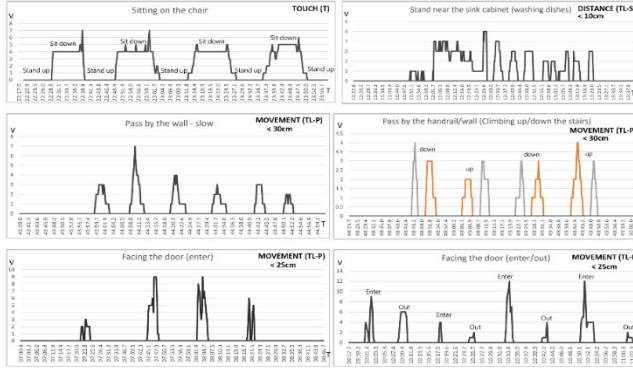


Figure 5. The sensing value patterns in different behaviors

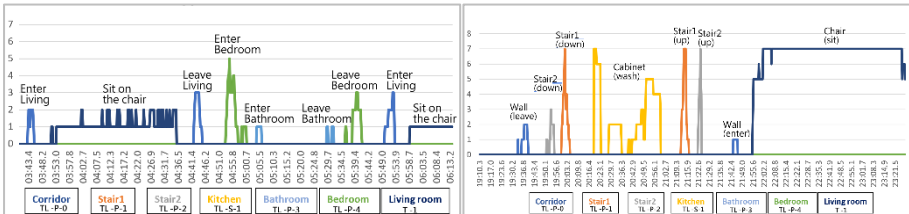


Figure 6. Sensor data schema and behavior classification

## References

- Arshad, A., Khan, S., Zahirul Alam, A. H. M., Abdul Kadir, K., Tasnim, R., & Fadzil Ismail, A. (2017). A capacitive proximity sensing scheme for human motion detection. *Paper presented at the 2017 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*.
- Braun, A., Wichert, R., Kuijper, A., & Fellner, D. W. (2015). Capacitive proximity sensing in smart environments. *Journal of Ambient Intelligence and Smart Environments*, 7(4), 483-510. doi:10.3233/ais-150324
- Greenberg, S., Marquardt, N., Ballendat, T., Diaz-Marino, R., & Wang, M. (2011). Proxemic interactions. *Interactions*, 18(1), 42-50. doi:10.1145/1897239.1897250
- Hall, E. T. (1966). *The Hidden Dimension*: Anchor Books.
- Holman, D., & Vertegaal, R. (2008). Organic user interfaces: Designing computers in any way, shape, or form. *Communications of the ACM*, 51(6), 48-55. doi:10.1145/1349026.1349037
- Ishii, H., & Ullmer, B. (1997). Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms. *Paper presented at the CHI*.
- Nabil, S., Kirk, D. S., Ploetz, T., & Wright, P. (2017). Designing Future Ubiquitous Homes with OUI Interiors: Possibilities and Challenges. *Interaction Design and Architecture(s) Journal*, 28-37.
- Sato, M., Poupyrev, I., & Harrison, C. (2012). Touché: Enhancing Touch Interaction on Humans, Screens, Liquids, and Everyday Objects. Paper presented at *the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.
- Zimmerman, T. G., Smith, J. R., Paradise, J. A., Allport, D., & Gershensfeld, N. (1995). Applying Electric Field Sensing to Human-Computer Interfaces. Paper presented at *the CHI Mosaic of Creativity*.

# QUANTIFYING THE IMBALANCE OF SPATIAL DISTRIBUTION OF ELDERLY SERVICES WITH MULTI-SOURCE DATA

PIXIN GONG<sup>1</sup>, XIAORAN HUANG<sup>2</sup>, CHENYU HUANG<sup>3</sup> and MARCUS WHITE<sup>4</sup>

<sup>1,2,3</sup>*School of Architecture and Art, North China University of Technology, <sup>4</sup>Swinburne University of Technology.*

<sup>1</sup>*gongpixinn@ncut.edu.cn, 0000-0003-4795-666X*

<sup>2</sup>*xiaoran.huang@ncut.edu.cn, 0000-0002-8702-2805*

<sup>3</sup>*huangchenyu303@163.com, 0000-0002-6360-638X*

<sup>4</sup>*marcuswhite@swin.edu.au, 0000-0002-2238-9251*

**Abstract.** With the growing challenge of aging populations around the world, the study of the elderly service is an essential initiative to accommodate the particular needs of the disadvantaged communities and promote social equity. Previous research frameworks are very case-specific with limited evaluation indicators that cannot be extended to other scenarios and fields. Based on multi-source data and Geographic Information System (GIS), this paper quantifies and visualises the imbalance in the spatial distribution of elderly services in 218 neighbourhoods in Shijingshan District, Beijing, China. Mortality data were obtained, and the most contributing indicators to mortality were investigated by correlation analysis. Finally, mapping between other facility indicators to mortality rates was constructed using machine learning to further investigate the factors influencing the quality of elderly services at the community level. The conclusion shows that the functional density of transportation facilities, medical facilities, living services facilities, and the accessibility of elderly care facilities are most negatively correlated with mortality. The correlation conclusion is combined with a machine learning prediction model to provide future recommendations for the construction of unbalanced elderly neighbourhoods. This research offers a novel systematic method to study urban access to elderly services as well as a new perspective on improving social fairness.

**Keywords.** Elderly Service Facilities; Multi-source Data; Machine Learning; SDG 3; SDG 10; SDG 11.

## 1. Introduction

Data from China's Seventh Census indicates that the proportion of the resident population aged 60 and above in Beijing has reached 19.6%. The elderly dependency coefficient has been rapidly increasing, boosting pressure on social pension and other political and economic challenges. With the looming challenges of an aging

population, the study of the elderly service access is an essential initiative to accommodate the particular needs of the disadvantaged communities and further promote social equity.

In the current rapid aging societies, the study of elderly care facilities has received extensive attention from scholars, both domestically and internationally. New data and algorithms are being integrated into the area of study due to the introduction of new urban science. Bingqiu et al., (2014), used structural equation modeling (SEM) to examine the satisfaction of seniors concerning the elderly services and living environments. Junling et al., (2017), used validated and psychometrically tested measures to assess each neighbourhood's social and physical attributes. Yanan and Pak-Kwong, (2017), examined the associations between walkability related environmental attributes and health-related quality of life. Fan et al., (2019), used the Gaussian two-step floating catchment area (G2SFCA) method to assess the accessibility to community-based services. Fan and Dezhi, (2019), adopted multiple linear regression (MLR) and structural equation modeling (SEM) to test the integrated model for community-dwelling older adults, and developed a mediation model called "Neighborhood Environment-Quality of Life (NE-QoL)" for community-dwelling older adults (Fan and Dezhi, 2019). Nicia et al., (2020), analysed the available instruments intended to measure the quality of life of institutionalised older adults. Ma et al., (2021), proposes a method to identify the ideal distribution of elderly facilities. However, several gaps can be found from the existing research: firstly, the data sources are mostly outdated and can no longer meet the research requirements; secondly, many studies overlook the influence of perceptible walking environment. Meanwhile, some research frameworks are very case-specific that cannot be extended to other scenarios.

Based on the above, we conducted a sub-district level analysis to determine the imbalanced distribution of elderly services. Based on multi-source data, correlation analysis and machine learning regression analysis were carried out to further investigate the factors influencing the quality of elderly services at the community level.

## **2. Methods**

### **2.1. FRAMEWORK**

Supported by open-source data and machine learning algorithms, 218 neighborhoods of 9 sub-districts selected from Shijingshan District were studied. Researched on the community-level distribution of elderly services, we used multi-source data to construct 20 indicators divided by three major categories, namely indicators of elderly population, indicators of elderly service circle quality, and indicators of elderly care facilities quality. Data was normalised to undertake a comparison of the quality of elderly services at the sub-district level. And then, the Spearman approach is used to conduct a correlation analysis. Finally, a supervised machine learning regression analysis and modelling study was carried out to further investigate the factors influencing the quality of elderly services at the community level.

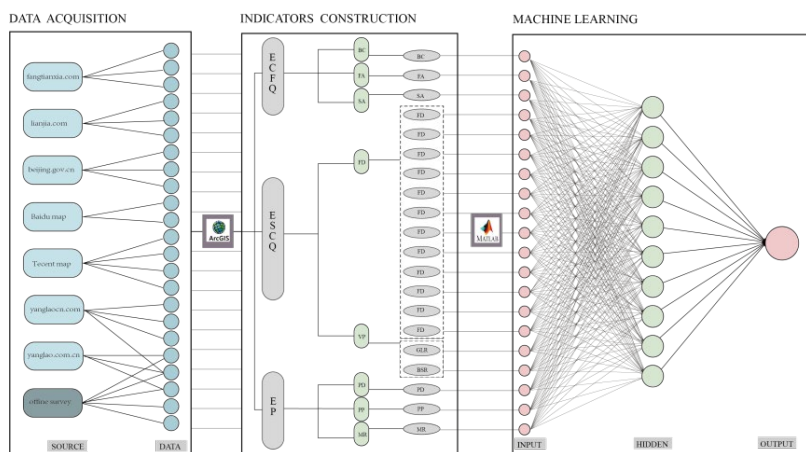


Figure 1. The research workflow

## 2.2. RESEARCH SITE

The Shijingshan District has a resident population of 567,851, with 137,768 individuals aged 60 and above accounting for 24.3%, according to the 7th National Census (2021). In some sub-districts, such as Bajiao, the elderly population is extremely large: "the household population of elderly people account for 29.7% of the total population, and the proportion in a few communities even reaches 40%, far exceeding the average level in Beijing". We conducted research of elderly service accessibility on 218 neighborhoods selected from Shijingshan District.



Figure 2. Research site

## 2.3. CALCULATING THE INDICATORS OF “ELDERLY CARE FACILITY QUALITY (ECFQ)”

### 2.3.1. Beds Coverage (BC)

Data on the number of beds in elderly care facilities came from two sources: online websites (yanglao.com.cn) for the number of beds in commercial elderly care facilities, and offline surveys for the number of beds in community-based elderly care facilities. In this paper, a bed-per-100-person calculation was conducted to create the indicator of beds coverage.

2.3.2. Facility Accessibility (FA)

The spatial accessibility of facilities is a crucial criterion for assessing the rationality of elderly service distribution. This research uses the Gaussian 2-step Floating Catchment Area (G2SFCA) approach, which can account for the influence of supply scale, demand scale, and spatial impedance factors between supply and demand sites on elderly care facility accessibility (Dajun et al., 2011). G2SFCA method estimates the accessibility of elderly care facilities in two phases, depending on the place of supply and demand, respectively. In this paper, the supply and demand sides are elderly care facilities and neighbourhoods. Formulas are shown in Table 1.

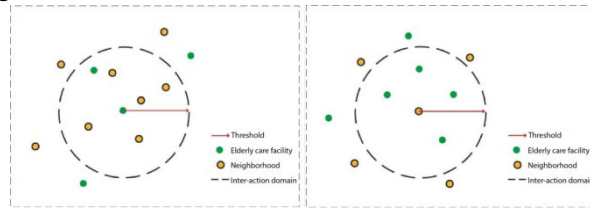


Figure 3. Process diagram of 2SFCA (from left to right: Step 1: Calculating the supply/demand ratio; Step 2: calculating the accessibility of elderly care facilities).

2.3.3. Street Accessibility (SA)

The impedance size is employed to study the street accessibility, and it is calculated using the O-D cost matrix approach. The ease with which plots can be connected is reflected by the least impedance at any two nodes connecting plots. A greater number indicates a larger impedance, less geographical accessibility, and more spatial fragmentation (Puebla, 1996). The following formulas shown in Table 1 are used to calculate node accessibility and road network accessibility. In this paper, Baidu Map was used to gather road network data, which was then cleaned and analysed using ArcGIS. Following that, a road network model was created to make the calculations of street accessibility.

Table 1. Formulas of the indicator FA and SA

INDICATOR	FORMULA	DESCRIPTION
Facility Accessibility	Step1 : $R_j = \frac{S_j}{\sum_{k \in (d_{kj} \leq d_0)} G(d_{kj}, d_0) P_k}$	where: $R_j$ is the supply-demand ratio of elderly care facility $j$ ; $S_j$ is the supply capacity of elderly care facility $j$ ; $\sum_{k \in (d_{kj} \leq d_0)} G(d_{kj}, d_0) P_k$ is the total demand for elderly care facility $j$ within the search threshold. $P_k$ is the number of elderly people at demand point $k$ ; $d_{kj}$ is the spatial distance between demand point $k$ and elderly care facility $j$ ; $d_0$ is the distance threshold from the residential point to the elderly care facility and $G(d_{kj}, d_0)$ is a Gaussian equation
	$G(d_{ij}, d_0) = \begin{cases} e^{-(1/2) \times (d_{ij}/d_0)^2} - e^{-(1/2)}, & \text{if } d_{ij} \leq d_0 \\ 0, & \text{if } d_{ij} > d_0 \end{cases}$	
Street Accessibility	Step2 : $A_i = \sum_{j \in (d_{ij} \leq d_0)} G(d_{ij}, d_0) R_j$	Where: $A_i$ is the accessibility of elderly care facilities. $R_j$ is the supply-demand ratio of elderly care facilities in the inter-action domain.
	$H_i = \frac{1}{n-1} \sum_{j=1}^n 1/(j+i) (d_{ij}) , H = \frac{1}{n} \sum_{i=1}^n (H_i)$	Where : $H_i$ is the accessibility of network node ; $H$ is the accessibility of the whole network; $d_{ij}$ indicates the minimum impedance between nodes $i, j$ ; $n$ is the number of nodes in the network.

2.4. CALCULATING THE INDICATORS OF “ELDERLY SERVICE CIRCLE QUALITY (ESCQ)”

Older adults are less likely to travel long distances or make complex trips (Sekhar and Matthew, 2013). Previous studies on elderly services typically used administrative borders as the research scope, which did not correlate to the actual residential scope of the elderly. To solve this problem, we introduced the concept of the elderly service circle, which is based on a 15-minute walking range.

We investigated to see if the rationale for choosing the 15-minute walking range as the study's scope was acceptable. Eight types of facilities were chosen to compare the two types of elderly service circles, 15-minute walking circle (WC), and circular buffer zone (CBZ). To establish a comparison, the scatter diagram is employed, demonstrating that the WC-based elderly service circle is more realistic.

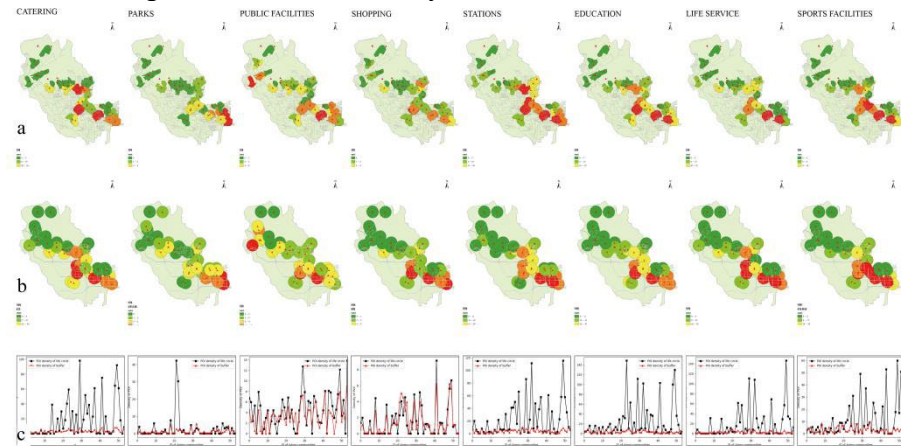


Figure 4. Comparison of the two types of elderly service circles (a:WC;b:CBZ;c:Scatter diagram)

#### 2.4.1. Functional Density (FD)

This category of indicators focuses on twelve kinds of functional POI (Point Of Interest) density linked to the living of the elderly, (including: elderly care facilities, medical facilities, catering, parks, public facilities, shopping facilities, transportation stations, education facilities, life services facilities, sports facilities, government institutions and housing). The data was transformed into Geospatial data with the use of ArcGIS platform, then connected to each neighbourhood's elderly service circle to make the calculation of each kind of functional density.

#### 2.4.2. Visual Perception (VP)

A pleasant outdoor atmosphere is a key factor in attracting older adults to spend time outside. The green looking ratio (GLR) and the blue sky ratio (BSR) are two noticeable measures based on streetscape pictures. A total of 1346 location points were gathered to get 5384 streetscape pictures (4 fetching directions), with using the API interface of Tencent Map. Image semantic segmentation algorithms are used to calculate these two indicators. Semantic segmentation divides a picture into categories by using a convolutional neural network (CNN) to categorise the pixels.

## 2.5. CALCULATING THE INDICATORS OF “ELDERLY POPULATION”

### 2.5.1. Elderly Proportion (EP).

The proportion of elderly population reflects the degree of ageing in the community. Due to limited data availability, this indicator uses the proportion of the elderly population at the Sub-district level published on the official website of the Bureau of Statistics.

### 2.5.2. Elderly Density (ED).

This category indicator was created using data from the 2020 Statistical Yearbook, the 7th National Population Census, and Beijing's 100m raster population data, with geographical data calculated using ArcGIS.

### 2.5.3. Mortality Rate (MR).

The quality of elderly service system is linked to life expectancy and physical health. Therefore, in this study, in order to highlight the influence of the elderly service on the lives of the elderly, the mortality ratio of the old population was employed as an indicator reference.

## 2.6. CORRELATION ANALYSIS

Given that each indicator's range of values varies, a linear transformation is performed through normalisation, with the values mapped to the range [0,1] to avoid the impact of differing magnitudes. The following are the transformation functions.

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)}$$

The variables were assessed for correlation using the Spearman approach and the Spearman rank correlation coefficient was obtained using SPSS software.

$$\rho = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}} = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

Where, n is the sample size,  $d_i = (x_i - y_i)$ ,  $x_i, y_i$  respectively represent the rank of the two variables sorted by size.

$\rho$  is the rank correlation coefficient. If the value is positive, there is a positive association, otherwise the opposite. The stronger the association, the higher the absolute value.

## 2.7. NEURAL NETWORK MODELLING

Machine learning regression analysis and modelling study was carried out to better investigate the factors influencing the quality of elderly services at the community-level. According to the machine learning algorithm framework, the above 19 indicators are selected as feature variables, and the mortality rate of the elderly is employed as a responsive variable for supervised regression analysis. In MATLAB, a shallow neural network is formed, and the model is tested using the hold-out approach. 70% of the



sample data is separated into training sets, 15% is used for verification, and 15% is used for the independent generalisation tests.

### 3. Results and discussion

#### 3.1. RESULTS OF CALCULATION

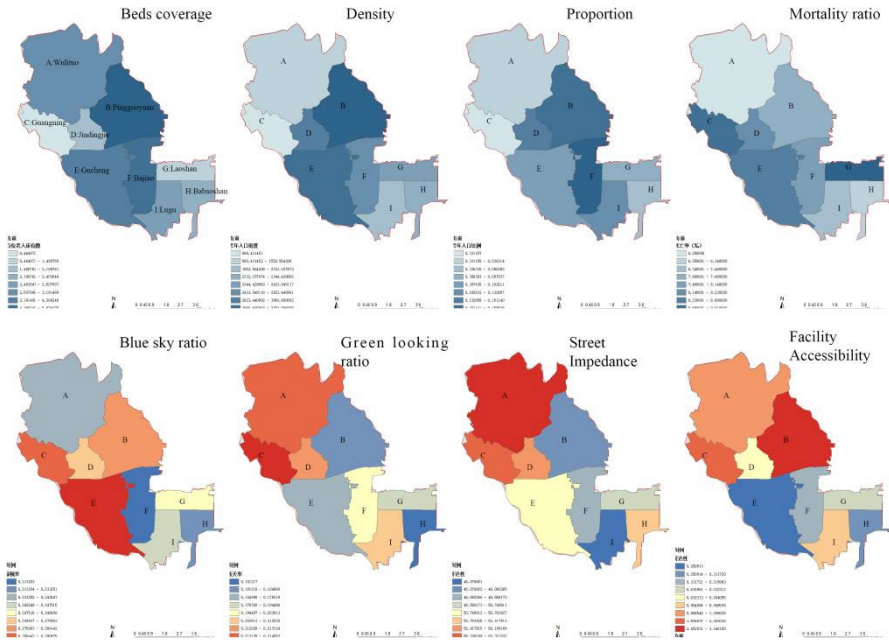
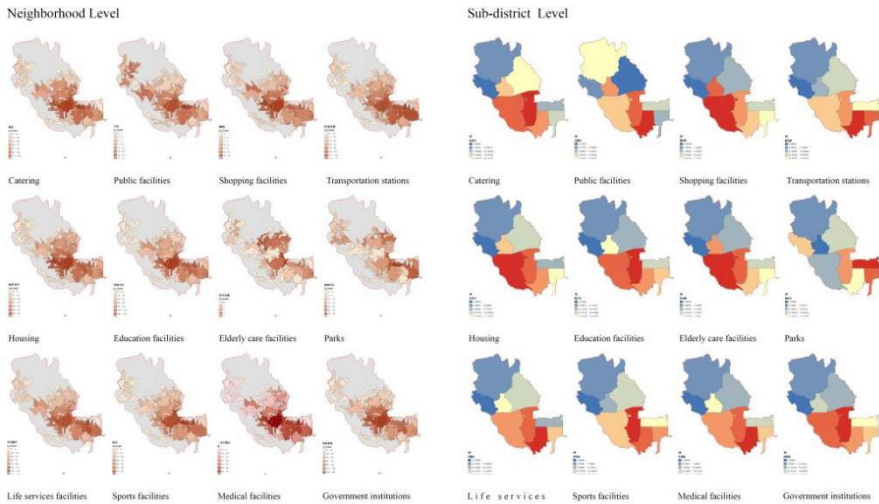


Figure 5 depicts the various indicators' clustering characteristics. Sub-district B and F have the largest proportion of elderly people, and Sub-district B and E have the highest density of elderly people, while Sub-district G and C have the highest mortality rates. The northern sub-districts have better access to elderly care facilities than the southern ones, with Sub-district B being the finest. The sub-districts with the highest green looking ratio and blue sky ratio are primarily on the west side, where mountains and rivers abound and the environment is of outstanding quality.



As shown in the graph figure 6, the indicators within each elderly service circle are mainly concentrated in the center of district, which is home to the government's commercial center, and the neighborhoods in these places have better access to services; at the sub-district level, the indicators follow the same pattern, with a high value in the south and a low value in the north, which is related to their location.

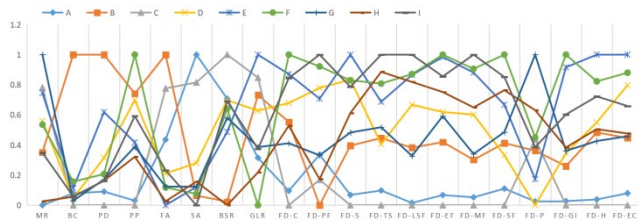


Figure 7. Results of normalisation of indicators

Indicators are normalised to create a line graph that can visually compare and analyse the indicators of each sub-district, as illustrated in figure 7, which shows the quality status at the sub-district level. For example, all the value of Sub-district A's indicators are low, which is related to its location on the periphery; Sub-district G got medium values on most of indicators, but its mortality rate (MR) is the highest; and Sub-district F have a high level of functional density (FD) of the elderly service circle.

### 3.2. RESULTS OF CORRELATION ANALYSIS

The degree of correlation between the indicators is shown in Figure 8. A) There is a substantial correlation between the indices of functional density in the elderly services circle, indicating a clustering effect. B) The accessibility of elderly facilities is mostly adversely connected with functional indicators, which reflects the city's previous

overemphasis on the economics, resulting in the neglect of the development of elderly service system. c) Indicators including the functional density of transportation facilities, medical facilities, living services facilities, and the accessibility of elderly care facilities have a high negative link with mortality. It is shown that the quality of elderly service circle, density of elderly care facilities and quality of street environment will influence the health of older adults significantly.

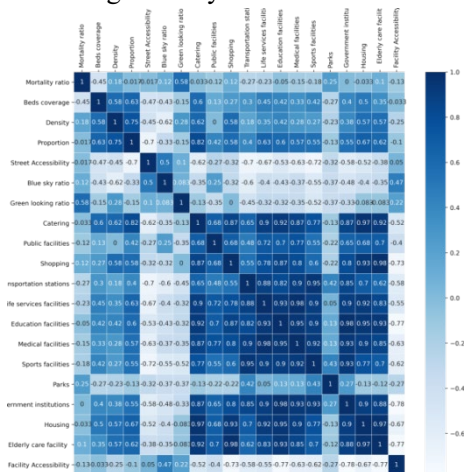


Figure 8. Spearman Correlation coefficient matrix

### 3.3. RESULTS OF NEURAL NETWORK MODELLING

As shown in Figure 9.(a), the MSE of the model gradually decreases as the number of model iterations grows for the three data sets, and the highest performance is attained at the tenth epoch. Figure 9.(b): the model's errors has a normal distribution of roughly -0.07756. Figure 9.(c) depicts the training results, on the training set:  $R^2 = 0.9485$ , on the entire data set:  $R^2 = 0.89195$ , which shows that the shallow neural network model regression works well. The results show that the neural network prediction model developed in this research has a lot of application potential.

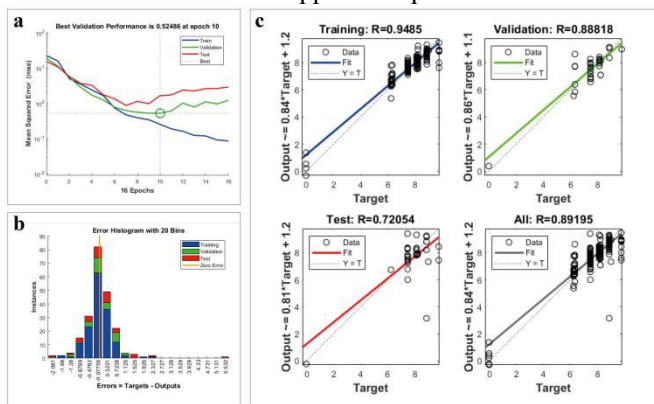


Figure 9. Neural network regression performance

#### 4. Conclusion

The investigation of the elderly service distribution is not only a trend in the scenario, but also a significant step toward ensuring the survival of vulnerable groups and promoting social fairness. Based on multi-source data and Geographic Information System (GIS), this paper quantifies and visualises the imbalance in the spatial distribution of elderly services in 218 neighborhoods in Shijingshan District, Beijing, China. Mortality data were obtained and the most contributing indicators to mortality were investigated by correlation analysis. Finally, mapping between other facility indicators to mortality rates was constructed using machine learning method, to further investigate the factors influencing the quality of elderly services at the community level. The study shows that the functional density of transportation facilities, medical facilities, living services facilities, and the accessibility of elderly care facilities are most negatively correlated with mortality. The neural network model created in this study performed well.

The study has several limitations, such as a lack of sample precision due to limited data availability and the difficulties in establishing a well-being measure for the elderly. This research offers a systematic way to study urban elderly service as well as a new perspective dimension to improve social fairness. If our approach was adopted by government and urban designers, they could identify problematic areas in a highly targeted manner not possible before now.

#### References

- Dai, D. (2011). Racial/ethnic and socioeconomic disparities in urban green space accessibility: Where to intervene? *Landscape and Urban Planning*, 102(4), 234-24.
- Gao, J et al. (2017). Relationships between neighborhood attributes and subjective well-being among the Chinese elderly: Data from Shanghai. *BioScience Trends*, 11(5), 516-523
- Puebla, G. (1996). Accessibility in the european union: a comparative analysis by transportation mode. *Estudios De Transportes Y Comunicaciones*.
- Santana-Berlanga, N.D., et al. (2020). Instruments to measure quality of life in institutionalised older adults: Systematic review. *Geriatric Nursing*, 41(4), 445-462.
- Yan, B., Gao, X., & Lyon M. Modeling satisfaction amongst the elderly in different Chinese urban neighborhoods. *Social Science & Medicine*, 118, 127-134.
- Zhang, F., et al. (2019). Assessing spatial disparities of accessibility to community-based service resources for Chinese older adults based on travel behavior: A city-wide study of Nanjing, China. *Habitat International*, 88, 101984-101984.
- Zhang, F. and Li, D. (2019). Multiple Linear Regression-Structural Equation Modeling Based Development of the Integrated Model of Perceived Neighborhood Environment and Quality of Life of Community-Dwelling Older Adults: A Cross-Sectional Study in Nanjing, China. *International Journal of Environmental Research and Public Health*, 16(24), 4933.
- Zhang, F. and Li, D. (2019). How the Urban Neighborhood Environment Influences the Quality of Life of Chinese Community-Dwelling Older Adults: An Influence Model of "NE-QoL". *Sustainability*, 11(20), 5739-5739.
- Zhao, Y., and Chung, P.K. (2017). Neighborhood environment walkability and health-related quality of life among older adults in Hong Kong. *Archives of Gerontology and Geriatrics*, 73, 182-186.

# MACHINE-READING PLACES & SPACES

## *Generative Probabilistic Modelling of Urban Thematic Zones & Contexts*

NUOZHI LIU<sup>1</sup> and IMMANUEL KOH<sup>2</sup>

<sup>1,2</sup>*Singapore University of Technology and Design.*

<sup>1</sup>*nuozhi\_liu@sutd.edu.sg, 0000-0002-3709-6822*

<sup>2</sup>*immanuel\_koh@sutd.edu.sg, 0000-0002-1181-1082*

**Abstract.** In this paper, a "place" is conceptualised as a composition of dynamic socioeconomic activities and collective perceptions. We apply generative probabilistic modelling to explore urban contextual semantics. By analogy to sorting documents into different topics, this research retrieves data embedding for each urban regions and classify them with thematic zones. Using Singapore as a case study, topic modelling is applied to retrieve perceptual and functional thematic zones from Instagram and TripAdvisor respectively. A subsequent analysis shows strong correlations among certain regions with functional and perceptual consistency. In addition, with our proposed uniqueness and diversity indices, a strong negative correlation at 0.82 is found, suggesting that a region could be more unique if the functions tend to be dominated by certain types of functional and perceptual thematic zones.

**Keywords.** Machine Learning; Natural Language Processing; Generative Probabilistic Models; Urban Data Modelling; Thematic Zones; Topic Modelling; SDG 11.

## 1. Introduction

In the city, a place is associated with a space that provides facilities in supporting socio-economic activities. The sense of places in turn plays a critical role in affecting how people feel about these spaces. With rapid urbanization and globalization, urban planners need to acknowledge and be attuned to the diversity of places. More urgently, in view of the United Nations' Sustainable Development Goals such as SDG 11 (Sustainable cities and communities), urbanization approaches that recognize the diversity of places could significantly make cities more inclusive and participatory. Previous studies on place functions and human behaviours have provided insights on the features of places, and consequently, facilitate in the proper classification of urban regions. Discovering the functions of places enables planners to monitor change of land use (Gao et al., 2017) as well as to allocate resources accordingly (Yuan et al., 2012). However, despite prior observations of these functional characteristics, it remains unclear how places correlate to users' perceptions. For example, the perception of any

given restaurant in the city may differ according to its urban context (e.g., Central Business District (CBD) as compared to a school) despite it playing the same function of food and beverages. Hence, the purpose of this research is to study the relationship between space functions and human perceptions of urban regions. It aims to answer whether and how similar function of spaces can have an impact on people's feeling about places. It also investigates how patterns of urban regions might be inferred from their urban contextual semantics. It demonstrates both functional and perceptual thematic zones beyond their existing homogeneously allocated administrative districts, in order to illustrate how a spatial region could be re-composed by socio-economic activities as well as human perceptions. The datasets used for this research are taken from both TripAdvisor and Instagram. By analogy to reading documents, this research employed topic modelling techniques to read a city. Both Latent Dirichlet allocation (LDA) and supervised LDA have been applied to the datasets in transforming the sparse urban data to dense informative data embedding. This paper contributes to the formulation of a proposed framework in extracting the latent structure of features from the regions of interests. The latent structure defines two types of thematic zones – a functional thematic zone for describing the physical characteristics and a perceptual thematic zone for exploring the human perceptions of the space. To measure the similarity among regions and the correlation between these two aspects of data, the research uses a diversity index and uniqueness index (as derived from Jensen-Shannon divergence). Both indices are meant for evaluating the urban-scale contexts by comparing regions of interests.

## 2. Background

The sense of place relating to place attachment, identity, and place making has been studied previously by others. Most frequently adopted methodologies by researchers are questionnaires, interviewing, and gathering feedbacks, and it could take a long period of time in collecting data (Lewicka, 2008; Savage et al., 2004; Shamai & Ilatov, 2005; Yuen, 2005). However, it is difficult to make use of the fully collected dataset, because the raw data typically consists of unstructured texts such as captions, comments, feedbacks, and open-ended questions. Furthermore, due to the diverse backgrounds of participants, summarising human perceptions from texts could also be challenging. Instead, this research introduces generative probabilistic modelling to synthesize such data in an unsupervised manner. Moreover, as the discussion of sense of places hitherto remains inconclusive (Cross, 2001; Hashemnezhad et al., 2013), this research aims to shed light on the field of urban studies via a statistical approach.

Another important task for urban studies is to detect and estimate the land use in Geographic Information System (GIS). It has been achieved by using census surveys, municipal records, and satellite imagery, etc. Meanwhile, with an unprecedented scale of information generated by users on the internet, emerging machine learning algorithms have brought new opportunities for understanding the built environment. For example, "social sensing" is proposed to discover socioeconomic environments through social network data by considering individuals as spatial detectors (Ali et al., 2011; Liu et al., 2015; Wang et al., 2019). One of the most popular techniques being LDA. It is a probabilistic topic model for uncovering the hidden thematic structures in large volume of documents, such as, books, news, articles, etc (Blei et al., 2003). While

others have extended it to urban studies by revealing spatial functions, incorporating geospatial data and location-based social media like Twitter and Foursquare (Gao et al., 2017; Ghahramani et al., 2021; Yuan et al., 2012), these studies have mainly focused on urban land use with limited attention paid to human crowd perception and without any synthesis of both. Therefore, our research attempts to evaluate and synthesize both urban functions and human perceptions via a quantitative approach that could complement further qualitative studies by applying variational inference of unstructured data at scale.

### 3. Methodology

#### 3.1. RESEARCH DESIGN

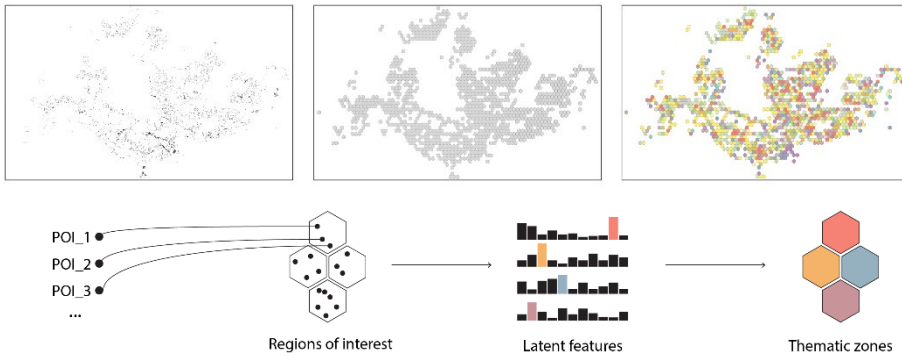


Figure 1. An example of data processing shows perceptual thematic zones from Instagram (above). Data aggregation based on regions of interests (below).

The collected raw data are points of interests (POIs) which are geo-locations with attributes in texts. To aggregate attributes based on analysis unit, regions of interests are introduced. Similar to 'Discovers Regions of different Functions' or DRoF (Yuan et al., 2012) but simplified, a region of interest is a geographic region with elements where the elements are aggregated attributes of POIs inside. By defining the geographic regions, Singapore territory is subdivided into hexagons of 500 meters radius. As shown in Figure 1, each hexagon is regarded as a region of interest, aggregating POIs inside. If no POI is found, the region will be removed. LDA is then applied to these regions of interest in order to extract the latent structure of features. Each region will be categorised with a thematic zone based on its latent features. This procedure is applied to two datasets, namely Instagram for perceptual thematic zones and TripAdvisor for functional thematic zones. The correlation between function and perception in the same region can thus be analysed and measured.

#### 3.2. DATA COLLECTION AND PRE-PROCESSING

Instagram is one of the major social media platforms in Singapore, as people use Instagram to document their life and interact with their family members and friends. Although Instagram is an image-based platform, the captions used constitute a directly

suggestive form of how people express themselves. A crawler was run to retrieve captions of Instagram posts from all locations in Singapore during the year of 2019. Text pre-processing, such as tokenization and lemmatisation, was then applied to these captions, splitting texts into words while removing tenses, plurality, etc. However, the “#” tags were kept during the process as they are representative of meaning rather than of punctuation in the context of Instagram.

For datasets that are representative of urban functions, we use TripAdvisor's place metadata. Compared to general Map service providers such as Open Street Map, TripAdvisor captures more detailed features about a place's characteristics, such as types of cuisines, price ranges, etc. Furthermore, urban functions should also consider dynamic demands, rather than a static usage. For example, when considering how people visits functional places, this research collected passenger volume data of the most popular transport modes in Singapore, including Mass Rapid Transit (MRT) and Light Rail Transit (LRT). These volumes were grouped using transition cuboid into 24 time bins (12 for weekdays and 12 for weekends) (Yuan et al., 2012). Both TripAdvisor and traffic data are inputs to our supervised LDA to retrieve functional thematic zones.

### 3.3. TOPIC MODELLING

Probabilistic topic models are based on word frequencies. For instance, given two papers, one with high occurrence of the words "urban" and "plan", and the other with the words "probability" and "Markov chain", intuition will suggest that the former may be related to topics in urban planning, while the latter to topics in statistics. However, in practice, a paper could be a mixture of various topics in different proportions. For example, a paper related to generative urban design may mention the words "urban" and "probability", thus suggesting a mix of topics from urban planning and statistics. The definition of a *topic* of a paper, in such a context, thus refers to a *distribution* of topics instead of a single 'topic', and the goal of topic models is to capture this kind of intuition.

One of the most popular topic models is LDA. Given the words contained in a set of documents, LDA performs a generative process that infers the latent thematic structure through a posterior distribution. It considers each document as a distribution over topics, and each topic is a distribution over words. The generative process computes a joint probability distribution over both the observed and hidden random variables, where the documents and words are observed, and the topics are hidden (Blei, 2011; Blei et al., 2003). In extending this algorithm to urban studies, the methodology from Yuan et al. (2012) has been adapted for this research. Each region is considered a document, while the elements within the region liken to the words within the document. The elements are either keywords extracted from Instagram or functions from TripAdvisor's metadata as explained in Section 3.2. The hidden random variables are thematic zones. In addition, the vocabulary of the documents refers to processed attributes gathered from all POIs, or in other words, our metadata.

*Thematic zones* provide another perspective to see the relation between regions and zoning. Unlike how standard zoning in urban planning sees residential and commercial as disparate zones, a thematic zone mixes them in different proportions. For example, one thematic zone might have more residential areas with little commercial activities, while another might have more commercial functions but with few residential



buildings. These two thematic zones should not be simply regarded as either residential or commercial zones. Therefore, LDA is applied to uncover these thematic zones. As shown in Figure 2, the region belongs to certain thematic zone which is defined by mixing 3 themes such as foods, happy family, and fitness. Each theme refers to a specific subset of elements retrieved from the metadata. Such a definition of thematic zone will also have a greater capability in abstracting senses collectively, such as human perception which tends to be more complex as compared to functions.

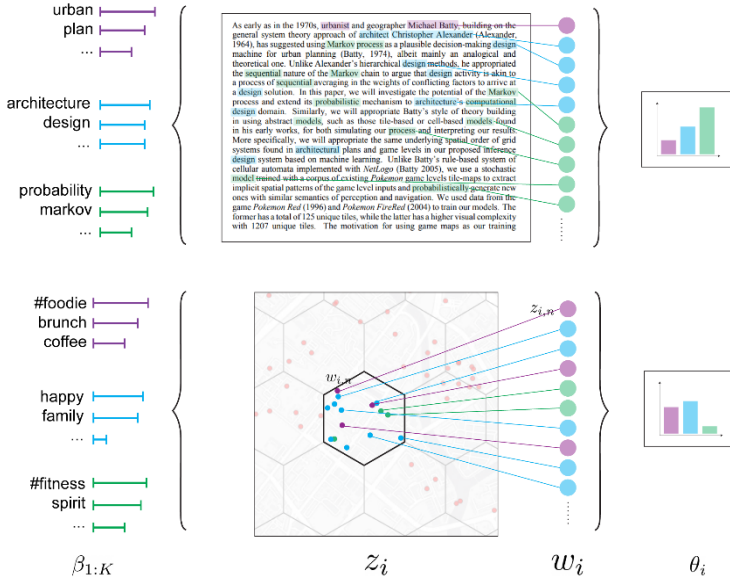


Figure 2. Analogy from documents/topics to regions of interest/thematic zones

The notations are defined as follows:  $K$  number of thematic zones are denoted as  $\beta_{1:K}$ . Each  $\beta_k$  is a distribution over metadata. Proportion of theme  $k$  in the  $i$ th region is denoted as  $\theta_{i,k}$ , while the theme of this region  $\theta_i$ . Inferred theme of element of the region is denoted as  $z_{i,n}$ , while  $w_{i,n}$  denotes the observed element. The value of  $n$  is the index of the element. The joint posterior distribution over the inferred and observed variables is described as:

$$p(\beta_{1:K}, \theta_{1:I}, z_{1:I}, w_{1:I}) \\
= \prod_{j=1}^K p(\beta_j) \prod_{i=1}^I p(\theta_i) \left( \prod_{n=1}^N p(z_{i,n} | \theta_i) p(w_{i,n} | \beta_{1:K}, z_{i,n}) \right)$$

This research employed supervised topic modelling to include the factor of traffic behaviours as part of consideration of functional thematic zones. A typical approach of supervised LDA is using authors as additional references to enhance the topic modelling of articles (Roberts et al., 2013). In contrast, this research uses transition cuboid as additional references to complement TripAdvisor dataset.

The number  $K$  of thematic zones was chosen based on the recommended indicators, including residuals, semantic coherence and exclusivity. The goal is to find a  $K$  number with a lower residual but with a relatively higher semantic coherence and exclusivity.

### 3.4. DIVERSITY INDEX AND UNIQUENESS INDEX

For the region of  $i$ , the output values of  $\theta_i^{TA}$  and  $\theta_i^{IG}$  respectively indicate the functional theme distribution and perceptual theme distribution. The research concatenates both  $\theta$  values into one feature vector for each region. The similarity between regions of interest could be measured using the distance matrix. Figure 3. shows the selection of distance measures. It first considers the regions of three areas – (1) Marina Bay area in CBD, (2) Punggol neighbourhood area and (3) Changi airport area. These three areas are then considered as three distinctive areas from one another in terms of both functions and people activities. Both cosine distance and Jensen-Shannon Divergence (JSD) are capable in separating these areas, although JSD performed best in showing the nuances between regions within the same area.

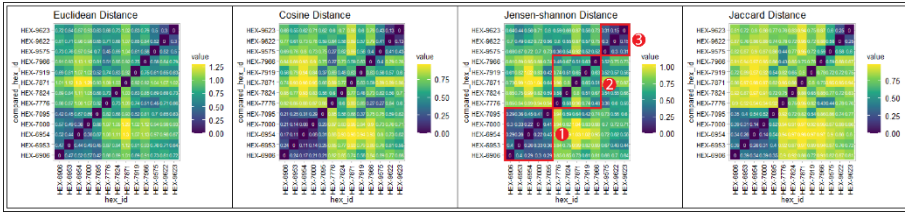


Figure 3. Comparisons of distance measures over regions of interest. From left to right: Euclidean distance, cosine distance, Jensen-Shannon distance, Jaccard distance.

Based on the  $\theta$  values and JSD, this research introduces the *Uniqueness* and *Diversity* indices. Diversity index compares the difference within the regions, while uniqueness index compares the difference crossing regions with aggregation of JSD.

$$\text{Uniqueness } U_i = \sum_{j=1} D_{ij}, j \in \{1, 2, \dots, I\}, \forall i, \text{ where, } D_{ij} = \text{JSD}(P \parallel Q)$$

$$P = (\theta_i^{TA_1}, \theta_i^{TA_2}, \dots, \theta_i^{TA_K}, \theta_i^{IG_1}, \theta_i^{IG_2}, \dots, \theta_i^{IG_K})$$

$$Q = (\theta_j^{TA_1}, \theta_j^{TA_2}, \dots, \theta_j^{TA_K}, \theta_j^{IG_1}, \theta_j^{IG_2}, \dots, \theta_j^{IG_K})$$

$$\text{Diversity } H_i = -\sum p_i \log p_i$$

## 4. Results and Discussions

The model retrieved 30 thematic zones (prefix “TA”) for TripAdvisor, and 30 thematic zones (prefix “IG”) for Instagram. For the functional thematic zones, more differences have been observed. For example, as shown in Table 1, Zone TA-27 contains functions

that are mostly bars and clubs, while TA-7 contains more varied activities, events, and leisure parks. Mid-range restaurant takes up the largest proportion for both Zone TA-2 and Zone TA-9. However, the former consists of more hotel types, while the latter more cuisine types.

Table 1. Top 10 ranked elements of functional thematic zones

Theme	To 10 Ranked elements	Interpreted tag
TA-2	Brew Pub, Central Asian cuisine restaurant, Gastropub, Hotel (> \$300), Marina View Hotel, Mid-range restaurant, Other Outdoor Activities, Scandinavian cuisine restaurant, Swedish cuisine restaurant, Wine Bar	Mid-range restaurant
TA-7	Equipment Hire, Fun & Games, Game & Entertainment Centres, Gardens, Nature & Parks, Outdoor Activities, Playgrounds, Room Escape Games, Sports Camps & Clinics, Sports Complexes	Activities
TA-9	Cafe, Cheap restaurant, Chinese cuisine restaurant, Healthy, Italian cuisine restaurant, Medicinal foods, Mid-range restaurant, Singaporean cuisine restaurant, Soups, Trendy Hotel	Cafe & Restaurant
TA-27	Bars & Clubs, Dance Clubs & Discos, Family Hotel, Fine-dining restaurant, French cuisine restaurant, Great View Hotel, Hotel (> \$300), Karaoke Bars, Nightlife, Wine Bars	Nightlife

As for perceptual thematic zones, many words are found frequently co-occurring in different thematic zones, but in varied contexts, such as, "time", "happy", "love", etc. For example, as shown in Figure 4, all 10 thematic zones mention the word "happy", with most zones mentioning the words "love" and "time" in parallel, while the nuances of contextual words serve to differentiate those zones. In addition, some thematic zones are found to be more ethnicity-related. In particular, IG-1 is dominated with Malay words, such as, "selamat" (congratulations), "kasih" (love), "hidup" (life), and "hati"(heart).

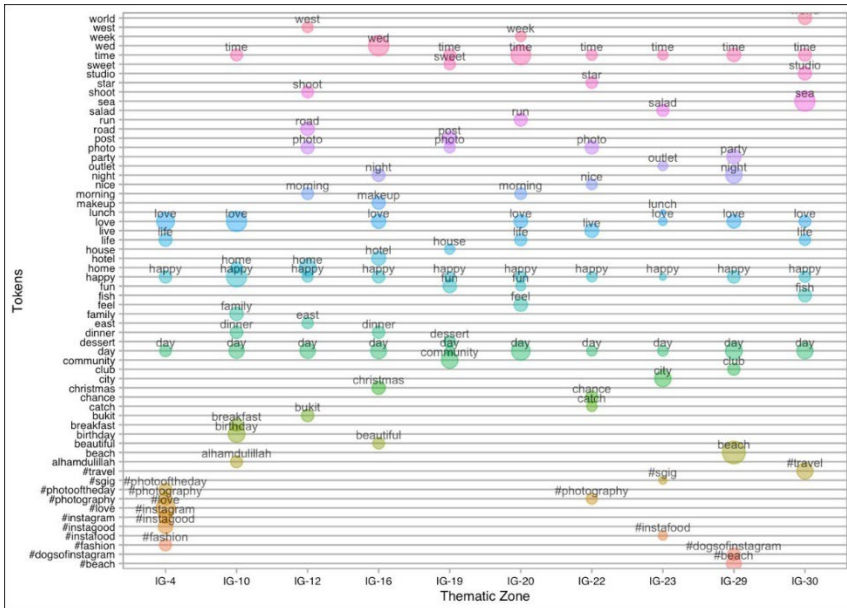


Figure 4. Perceptual thematic zones related to "happy"

Pearson correlation was computed between functional and perceptual thematic zones (Figure 5). By way of cross comparing, only few correlations can be found between functional and perceptual zones. Relatively high positive correlation can be found for TA-6 (Indian cuisine) / IG-13 (Hawker food), and TA-4 (Sightseeing & Landmarks) / IG-30 (Sea). For the former pair, it shows hawker food tends to be relatively consistent in space functioning and human perception, especially in Indian Cuisine, and those thematic zones are near Little India, Chinatown and Toa Payoh. The latter pair are instead found near Sentosa, East Coast Park, and Marina Bay Sands, which are typically regarded as tourist destinations in Singapore. There are also some negative correlations. For instance, TA-11 (Japanese & French cuisine) and IG-10 (Family events) could less likely co-exist in functions and perceptions.

In terms of the uniqueness index, we found that the most unique two regions are those near the National and Hougang stadiums. Neighbourhoods are supposed to be generic in Singapore. Surprisingly, some of the neighbourhoods are the unique regions themselves, such as Telok Blangah. Yet, some regions are not as unique as people expected. For example, Changi airport shows relatively lower uniqueness, especially the area near the Jewel, an entertainment and shopping complex inside the Changi airport. The reason could be due to the presence of similar shops, department stores, landscape, playground. Figure 5. shows the distribution of normalised uniqueness index. By showing the relation between diversity index and uniqueness index for all

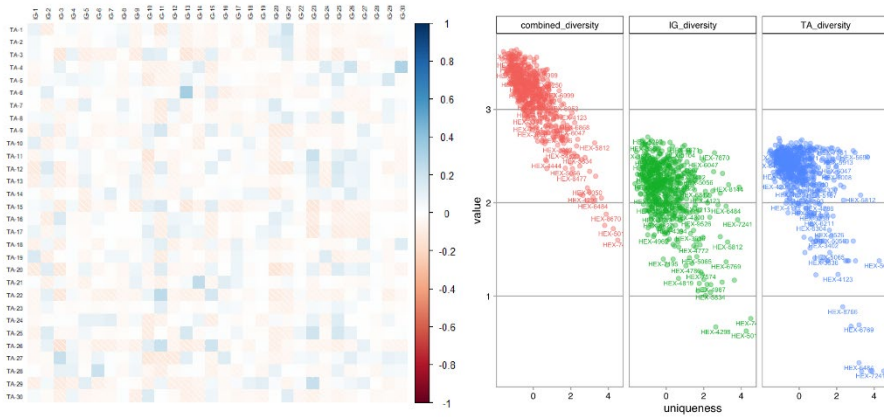


Figure 5. Pearson correlation between perceptual (horizontal axis) and functional thematic zones (left); uniqueness and diversity indices of regions (right)

regions, a strongly negative correlation can be found at  $-0.82$ . It implies that if a region becomes more unique in both functions and perceptions, the region tends to be dominant by certain types of functional and perceptual thematic zones.

## 5. Conclusion

We have explored how one might 'read' the city in the form of urban thematic zones with probabilistic topic modelling. By appropriating natural language processing and topic modelling for urban studies, we propose that an urban thematic zone should be regarded as a mixture of different scenarios of certain proportions. By using Singapore as a case study, we have used TripAdvisor and Instagram datasets to retrieve functional and perceptual thematic zones respectively. A further correlation analysis found that some thematic zones are consistent in functions and perceptions. Our proposed uniqueness and diversity indices have also found a strongly negative correlation. We believe our contribution lies in a methodology that could be complementary to research on sense of places via a statistical perspective on urban functions and perceptions. It could enable urban planners and designers to evaluate urban regions located in any cities around the world with a vast amount of unstructured data using machine learning in "reading" places at scale. In addition, such a statistical view of "thematic zone" also reframes questions like "what makes a place a place", by seeing a place as a proportional mixture of tangible and intangible elements. It challenges the notion of a place beyond that of a static function, but reckons it as that which is constituted by its dynamic socioeconomic activities and collective perceptions. To address the limitation of our current research such as dataset biasness, future works would be further verified in a qualitative manner, alongside a more comprehensive selection of machine learning models to compare factors like probability model perplexity.

## References

- Ali, R., Solis, C., Salehie, M., Omoronyia, I., Nuseibeh, B., & Maalej, W. (2011). Social sensing: When users become monitors. *Proceedings of the 19th ACM SIGSOFT Symposium and the 13th European Conference on Foundations of Software Engineering* (pp. 476–479).
- Blei, D. M. (2011). Introduction to probabilistic topic models. *Communications of the ACM*, 55(4), 77–84.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *The Journal of Machine Learning Research*, 3, 993–1022.
- Cross, J. E. (2001). *What is sense of place?* [PhD Thesis]. Colorado State University Libraries.
- Gao, S., Janowicz, K., & Couclelis, H. (2017). Extracting urban functional regions from points of interest and human activities on location-based social networks. *Transactions in GIS*, 21(3), 446–467.
- Ghahramani, M., Galle, N. J., Ratti, C., & Pilla, F. (2021). Tales of a city: Sentiment analysis of urban green space in Dublin. *Cities*, 119, 103395. <https://doi.org/10.1016/j.cities.2021.103395>
- Hashemnezhad, H., Heidari, A. A., & Mohammad Hoseini, P. (2013). Sense of place” and “place attachment. *International Journal of Architecture and Urban Development*, 3(1), 5–12.
- Lewicka, M. (2008). Place attachment, place identity, and place memory: Restoring the forgotten city past. *Journal of Environmental Psychology*, 28(3), 209–231. <https://doi.org/10.1016/j.jenvp.2008.02.001>
- Liu, Y., Liu, X., Gao, S., Gong, L., Kang, C., Zhi, Y., Chi, G., & Shi, L. (2015). Social sensing: A new approach to understanding our socioeconomic environments. *Annals of the Association of American Geographers*, 105(3), 512–530.
- Roberts, M. E., Stewart, B. M., Tingley, D., & Airoidi, E. M. (2013). The structural topic model and applied social science. *Advances in Neural Information Processing Systems Workshop on Topic Models: Computation, Application, and Evaluation*, 4, 1–20.
- Savage, V. R., Huang, S., & Chang, T. C. (2004). The Singapore River thematic zone: Sustainable tourism in an urban context. *Geographical Journal*, 170(3), 212–225.
- Shamai, S., & Ilatov, Z. (2005). Measuring sense of place: Methodological aspects. *Tijdschrift Voor Economische En Sociale Geografie*, 96(5), 467–476.
- Wang, D., Szymanski, B. K., Abdelzaher, T., Ji, H., & Kaplan, L. (2019). The age of social sensing. *Computer*, 52(1), 36–45.
- Yuan, J., Zheng, Y., & Xie, X. (2012). Discovering regions of different functions in a city using human mobility and POIs. *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 186–194.
- Yuen, B. (2005). Searching for place identity in Singapore. *Habitat International*, 29(2), 197–214. <https://doi.org/10.1016/j.habitatint.2003.07.002>

# MORPHOLOGICAL REGENERATION OF THE INDUSTRIAL WATERFRONT BASED ON MACHINE LEARNING

SHUYI HUANG<sup>1</sup> and HAO ZHENG<sup>2</sup>

<sup>1</sup>*School of Architecture, Tsinghua University, Beijing, China.*

<sup>2</sup>*Stuart Weitzman School of Design, University of Pennsylvania, Philadelphia, USA.*

<sup>1</sup>*huangsy21@mails.tsinghua.edu.cn, 0000-0001-5498-2730*

<sup>2</sup>*zhhao@design.upenn.edu, 0000-0001-5769-6035*

**Abstract.** The regeneration of the industrial waterfront is a global issue, and its significance lies in transforming the waterfront brownfield into an eco-friendly, hospitable, and vibrant urban space. However, the industrial waterfront naturally has comparatively unmanageable morphological features, including linear shape, irregular waterfront boundary, and separation with urban networks. Therefore, how to subdivide the vacant land and determine the land-use type for each subdivision becomes a challenging problem. Accordingly, this study proposes an application of machine learning models. It allows the generation of morphological elements of the vacant industrial waterfront by comparing the before-and-after scenarios of successful regeneration projects. The data collected from New York City is used as a showcase of this method.

**Keywords.** Machine Learning; Urban Morphology; Industrial Waterfront Regeneration; Sustainable Cities; SDG 11.

## 1. Introduction

### 1.1. RESEARCH BACKGROUND

Starting from the second half of the twentieth century, the recession of industry left lots of deteriorated or underutilized parcels, resulting in weakening ties between water and cities (Carpenter and Lozano, 2020). Accordingly, there arise the challenging problems of how to subdivide the vacant lands (Yang, 2014) and determine the long-term use for these waterfront parcels (Samant, 2017). Basically, these problems require a morphological regeneration, including forms and functions, of the waterfront.

### 1.2. PROBLEM STATEMENT

Urban morphology planning is a complex system that involves multiple design elements. According to Conzenian tradition, the urban form contains the street system, the building system (plots and the buildings located on them), and the land-use pattern (Conzen, 1960). In the last decades, several analysis methods or tools are introduced

globally to evaluate urban forms, such as Space Syntax (Hillier et al., 1993), Urban Network Analysis (Sevtsuk and Mekonnen, 2012), and Form Syntax (Ye et al., 2017). These tools sound the argument that the urban form is tangible and quantifiable, and they can help designers to evaluate a given scheme. However, these tools are not developed for generating new designs, they work only with analysable layouts. Besides, when it comes to complex space like the industrial waterfront, involving irregular boundaries and vast vacant area, the time required to figure out an analysable layout will become much longer. Hence, if we can provide an approach to generate a preliminary layout automatically, using it as a reference for designers to further develop their ideas or conduct morphological analysis, the efficiency of redesigning the industrial waterfront can be improved.

### 1.3. OBJECTIVES

Correspondingly, we propose to use the machine learning method to study whether the computer can identify the characteristics of the industrial waterfront and surrounding environment, and independently generate new designs based on them.

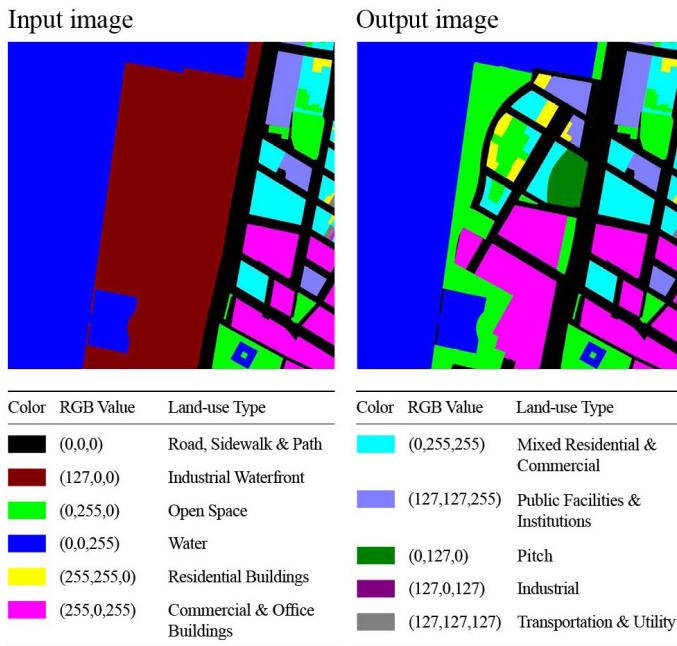


Figure 1. An example of the input map and ideal output map

At present, many studies have shown that machine learning can be applied in the fields of architecture and urban planning, such as building footprint generation (Shen et al., 2020), urban block reprogramming (Yu, 2020), and urban functional layout (Lin et al., 2020). Specifically, we use Generative Adversarial Network (GAN) to conduct this research. Such an image-to-image method is suitable for identifying the morphological features around the industrial land on the masterplan level, including



street network, building footprint, and land-use types.

To make sure the discussion can be focused on one urban planning context, this study takes New York City, where the waterfront space is diverse enough, well developed, and following the same urban network system, as the research object. The goal of this study is to use the trained neural network to transform a given map of industrial waterfront into a map with street networks, building footprints, and land-use types (Figure 1). At the same time, this model should be able to give feedback to the change of the morphological elements and generate different and reasonable waterfront urban space.

## 2. Methodology

### 2.1. DATASET CONSTRUCTION

The data set of this study is constructed through two steps. Firstly, the land-use shapefile in New York was obtained from the NYC PLUTO Data platform and then imported into the Mapbox platform, which contains the street network data. According to historical satellite maps, we also created the industrial land tileset on the platform. Together, these three tilesets constituted the data source of this paper.

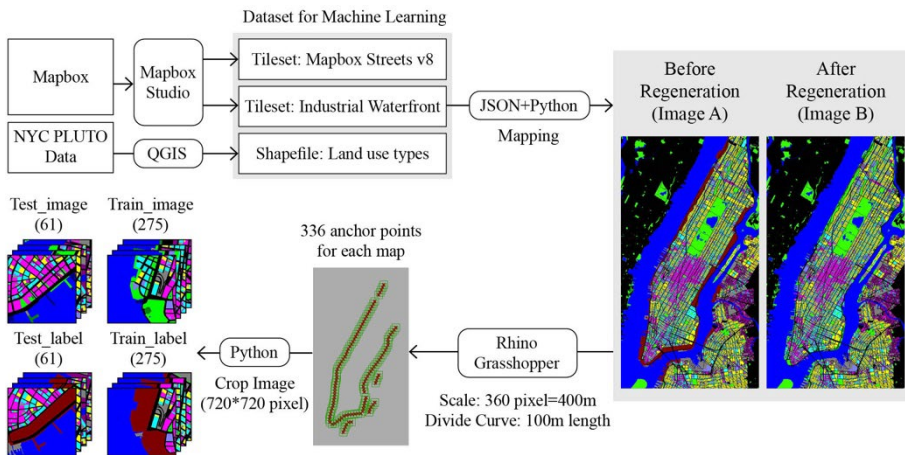


Figure 2. Dataset construction process

The second step is data visualization, according to the land-use classification in Figure 1, the tilesets were modified as a series of image pairs (A and B). In these pairs, different land-use types were labelled with different colours. In image A, the industrial lands before regeneration were covered by a distinguishable colour (RGB:127,0,0). Namely, image A presented the land condition of the waterfront before regeneration, and image B illustrated the morphological features after the regeneration of the same site. According to the walking radius of 5 minutes, we took 800m\*800m as the research scope of each waterfront space (that is, each image includes a circle of 400m radius), and cropped the images every 100m along the waterfront. The process ensures that the two adjacent maps are related, which is more suitable for analysing the continuous and

linear space like the waterfront. Respectively, 336 pictures were cropped from the above two maps and were divided into 61 test sets and 275 training sets (Figure 2). The 275 training sets were used in the training process and the 61 test sets were prepared for verifying the results of morphological features generation in section 3.3.

## 2.2. NETWORK TRAINING

Pix2PixHD (Wang et al., 2018), a Generative Adversarial Network (GAN) model, was adopted to train the above dataset. In the neutral network, the generator generates pseudo images similar to the images in the training set, to increase the probability that the discriminator misjudges the accuracy of the images. Figure 3 shows the change of the training loss value of the generator and the discriminator in 200 training epochs. The entanglement and non-convergence of the two values indicate that the generator and the discriminator were competing with each other and evolved simultaneously in the training process. Additionally, the comparison of the results of the training set (Figure 4) indicates that the generated map is almost consistent with the actual map morphological map. Therefore, the network training was completed.

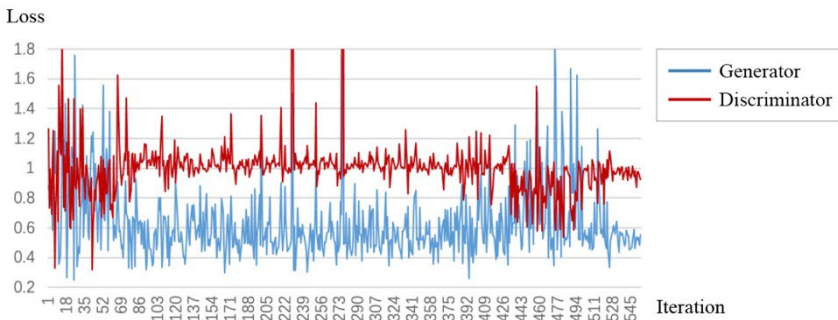


Figure 3. The loss values of the generator and discriminator in the training process

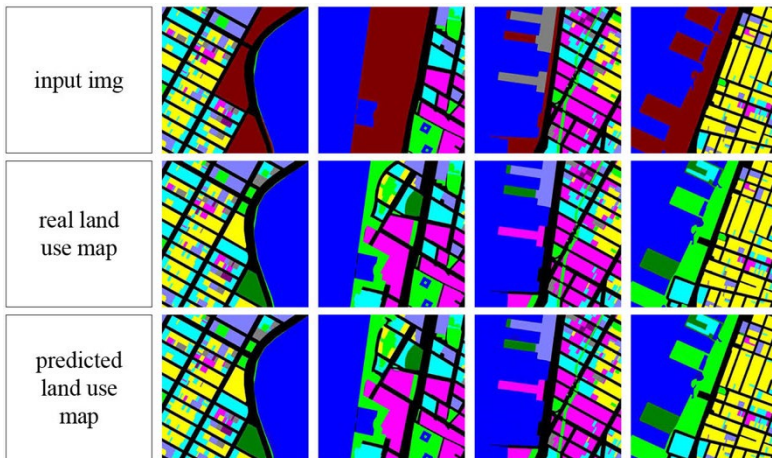


Figure 4. Selected results of the input image, real land-use map, and the predicted land-use map in the training dataset

### 3. Experiments and Results

#### 3.1. RESULT ENHANCEMENT

To test the feasibility of the machine learning model, two experiments were conducted. The feasibility refers to whether the model can identify the morphological elements of the industrial waterfront (such as water boundary, street network, building footprint, and land-use types) and can generate different maps according to the changes of elements. Since the waterfront space is linear whereas the data used by the generator is square, the morphological information at the edge of each input image may not be sufficient enough to support a reasonable calculation. To solve this problem, we inputted all the images that intersect the input area and then overlapped the corresponding output images with 50% transparency (Figure 5). The pixel coordinates of each image had been labelled in its file name, which was used for repositioning the corresponding output image in the overlapping process. In addition, the sequence to overlap the output image was determined by the Euclidean distance between each intersected image and the original area, with the original image at the top, followed by the picture with the closest Euclidean distance, and so on. The overlapped image was then cropped by the original area, resulting in a new map including more comprehensive information. Based on this method, two corresponding experiments will be discussed below.

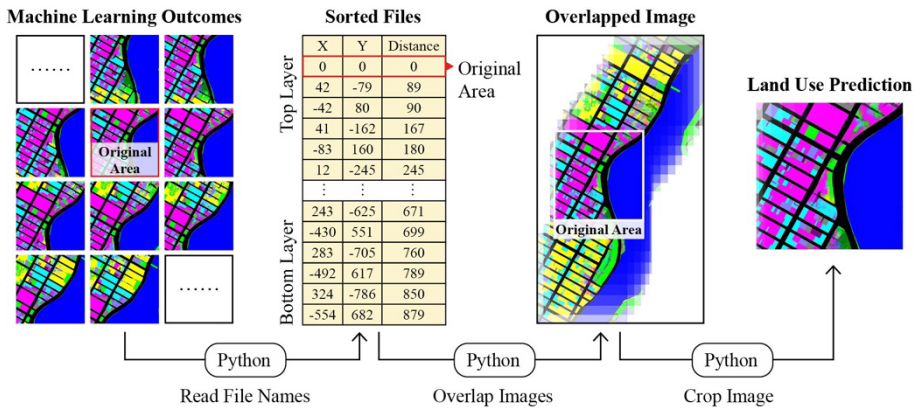


Figure 5. The process to overlap the output images

#### 3.2. MORPHOLOGICAL FEATURES IDENTIFICATION

The first experiment was conducted to test whether the neural network could identify urban morphological features. We selected three samples from the training set as the control group and modified one of their morphological features (water boundary, land-use composition, or street networks) to build three experimental groups. The results generated by the machine learning model are shown below (Figure 6). Generally speaking, instead of rigidly copying the urban form in original data, the model can give reasonable feedback to the changes of these three elements.

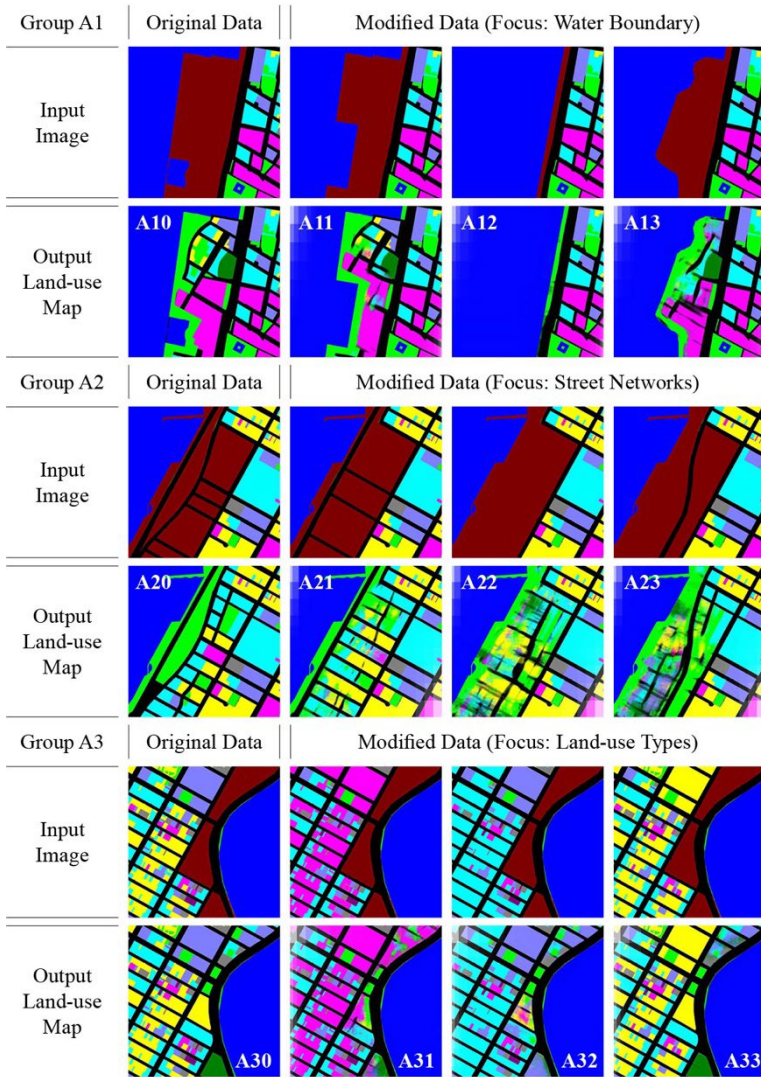


Figure 6. The outcomes of the morphological features identification experiment

Specifically, the input images in experimental group A1 illustrate the same place with modified water boundaries. When the shape was slightly changed (A11), the output image keeps the building footprint and land-use composition in the original scheme. Moreover, when the shape changes dramatically, the generated map could adjust the morphological elements to match the new boundary. For example, in A12, the vacant parcel is narrow, thus the model only generates open space without setting any buildings. In A13, the output building footprints and street networks can still adapt to the irregularity of water boundary, and at the same time ensure the connectivity of waterfront open space.

The input map of group A2 changed on the street network. It turns out that the trained model can further subdivide the remaining vacant land (A21-A23) and apply various land-use types in different conditions. Group A3 shows that the model can also generate the land-use types in the vacant land according to the modifications in the existing land-use distribution. Particularly, in A32, the model has recognized a street in the middle of the map which is mainly "commercial and office buildings (pink)" and extended the feature of this street to the waterfront space.

### 3.3. MORPHOLOGICAL FEATURES GENERATION

In the second experiment, we chose images that were not included in the training dataset to verify whether the machine learning model worked in new scenarios. These images are included in the "Test label" dataset mentioned above (Figure 2) and have not been used in the training process.

Since the number of samples used in this study is limited, we majorly selected areas with irregular shapes or unsatisfactory planning results to build up the data in "Test label". On one hand, such a classification helps to ensure the consistency of the training data. On the other hand, we can use the output of the "Test label" data to test whether the machine learning model can provide a better solution than the original scheme. The selected data contains four groups of continuous spaces with different characteristics. The maps in Group B1 are similar to the training set, which can be used to verify the accuracy of the generated images. The rest three groups of spaces have the following characteristics respectively:

- Group B2 is an undeveloped area that contains an irregular water boundary.
- Group B3 is characterized by a large vacant area.
- Group B4 is the location of the United Nations Headquarters. The building is characterised by its interference with the urban network. Therefore, this data will help to imagine what the building should look like if it echoes the surrounding environment.

It turns out that the generated results are similar to the real status and the model could even generate some details that were missing in the original data (Figure 7). The output of group B1 identified the shape characteristics of the highway, thus a large area of open space was generated on the vacant land. Besides, it recognized that the land was wide enough to contain pitches, which was consistent with the real design. In group B2, although the site lacks sufficient land-use reference, the model can still generate street networks highly similar to the original scheme by aligning the road to the water boundary. In group B3, because of the large vacant area, the generated map cannot effectively define the distribution of land-use types. However, the generated map clearly identifies an east-west road, although the road does not exist in the original land-use map, it can be found that this road is the main street in the community from the satellite map. That is to say, the machine learning model generated a road that was missed in the original data (Figure 7b). In group B4, the new map effectively continues the standard urban grid of New York City and maintains the combination of the open space and institutional land-use type.

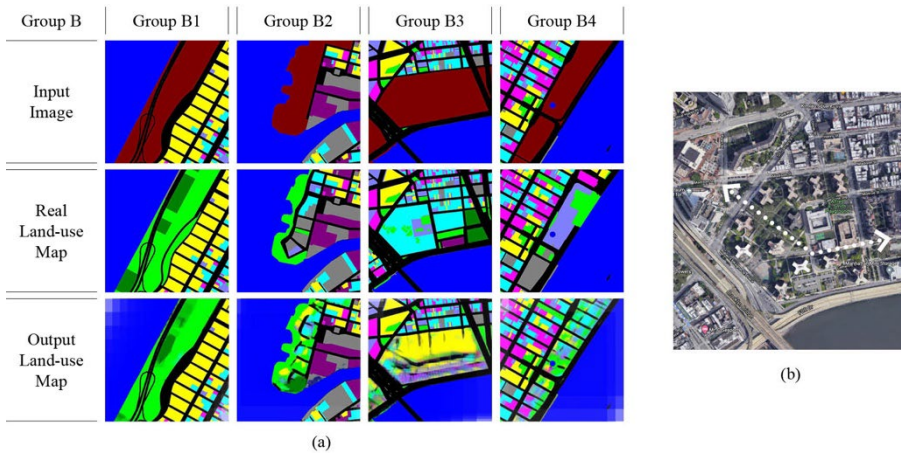


Figure 7. (a) The outcomes of the morphological features generation experiment; (b) The missed street in the original data of group B3 (Source: Google Map)

## 4. Discussion

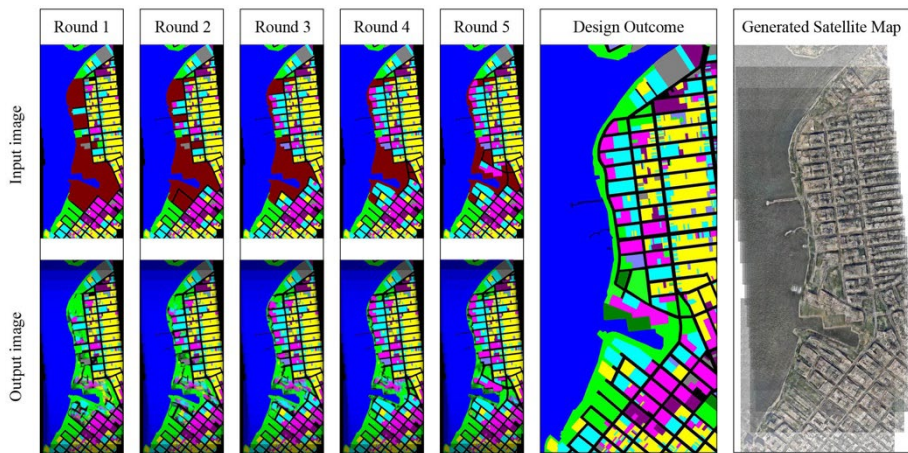
### 4.1. RESULTS ANALYSIS

It can be concluded from the above experiments that the machine learning model is more effective for identifying the water boundary and street networks. When it comes to land-use types, although the generated dominant land-use is usually correct, the land-use boundary presented is rather vague, especially when the input data contains large vacant land. These phenomena can also reflect three features of the training model:

- **Boundary feature:** Open space will generally be set at the vacant land adjacent to the water, defining a public space between buildings and the waterfront. If there is a street or highway adjacent to the water, dividing the vacant industrial land from the waterfront, the buildings will be directly built without in-between open space. Such a design logic is clear, thus it can be observed in nearly all regenerated maps.
- **Adjacent feature:** The generated land-use type will generally continue the land use function of adjacent parcels. If the adjacent parcels are residential use, the vacant land will usually be redeveloped as open spaces and pitches instead of copying the residential type. However, if the land-use composition of adjacent parcels is complex, the generation of land-use will become uncertain.
- **Grid feature:** The computer-generated street strictly follows the east-west network system of New York City. If the waterfront space locates in the south or north of the site, the street tends to be perpendicular to the water boundary instead of continuing the south-north grid. This feature is related to the irregular water boundary in the south of Manhattan and the lack of waterfront in the north, which also results in the distortion of some output maps facing south or north.

#### 4.2. FEATURES UTILIZATION

The above features were then utilized to generate an entire region of vacant industrial land rather than a single cropped square. Specifically, we selected the Green Point-Williamsburg Waterfront as our research object. The region was first divided into 13 maps according to the crop rule in Chapter 2 and imported to the trained model to generate the corresponding new maps. The new maps were then overlaid into a continuous image. By observing the image, we artificially mark the parts with clear boundaries and add them into the input maps, while leaving the blurry areas unchanged with the input colour. The new map becomes the input image for the next iteration (Figure 8 left). In each iteration, the machine learning model can generate some morphological elements related to the built space based on the above-mentioned features, then the newly generated elements can be used as new references to support the next iteration. Therefore, after five rounds of iteration, we obtained a waterfront map with clear street networks and regularly distributed land-use types (Figure 8 middle). We further implemented another image-to-image machine learning model (Zheng, 2018, Huang and Zheng, 2018) to generate the satellite image from the colour-coded map, the output image (Figure 8 right) shows a better visualization of the generated result.



*Figure 8. Left: The machine-learning-assisted process to generate the Green Point-Williamsburg Waterfront; Right: The satellite map was generated from the land-use map by our machine learning model*

#### 4.3. CONCLUSION

To conclude, the research explores the potentiality of GAN to learn and apply the morphological rules of industrial waterfront. The proposed methods for data preparation and manipulation illustrate the feasibility of understanding a linear, vast, and irregular urban space by converting a given area into a series of simple and analysable images. Therefore, these methods can help to regenerate other unique urban spaces that cannot be simply understood by classical morphological theories. By doing so, it is promising to simplify the procedure of early-stage planning to reuse the

deteriorated urban space and move forward the process to make a city more diverse, inclusive, and sustainable, which matches the Sustainability Development Goals 11 of the United Nations, Sustainable Cities and Communities.

## References

- Carpenter, A., & Lozano, R. (2020). Proposing a Framework for Anchoring Sustainability Relationships Between Ports and Cities. In A. Carpenter & R. Lozano (Eds.), *European Port Cities in Transition: Moving Towards More Sustainable Sea Transport Hubs* (pp. 37-51). Springer International Publishing. [https://doi.org/10.1007/978-3-030-36464-9\\_3](https://doi.org/10.1007/978-3-030-36464-9_3).
- Conzen, M. R. G. (1960). *Alnwick, Northumberland: A Study in Town-Plan Analysis*. George Philip.
- Hillier, B., Penn, A., Hanson, J., Grajewski, T., & Xu, J. (1993). Natural Movement: Or, Configuration and Attraction in Urban Pedestrian Movement. *Environment and Planning B: Planning and Design*, 20(1), 29-66. <https://doi.org/10.1068/b200029>.
- Huang, W., & Zheng, H. (2018). Understanding and Visualizing Generative Adversarial Networks in Architectural Drawings. In *23rd International Conference on Computer-Aided Architectural Design Research in Asia: Learning, Prototyping and Adapting, CAADRIA 2018* (pp. 156-165). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA).
- Lin, B., Jabi, W., & Diao, R. (2020). Urban Space Simulation Based on Wave Function Collapse and Convolutional Neural Network. In *11th Annual Symposium on Simulation for Architecture and Urban Design, SimAUD 2020* (Article 18). Society for Computer Simulation International.
- Sevtsuk, A., & Mekonnen, M. (2012). Urban network analysis. A new toolbox for ArcGIS. *Rev. Int. GÉomatique*, 22, 287-305. <https://doi.org/10.3166/RIG.22.287-305>.
- Shen, J., Liu, C., Ren, Y., & Zheng, H. (2020). Machine Learning Assisted Urban Filling. In *25th International Conference on Computer-Aided Architectural Design Research in Asia: Intelligent and Informed, CAADRIA 2020* (pp. 679-688). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA).
- Wang, T., Liu, M., Zhu, J., Tao, A., Kautz, J., & Catanzaro, B. (2018). High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 8798-8807). Institute of Electrical and Electronics Engineers (IEEE).
- Yang, C. X. (2014). The Integrative Organization among Urban Waterfront Elements. *Advanced Materials Research*, 869-870, pp. 104-109. <https://doi.org/10.4028/www.scientific.net/AMR.869-870.104>.
- Ye, Y., Yeh, A., Zhuang, Y., van Nes, A., & Liu, J. (2017). "Form Syntax" as a contribution to geodesign: A morphological tool for urbanity-making in urban design. *URBAN DESIGN International*, 22(1), 73-90. <https://doi.org/10.1057/s41289-016-0035-3>.
- Yu, D. (2020). Reprogramming Urban Block by Machine Creativity - How to use neural networks as generative tools to design space. In *38th Education and Research in Computer Aided Architectural Design in Europe: Anthropologic: Architecture and Fabrication in the cognitive age, eCAADe2020* (pp. 249-258). Education and Research in Computer Aided Architectural Design in Europe (eCAADe).
- Zheng, H. (2018). Drawing with Bots: Human-computer Collaborative Drawing Experiments. In *23rd International Conference on Computer-Aided Architectural Design Research in Asia: Learning, Prototyping and Adapting, CAADRIA 2018* (pp. 156-165). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA).



# DATA-DRIVEN RESEARCH ON STREET SPACE QUALITIES AND VITALITY USING GIS MAPPING AND MACHINE LEARNING

*A case study of Ma On Shan, Hong Kong*

XINYU LIU<sup>1</sup> and JEROEN VAN AMEIJDE<sup>2</sup>

<sup>1,2</sup>*The Chinese University of Hong Kong*

<sup>1</sup>*xinyu.liu@link.cuhk.edu.hk*

<sup>2</sup>*jeroen.vanameijde@cuhk.edu.hk, 0000-0002-3635-3305*

**Abstract.** In a post-carbon framework, data-driven methods can be used to assess the environmental quality and sustainability of urban streetscape. Streets are an important part of people's daily lives and provide places for social interaction. Therefore, in this study, the relationship between street quality and street vibrancy is measured using the new town of Ma On Shan, Hong Kong as a study area. Firstly, machine learning was used to identify the physical features of streets through geographic information collection and streetscape image acquisition. Secondly, previous measurement algorithms are combined to calculate the greenness, walkability, safety, imageability, enclosure, and complexity of streets. Thirdly, secondary calculations and visualisations were carried out on a Geographic Information System (GIS) platform to observe the current distribution of street qualities. Finally, the relationship between street quality and vibrancy was analysed using SPSS statistical analysis software. The results show that walkability has a positive effect on street vitality, whereas safety and complexity have a negative effect on street vitality. This study demonstrates how the quantitative assessment of urban street environments can be used as a reference for building a green, low-carbon, healthy, and walkable city.

**Keywords.** Street Quality; Geographic Information Systems; Machine Learning; Image Segmentation; SDG 11.

## 1. Introduction

As the global pandemic and extreme weather continue to have an impact on people's lives, there is an increased awareness of the impact of industrialisation and urbanisation on climate change. Rapid urban development is often accompanied by a lack of investment in urban planning and infrastructure, resulting in traffic congestion, substandard urban spaces and the deterioration of urban vitality and the natural environment (Walton et al., 2020). Moreover, rapid industrialization and urbanization have a direct impact on climate change through increased carbon emissions. In the post-

carbon era, the construction of low-carbon smart cities is being promoted as an effective way to improve the quality of urban spaces and implement sustainable development. An effective way to create low-carbon, sustainable and healthy cities is by making cities more walkable (Shao et al., 2021) and improving their walkability (Project Drawdown, 2020). To create walkable cities, the design of urban streets should be considered, and the needs of pedestrians and cyclists should be given priority over vehicle users (Appleyard, 1980).

City streets are an important part of people's daily life and commute. They also serve as open spaces for citizens to communicate with one another. As such, streets are the most heavily used public spaces and, therefore, are a reflection of the city's overall image. This paper focuses on the quantitative analysis of urban streets and determines the relationship between street quality and vibrancy for the promotion of low-carbon and healthy cities.

## **2. Research on Street Space Qualities**

### **2.1. HISTORICAL RESEARCH ON THE QUALITY OF STREET SPACES**

Before the dominance of vehicles, the street was an important place for urban life and interaction, and the urban skeleton bridging the different functional spaces in the city. With the advent of the industrial age, mobility in the city increased and the demand for transport grew. As a result, more and wider streets were constructed for vehicular traffic, and the needs of pedestrians were compromised (Marshall, 2005). Streets evolved into an urban space suitable for cars and, therefore, were no longer created at a human scale.

With the emergence of the garden city theory in the second half of the 19th century, hygiene, safety and green aspects of streets were promoted to increase the quality of urban spaces. Later, during the post-war reconstruction in the early 20th century, the functionalist school of thought (including Le Corbusier) placed emphasis on streets as places for vehicular transport rather than living spaces for pedestrian movement and social interactions (Ingersoll, 2019; Great Britain Department for Transport, 2007; Mehaffy et al., 2014). To counter the effects of the industrial movement, many scholars have advocated for the design of more humane public spaces, accommodating the social and cultural functions of the streets (Jacobs, 1961; Lynch, 1964; Whyte, 1980). In response to the Modernist and rationalist theories in urban planning, progressive cities have now adopted urban design principles focusing on walkability, connectivity, diversity and smart transport.

### **2.2. QUANTITATIVE EVALUATION OF STREET QUALITIES**

Traditional research mostly adopts questionnaire and field survey methods (commonly used in sociology and statistics) to collect information on the characteristics of spatial elements of streets as well as their service level (Ewing and Cervero, 2010; Gehl, 1994; Handy et al., 2002). In addition, Space Syntax theory (Hillier and Hanson, 1984) is being widely used for large-scale evaluation of urban street networks. However, these common methods of evaluation do not cover the full complexity of residents' spatial movement and the flexibility of streetlife. With the

development of geographic information and computer science, it has become possible to collect and analyse residents' behavioural data through information processing software and equipment (Clarke, 1986). Moreover, the accuracy of these methods has greatly improved over the years. Geographical Information System (GIS) data allows for a detailed analysis of the material dimensions and the environment of streets on a larger scale. Several scholars have collected users' walking tracks in urban streets and analysed the hot spots of urban activities using GIS platforms (Edwards, 2009; Sagl et al., 2012).

In addition, street view image data have been widely accepted as an effective means of measuring the built environment (Kelly et al., 2013; Li et al., 2015; Salamon et al., 2014; Shen et al., 2018). As street view images reflect the street space environment from eye-level perspective, online data acquisition can replace other data collection methods that are limited by weather, time and place. Street view data can be analysed using machine learning techniques using SegNet (Badrinarayanan et al., 2017; Verma et al., 2018), semantic annotation tools that identify green, sky, buildings, roads and vehicles. These urban elements can then be represented and assessed through interactive visual analysis systems.

The accuracy of using Street View Image (SVI) data as research data has been discussed by several researchers, acknowledging the limitations around the interval distance and time of day variations associated with the data source. The methodology is deemed to produce reliable results when the study area is sufficiently large (Kim, 2021). Previous studies used 50-meter intervals for street greenery research (Lu et al., 2019; Ye et al., 2019), while others use sampling points every 50 and 100 meters, combining images in east, south, west and north directions into panoramic images for analysis (Yang et al., 2019; Law et al., 2020).

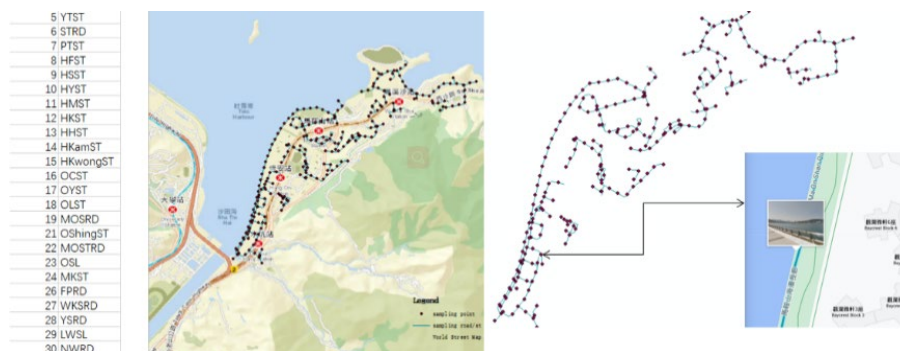


Figure 1. Ma On Shan New Town - Street segmentation and sample points

### 3. Methodology

#### 3.1. CASE STUDY AND DATA SOURCES

The study revolved around the streets of Ma On Shan (MOS) New Town in Hong Kong which is a 'third-generation New Town' (Figure 1). It used street data from the HK GEODATA STORE obtained in October 2020 and SVI data from Google Maps in November 2020. Data collection and filtering was carried out in two steps:

- SVI: In Google Street View, the coordinates of the collection sampling points were determined at intervals of 50 meters, and all four SVIs in the south, east, north and west orientations were obtained using python programming language procedures in PyCharm. At total, 1920 images were collected and at each collection point, four images were synthesized into one panoramic image, resulting in 480 SVIs.
- HK GEODATA STORE: Excluding expressways, trunk roads, data for a total of 30 local streets was collected. The data included street names, geographical coordinates, and lengths.

Limited by data sources, research techniques and variables, this study still has some shortcomings. Dependant on the SVI use of perspective, the pixel percentage for different objects in the scene is distorted by the distance to the camera. However, as this problem occurs simultaneously across all image data, the limitation is a constant factor across larger geographical areas and rendering the detection of regional differences significant.

### 3.2. OVERALL FRAMEWORK

The research framework was based on three levels (shown in Figure 2). Firstly, at the level of machine learning (ML), two ML algorithms, the Pyramid Scene Parsing Network (PSPNet) algorithm (Zhao et al., 2017) and Mask Region-based Convolutional Neural Networks (MaskR-CNN) (He et al., 2017), were applied to the segmentation of SVI. These are the most well-known and established image segmentation algorithms.

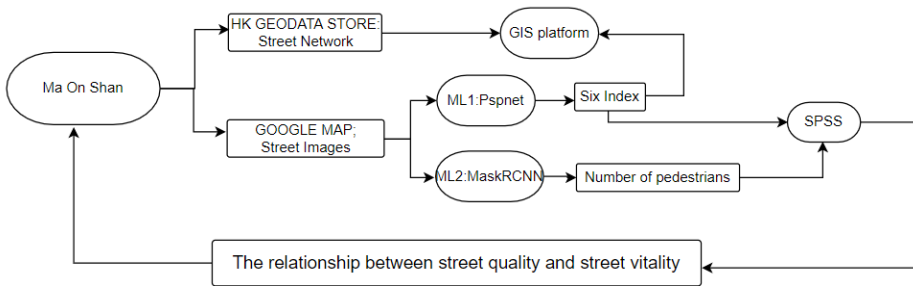


Figure 2. Research workflow

Secondly, calculation and visualisation were conducted on a GIS platform at the level of geographic information data. In this study it was used to visualize the calculation results and the distribution of the street vitality indicators. Finally, the relationship between street quality and street vibrancy was analysed using SPSS software.

### 3.3. PARAMETER SYNTHESIS OF STREET QUALITIES

PSPNet has been implemented into a module that can be trained to classify pixels in a raster to identify and mark 150 types of elements in a picture. The physical features and environmental elements used in this study are tree, grass, sidewalk, fence, road, person, signboard, streetlights and building (Figure 3). The percentage of these

elements in each image can be calculated. Moreover, MaskR-CNN efficiently detects objects in an image while simultaneously generating a high-quality segmentation mask for each instance which can identify and count the number of pedestrians and cars in the SVI (Figure 4).

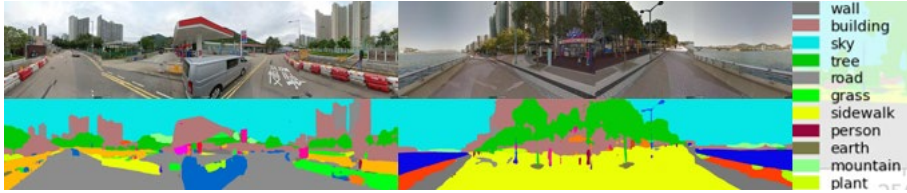


Figure 3. PSPNet -Image segmentation



Figure 4. MaskR-CNN - Instance segmentation

This study focused on the quantification of the characteristics of the physical space of the street. It is based on previous research by Ma et al. (2021) and Qiu et al. (2021), who have developed a methodology for the calculation of the indicators of greenness, openness, closure, walkability, safety and imageability based on the physical characteristics and environmental elements identified by SVI. By combining previous methodologies, our workflow involved extracting the spatial elements of the street and synthesizes the six parameters of street quality: Q1-Greenness, Q2-Walkability, Q3-Safety, Q4-Imageability, Q5-Enclosure and Q6-Complexity (Figure 5). The street vitality index is calculated by counting the number of pedestrians in the streets.

Index	Definition	Equations	Physical Features
<b>Greenness</b>	Urban Green pace that are an essential element in streetscape, including forests, greenbelts, and lawns. (Ma et al., 2021)	$=PSVtree+PSVgrass$	
<b>Walkability</b>	The psychological impact of the surrounding visual elements on the walking experience, such as the sense of comfort and pleasure for walking. (Ma et al., 2021)	$= (PSVsidewalk+PSVfence)/PSVroad$	
<b>Safety</b>	An individual's experience of the risk of becoming a victim of crime and disturbance of public order. (Jansson, 2019)	$=PSVperson+PSVsignboard+PSVstreetlight+PSVfence$	
<b>Imageability</b>	The quality of a place that makes it distinct, recognizable and memorable. (Ewing & Handy, 2009)	$=PSVbuilding+PSVsignboard$	
<b>Enclosure</b>	The degree to which streets and other public spaces are visually defined by buildings, walls, trees and other vertical elements. (Ewing & Handy, 2009)	$= (PSVbuilding+PSVtree)/(PSVroad+PSVsidewalk+PSVgrass)$	
<b>Complexity</b>	The visual richness of a place, which depends on the variety of the numbers and types of buildings, ornamentation, landscape elements, street furniture, signage and human activity. (Ewing & Handy, 2009b)	$= (PSVtree+PSVperson+PSVsignboard+PSVstreetlight+PSVfurniture)/(PSVroad+PSVbuilding)$	

Figure 5. Parametric synthesis- Street quality indicators

## 4. Data Interpretation

### 4.1. DISTRIBUTION CHARACTERISTICS OF STREET QUALITIES

This analysis quantifies the spatial aggregation or dispersion of street qualities. To visualise street quality and explore street dynamics, a 100\*100m rectangular grid was created in ArcGIS using the Fishnet tool. The average street quality values calculated in the previous step were cross-referenced to the street geographic information and linked to the attribute table for each cell. This was done to accurately map the data to the urban scale and reduce the bias caused by different street lengths.

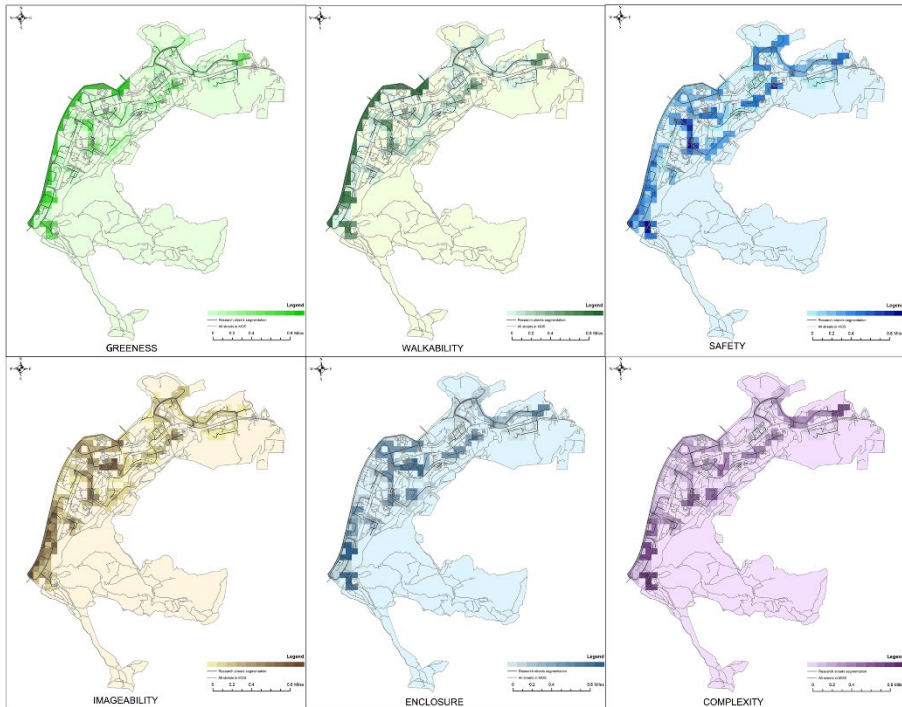


Figure 6. Visualization of street quality distribution

As shown in Figure 6, the ID1-MOS Promenade has the highest average indices of Q1-Greeness and Q2-walkability, followed by ID3-Hang Hong St, which is higher near the Playground and Sports Ground. Meanwhile, the ID1-MOS Promenade (southern section) and ID2-Ning Tai Rd have the highest Q3-Safety index, ID24-Ma Kam St has the least sense of security. Q4 identification is the highest in ID1-MOS Promenade (south section), and the lowest in ID24-Ma Kam St, ID27- Kai Sha Rd (west section) and 30-Nin Wah Rd. ID16-On Chun St has the highest Q5-Enclosure and ID28-Yiu Sha Rd the lowest. ID4-Hang Tai Rd has the highest Q6-Complexity. ID1-MOS Promenade (northern section) has the lowest complexity.

#### 4.2. CORRELATION AND REGRESSION ANALYSIS OF STREET QUALITIES AND STREET VITALITY

To verify relationships between street qualities and street vitality, Correlation and Linear Regression Analysis was used to study the influence of independent variables Q1-Q6 (quantitative) on dependent variable P=Pedestrian (quantitative), including whether there is an influence relationship, the direction of influence and the degree of influence. The Durbin-Watson test, which is an autocorrelation test, was first used to construct the linear regression model. The D-W value is near 2, indicating that the model has no autocorrelation and there is no correlation between sample data. Therefore, the model is performing well.

Table 1. Linear regression model summary

Model Summary						
R	R <sup>2</sup>	Adj. R <sup>2</sup>	RMSE	D-W	AIC	BIC
0.776	0.602	0.498	9.443	1.926	233.856	243.664

Subsequently, the model fitting situation was analysed by R square value, and the VIF value was analysed to determine whether the model has collinearity problem. As can be seen from the above table, the model formula is:  $P = 14.287 + 12.123 * Q1 + 106.657 * Q2 - 194.836 * Q3 + 12.599 * Q4 - 10.970 * Q5 - 34.682 * Q6$

Results are listed in Table 2. The means Q1-Q6 show a difference in P of 60.2%. At the same time, the model passed the F-test ( $F=5.798$ ,  $P = 0.001 < 0.05$ ), and also indicates that Q1- Q6 have at least one effect on P relations. On the other hand, the test of the model's multi-collinearity shows that all VIF values in the model are less than 5, meaning there is no collinearity problem.

Table 2. Summary of Regression Analysis Outcomes

Parameter Estimates (Summary)			
	Coefficients	95% CI	VIF
Constant	14.287 (1.509)	-4.275 ~ 32.850	-
Q1GREENESS	12.123 (0.621)	-26.110 ~ 50.356	1.572
Q2WALKABILITY	106.657** (4.372)	58.845 ~ 154.470	1.253
Q3SAFETY	-194.836* (-2.204)	-368.071 ~ -21.601	1.087
Q4IMAGEABILITY	12.599 (0.462)	-40.877 ~ 66.076	1.558
Q5ENCLOSURE	-10.970 (-1.301)	-27.497 ~ 5.558	2.362
Q6COMPLEXITY	-34.682* (-2.248)	-64.914 ~ -4.450	2.52
F Value	F (6,23)=5.798,p=0.001		

\*  $p < 0.05$  \*\*  $p < 0.01$  t statistics in parentheses

The regression coefficient of Q1 is 12.123( $t=0.621$ ,  $P=0.540>$ ; 0.05); Q4 is 12.599( $t=0.462$ ,  $P=0.649>$ ; 0.05); Q5 is -10.970( $t=-1.301$ ,  $P=0.206>$ ; 0.05). It means that Q1, Q4 and Q5 have no influence on P (Figure 7).

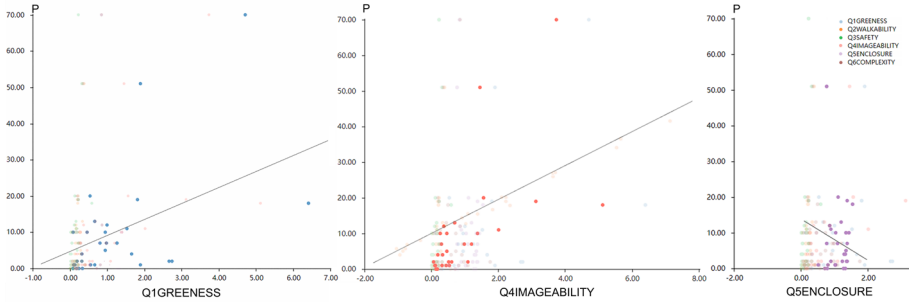


Figure 7. Scatter plot of Q1, Q4, Q5 and number of pedestrians

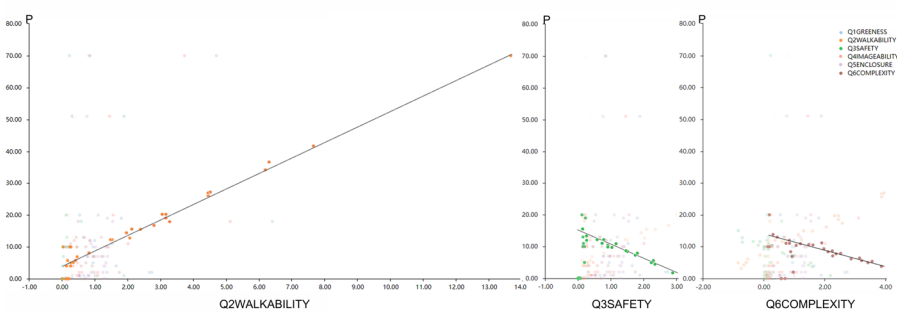


Figure 8. Scatter plot of Q2, Q3, Q6 and number of pedestrians

The regression coefficient of Q2 is 106.657( $t=4.372$ ,  $P=0.000<0.01$ ), indicating that Q2 has a significant positive influence on P. The regression coefficient of Q3 was -194.836( $t=-2.204$ ,  $P=0.038<0.05$ ), Q6 was -34.682( $t=-2.248$ ,  $P=0.034<0.05$ ) means that Q3 and Q6 have a significant negative influence on P (Figure 8).

## 5. Conclusions and Future Directions

This paper explores a mechanism to analyse the objective physical characteristics and environmental elements of streets to quantify the impact of street quality on the street vibrancy in a quantitative data-driven framework. The results show that walkability, safety, and complexity directly affect the number of pedestrians on streets. The study illustrates how the improvement of street vitality can be guided through quantitative parameters contained in an urban street quality index, and provides a reference for the construction of green, low-carbon, healthy and walkable cities.

The analytical framework presented in this paper has the potential to quantitatively measure street quality in more areas in the future, as data for streetscapes is open-source and relatively easy to collect. At the same time, it should not be overlooked that the



accuracy of this model depends heavily on the timeliness and comprehensiveness of the data sources as well as the subjective nature of the residents. With advancements in image recognition technology, it is hoped that more precise data-driven and quantitative analyses of urban streets, including analysis of different pedestrian activities, can be carried out in the near future.

## References

- Appleyard, D. (1980). Livable Streets: Protected Neighborhoods? *The Annals of the American Academy of Political and Social Science*, 451, 106–117.  
<http://www.jstor.org/stable/1043165>
- Badrinarayanan, V., Kendall, A. and Cipolla, R. (2017) Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(12), 2481-2495.
- Clarke, K.C. (1986). Advances in Geographic Information Systems. *Computers, Environment and Urban Systems*, Vol. 10: 175–184.
- Edwards, D. (2009). Using GPS to Track Tourists Spatial Behaviour in Urban Destinations. *SSRN Electronic Journal*. <http://dx.doi.org/10.2139/ssrn.1905286>
- Ewing, R., & Cervero, R. (2010). Travel and the built environment: A meta-analysis. *Journal of the American planning association*, 76(3), 265-294.
- Gehl, J. (1994). *Public spaces & public life in Perth: report for the Government of Western Australia and the City of Perth*. Perth, W.A: Dept. of Planning and Urban Development
- Great Britain Department for Transport (2007). *Manual for streets*. London: Thomas Telford Pub.
- Handy, S. L., Boarnet, M. G., Ewing, R., & Killingsworth, R. E. (2002). How the built environment affects physical activity: views from urban planning. *American journal of preventive medicine*, 23(2), 64-73.
- He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 2961-2969).
- Hillier, B., & Hanson, J. (1984). *The Social Logic of Space*. Cambridge, New York: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511597237>
- Ingersoll, R. (2019). *World architecture: A cross-cultural history*. New York. Oxford University Press
- Jacobs, J. (1961). *The Death and Life of Great American Cities*. New York: Vintage Books.
- Kelly, C. M., Wilson, J. S., Baker, E. A., Miller, D. K., & Schootman, M. (2013). Using Google Street View to audit the built environment: inter-rater reliability results. *Annals of behavioral medicine: a publication of the Society of Behavioral Medicine*, 45 Suppl 1(Suppl 1), S108–S112. <https://doi.org/10.1007/s12160-012-9419-9>.
- Kim, J.H., Lee, S., Hipp, J.R., & Ki, D. (2021). Decoding urban landscapes: Google street view and measurement sensitivity. *Comput. Environ. Urban Syst.*, 88, 101626.
- Law, S., Seresinha, C. I., Shen, Y., & Gutierrez-Roig, M. (2020). Street-frontage-net: Urban image classification using deep convolutional neural networks. *International Journal of Geographical Information Science*, 34(4), 681–707.
- Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q. and Zhang, W. (2015), Assessing street-level Lu, Y., Yang, Y., Sun, G., & Gou, Z. (2019). Associations between overhead-view and eye-level urban greenness and cycling behaviors. *Cities*, 88, 10-18.  
<https://doi.org/10.1016/j.cities.2019.01.003>
- Lynch, K. (1964). *The image of the city* (Publications of the Joint Center for Urban Studies). Cambridge [Mass.]: M.I.T. Press.

- Ma, Xiangyuan & Wu, Chao & Xi, Yuliang & Yang, Renfei & Chen, Zhang. (2021). Measuring human perceptions of streetscapes to better inform urban renewal: A perspective of scene semantic parsing. *Cities*. 10.1016/j.cities.2020.103086.
- Marshall, S. (2005). *Streets and patterns*. London: Spoon Press/Taylor & Francis Group.
- Mehaffy, M. W., Porta, S., & Romice, O. (2015). The “neighborhood unit” on trial: A case study in the impacts of urban morphology. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability*, 8(2), 199-217.
- Project Drawdown. (2020). *Walkable Cities. Climate Solutions at Work*. Project Drawdown.
- Qiu, W., Li, W., Liu, X., & Huang, X. (2021). Subjectively Measured Streetscape Perceptions to Inform Urban Design Strategies for Shanghai. *ISPRS International Journal of Geo-Information*, 10(8), 493. MDPI AG. Retrieved from <http://dx.doi.org/10.3390/ijgi10080493>
- Sagl, G., Loidl, M., & Beinat, E. (2012). A Visual Analytics Approach for Extracting Spatio-Temporal Urban Mobility Information from Mobile Network Traffic. *ISPRS International Journal of Geo-Information*, 1(3), 256–271. MDPI AG. Retrieved from <http://dx.doi.org/10.3390/ijgi1030256>
- Salamon, J., Jacoby, C., & Bello, J. P. (2014). A dataset and taxonomy for urban sound research. In *MM 2014 - Proceedings of the 2014 ACM Conference on Multimedia* (pp. 1041-1044). Association for Computing Machinery. <https://doi.org/10.1145/2647868.2655045>
- Shao, J., Yang, M., Liu, G., Li, Y., Luo, D., Tan, Y., Zhang, Y., et al. (2021). Urban Sub-Center Design Framework Based on the Walkability Evaluation Method: Taking Coomera Town Sub-Center as an Example. *Sustainability*, 13(11), 6259. MDPI AG. Retrieved from <http://dx.doi.org/10.3390/su13116259>
- Shen, Q., Zeng, W., Ye, Y., Arisona, S. M., Schubiger, S., Burkhard, R., & Qu, H. (2018). StreetVizor: Visual Exploration of Human-Scale Urban Forms Based on Street Views. *IEEE transactions on visualization and computer graphics*, 24(1), 1004–1013. <https://doi.org/10.1109/TVCG.2017.2744159>
- Verma, D., Jana, A., & Ramamritham, K. (2018). Quantifying Urban Surroundings Using Deep Learning Techniques: A New Proposal. *Urban Science*, 2(3), 78. MDPI AG. Retrieved from <http://dx.doi.org/10.3390/urbansci2030078>
- Walton, M., Murray, E., & Christian, M. D. (2020). Mental health care for medical staff and affiliated healthcare workers during the COVID-19 pandemic. *European heart journal. Acute cardiovascular care*, 9(3), 241–247. <https://doi.org/10.1177/2048872620922795>.
- Whyte, W. (1980). *The social life of small urban spaces*. Washington, D.C.: Conservation Foundation.
- Yang, Y., He, D., Gou, Z., Wang, R., Liu, Y., & Lu, Y. (2019). Association between street greenery and walking behavior in older adults in Hong Kong. *Sustainable Cities and Society*, 51, 101747.
- Ye, Y., Xie, H., Fang, J., Jiang, H., & Wang, D. (2019). Daily Accessed Street Greenery and Housing Price: Measuring Economic Performance of Human-Scale Streetscapes via New Urban Data. *Sustainability*, 11(6), 1741. MDPI AG. Retrieved from <http://dx.doi.org/10.3390/su11061741>
- Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). Pyramid Scene Parsing Network. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 6230-6239.

# EVALUATING THE ACCESSIBILITY OF AMENITIES: TOWARDS WALKABLE NEIGHBORHOODS

*An Integrated Method for Testing Alternatives in a Generative Urban Design Process*

SIFAN CHENG<sup>1</sup>, KA SHUT LEUNG<sup>2</sup> and JEROEN VAN AMEIJDE<sup>3</sup>

<sup>1,2,3</sup>*School of Architecture, the Chinese University of Hong Kong.*

<sup>1</sup>*SifanCHENG@link.cuhk.edu.hk, 0000-0002-2303-7151*

<sup>2</sup>*kashutleung@cuhk.edu.hk, 0000-0003-2936-0344*

<sup>3</sup>*jeroen.vanameijde@cuhk.edu.hk, 0000-0002-3635-3305*

**Abstract.** Studies have shown that walkable communities reduce traffic-related pollution and the risk of chronic illnesses, promote economic growth and prosperity, and stimulate community participation and the growth of social capital. To assess the walkability of urban areas, various methodologies have been developed around shortest-distance calculations between various points of interest (POIs), yet their outcomes do not guide potential urban design improvements. The absence of appropriate measurements and procedures that may give quantitative and actionable feedback to support design decision-making is one of the primary issues in building walkable neighborhoods. The work presented in this paper revolves around a new workflow, that employed Urbano, a mobility simulation and assessment tool, and integrated it within a generative design process to allowing for the quantitative evaluation on amenity accessibility for several alternative design scenarios for a case study site in Mong Kok, Hong Kong. The results show how this data-driven urban design process benefits from generative techniques to produce solutions with improved contextual connectivity, energy-efficient urban form, and good quality public spaces that contribute to the walkability of neighbourhoods.

**Keywords.** Generative Urban Design; Walkability; Urbano; SDG 3; SDG 11.

## 1. Introduction

There has been a recent surge in the amount of research providing methodologies and tools for assessing the walkability of urban environments. This practical endeavour reflects the increased interest in pedestrian-oriented urban design and planning among academics, practitioners, and public policymakers. The idea of walkability has evolved into a critical theoretical and operational set of planning goals. Studies have shown that walkable communities considerably reduce traffic-related pollution and the risk of chronic illnesses (Frank et al. 2006; Lee and Buchner 2008), promote economic growth

and prosperity (Claris and Scopelliti 2016), as well as encourage growth in social capital and political involvement (Leyden, 2003). One of the most significant qualities of a walkable city is access to varied amenities within walking distance, which has been linked to socioeconomic growth and quality of life.

From being considered merely a question of accessibility relating to pedestrians, the concept of walkability has evolved into a richer, more nuanced, and intrinsically multidimensional description of the relationships and interactions between people, urban space, and its social practices of use through debate and practice. As a result, there has been a wide range of attempts to operationalize the idea of walkability, including the development of formal measurements, audit processes, assessment methodologies, and tools (Blecic et al., 2017; Frank et al., 2006; Saelens et al., 2003).

The fundamental drivers and motivation behind promoting walking activities are generally accepted to be proximity to and availability of facilities (Clark et al., n.d.). To assess cities' walkability, efforts have been undertaken to rank them based on a shortest-distance study between various Points of Interest (POIs). These walkability ratings, known colloquially as *Walkscores* (Brewster et al., 2009; ESRI, 2019; Walkscore, 2019), are calculated on a one-to-one scale and include criteria such as access to services and facilities such as grocery stores, doctors, parks, schools, hospitals, and public transit. The Walkscore metric can be computed using the Urban Modeling Interface (Reinhart et al. 2013). The key problem with this tool is providing proper inputs (e.g., street networks, buildings, and the locations of facilities) that the user must manually enter. Furthermore, a single value like the Walkscore does not give enough knowledge to assist with design. For example, the services and amenities that are important to have walking access to differ depending on the demographic group, hence it is necessary to be able to analyse demographic-specific factors.

Rapid technological advances are leading to the integration of information and communication technologies (ICT) within the built environment, resulting in the emergence of a new urban science around sensing, data gathering, and analysis of urban processes. Meanwhile, the evolution of generative urban design techniques that utilise new computational tools and large datasets can assist in optimising urban form (Alonso et al., 2019; Noyman et al., 2019), energy efficiency (Nagy et al., 2018), and walkability of neighbourhoods (Rakha & Reinhart, 2012; Sonta & Jain, 2019). Urbano is an urban analytics toolset that enables an automated workflow for analysing the accessibility of facilities, as well as how the distribution of amenities affects people walking in a neighbourhood (Yang, Samaranayake, et al., 2020). It uses multi-source open data to import contextual GIS, OpenStreetMap, and Google Places data into Grasshopper to create an urban mobility model (Dogan et al., 2018). This new modelling framework incorporates travel behaviour modelling as well as the innovative urban design metrics of Streetscore, Amenityscore, and an upgraded Walkscore.

Using building-based local population statistics and the distribution of facilities, Urbano simulated the geographical distribution of potential pedestrian trips in the analysis findings. Previous research has examined the validity of Urbano's three walkability indices in the context of the United States (Yang, Dogan, et al., 2020; Yang, Samaranayake, et al., 2020). It examines the mobility models of two New York neighbourhoods and deduces the morphological concerns behind the variance in

walkability. Furthermore, Yang & Samaranayake, et al. used Urbano to investigate the spatial and dynamic variation of amenity demand in neighbourhoods over the duration of a day. It used Google Place API (popular time) as location-based social network (LBSN) data, and the results included suggestions for potential design interventions. However, few studies used Urbano in a generative design process to model the walkability outcomes of various design choices.

Design methods that integrate generative and parametric characteristics allow designers to explore a broader range of potential solutions. Computational optimisation is increasingly being used in this area to tackle difficult design challenges ranging from energy consumption to structural performance (Brown et al., 2020; Wortmann & Nannicini, 2017). The absence of appropriate measurements and procedures that give quantitative and actionable feedback to support design decision-making is one of the primary difficulties in urban design practice, leading to a lack of evidence-based decision making in urban planning. The hypothesis of this study is that Urbano, as a simulation tool, may be integrated withing generative urban design processes, to explore the impact of several alternative design scenarios on amenity accessibility.

## **2. Case Study Context: Mong Kok, Hong Kong**

High-density development is promoted by smart growth, new urbanism, and other comparable planning programs (Jenks, 2019). High density enables access to a variety of social, health, recreational, and other services, facilitating most daily trips to be accomplished on walking (Foord, 2010; Raman, 2010). Using Mong Kok as a case study site, this research project explored the novel possibility of using Urbano as a walkability evaluation toolkit to test generative design outcomes.

Previous studies usually included a traffic road network to analyse pedestrian activity and street connections. A study of the Belfast footpath network, however, found that evaluating pedestrian network systems predicted physical activity behaviours more effectively than road network data (Ellis et al., 2016). In the case of Hong Kong, the variations are significant when employing a pedestrian network rather than a road network. Pedestrian networks in densely populated cities, particularly hilly ones like Hong Kong, feature several non-vehicular shortcuts between two-dimensional (2D) road network linkages, some of which create a three-dimensional road network (3D) (Sun et al., 2021). Recent research used a geographic information system to create a three-dimensional network for Hong Kong territory, with the dataset displaying distinct levels for pedestrian activities such as footbridges and subterranean linkages (Zhao et al., 2021). This newly available information offers more accurate urban data which can be used effectively in walkability simulations due to the 3D capabilities of Urbano.

Mong Kok is situated in the centre of Hong Kong's historic urban core, and is characterised by dense urban form and complex 3D pedestrian networks typical of Hong Kong's old urban districts. The area features tall commercial, residential and mixed-use blocks and is a popular shopping destination within the urban region (Villani & Talamini, 2019; Xue et al., 2001). Next to the Mong Kok East MTR station lies an urban regeneration site which will be the largest urban block to be redeveloped in a decade. With busy markets, stores, restaurants, residential units, and offices within the area, it is an appropriate case study site for testing new site modelling techniques that

avoid oversimplification in spatial, economic and cultural terms.

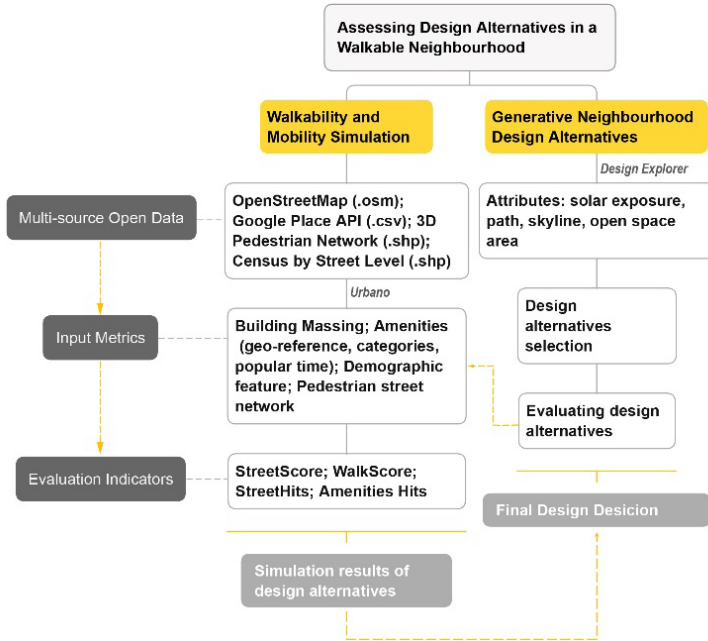


Figure 1. Analytics Framework

### 3. Methodology and Results

Our original walkability analysis study is based on the Urbano mobility model. Initially, we collected various information from multiple open data sources and entered them into Urbano to analyse the present walkability situation in Mong Kok. Following automatic data processing, this model produced two major indications, Walkscore and Streetscore, as well as two visual outcomes, Amenity Hits, and Street Hits. Subsequently, we employed a generative toolset, Design Explorer, to create design scenarios at the Mong Kok East site that included different density, typology, circulation routes, skyline, and open space characteristics. Finally, we integrated these various design alternatives into the Urbano walkability model, to evaluate the impact of the designs on the surrounding neighbourhood, shown as Amenity Hits and Street Hits, using a quantitative comparison analysis in Walkscore and Streetscore to select the most optimised solutions.

#### 3.1. DATA COLLECTION AND PRE-PROCESSING

The initial stage in our workflow was to create a mobility model using urban data such as streets, buildings, services, and population densities. Figure 2 illustrates the information flow for our data collection process. Urbano combined several input streams including buildings, streets, amenities, and associated metadata. It automatically performed post-processing and cleaning up of the street networks before

connecting the buildings to the nearest street segment to construct a mobility network. A topological network was then created with street segments as edges and intersections and endpoints as vertices, after the detection of crossings in streets and the segmentation of street curves. This graph was later utilized to obtain routes. A shortest-distance path matrix was produced for all of the graph's vertices as part of this procedure.

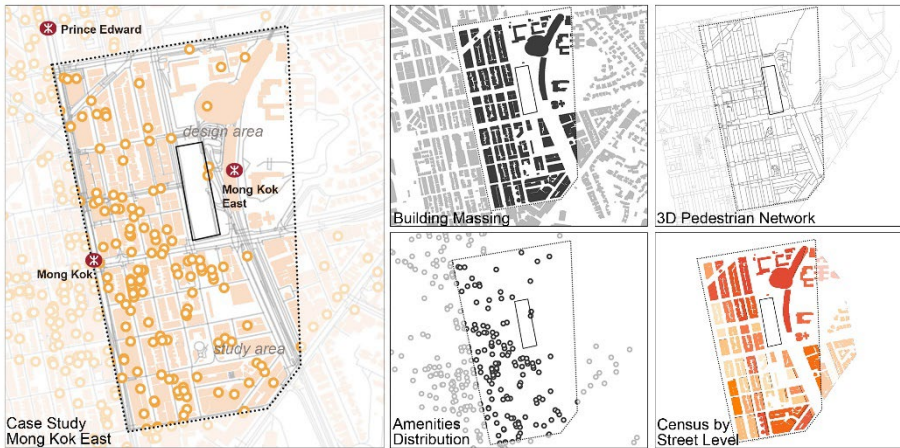


Figure 2. Data Input Process

### 3.2. MOBILITY MODEL OF THE CURRENT CONDITION

#### 3.2.1. Auto-generative Mobility Model in Urbano

To evaluate the potential pedestrian volumes of streets and amenities through a mobility simulation, two evaluation indicators: Walkscore and Streetscore were included. Firstly, the generated advanced Walkscore in Urbano is based on the well-known Walkscore statistic. A higher Walk Score, a destination-based rating method, is associated with a shorter distance and/or a broader range of facilities (Walk Score, 2019). The Walk Score method employs a distance decay function and categorizes facilities as categorical variables (grocery stores, restaurants, hardware stores, etc.). A decay function (Equation 1) is used to rate trips based on distance, with the score decreasing to 0 as trip duration increases.

$$Decay(x) = -17.15x^5 + 89.45x^4 - 126.37x^3 + 4.639x^2 + 7.58x + 99.5$$

The Walkscore in Urbano is based on a shortest-distance study between various places of interest (POIs) using OpenStreetMap and the Google Place API, of which categories are consistent with Walkscore amenities. According to Table 1, various categories of facilities were assigned varying weights. A cumulative Walkscore for each building and a neighbourhood average could be produced at the neighbourhood level, to show the general accessibility of various amenities.

Secondly, our workflow assessed streets using a simple counter called Streetscore, which predicts how many people utilize specific street sections. This function could be

used to evaluate the liveliness of a street within the overall network, or at a certain time.

The walkability evaluation model also includes two parts of visual results. Amenity Hits represent pedestrian hits that an amenity received during a simulated round. Through the input of the popular times of each amenity through Google Maps API and census metrics by building clusters level, this measure is primarily generated by Urbano's shortest-path routing methodology, determining the population numbers arriving from all sources. Street Hits measures how many people utilize a certain street section on all visits, which may be used to simulate the Street Score outcomes.

Table 1. Study Type of Amenities and Weights based on Walkscore

OSM Node Tag	Google Place API Type	Walkscore Category	Amenity Weights
amenity=post_office	type=post_office	errands	1
amenity=hospital	type=hospital		
amenity=bank	type=bank		
amenity=library	type=library		
amenity=theatre	type=movie_theatre	culture_entertainment	1
amenity=cinema	type=cinema		
amenity=art_centre	type=art_centre	grocery	3
shop=department_store	type=department_store		
shop=supermarket	type=grocery_or_supermarket		
shop=convenience	type=convenience_store		
shop=hardware	type=hardware_store		
shop=home_goods_store	type=home_goods_store		
amenity=pharmacy	type=pharmacy		
amenity=restaurant	type=restaurant	dining and drinking	0.75, 0.45, 0.25, 0.25, 0.225, 0.225, 0.225, 0.225, 0.2, 0.2
amenity=fast_food	type=fast_food		
amenity=café	type=café		
amenity=bar or amenity=pub	type=bar or type=night_pub	shopping	0.5, 0.45, 0.4, 0.35, 0.3
shop=clothes	type=clothing_store		
shop=books	type=book_store		
shop=beauty or shop=hairdresser	type=beauty or type=haircare		
shop=shoes	type=shoes		
shop=jewelry_store	type=jewelry_store		
amenity=school	type=school		
leisure=park	type=park	school	1
		park	1

### 3.2.2. Mobility Simulation Results in Mong Kok

Through the creation of a mobility simulation in the current condition in Mong Kok, our study provided evaluation results of this overall area and generated quantitative data on the popularity, capacity, and accessibility of varied amenities and streets. To explore the different amenities' needs patterns across varied periods of the day, which could reflect different user groups' need of amenities, we further generated three Amenity Hits result at 8 am (weekday), 1 pm (weekend), 7 pm (weekend). As shown in Figure 3, the amenities are concentrated in the southwest area of Mong Kok. With a dense cluster of restaurants, mall stores and convenience stores, this area achieved higher Amenity Hits and building-level Walkscores. Due to its the dense and continuous pedestrian network, the streets in the south section show vibrancy, indicated by higher Street Hits values compared with the northern part. Comparing the results from three different time periods shows that most of the retail and leisure amenities, such as restaurants, mall stores, markets, and convenience stores received more hits at 1pm (weekend). Service amenities such as banks, government offices received more hits at 8am.



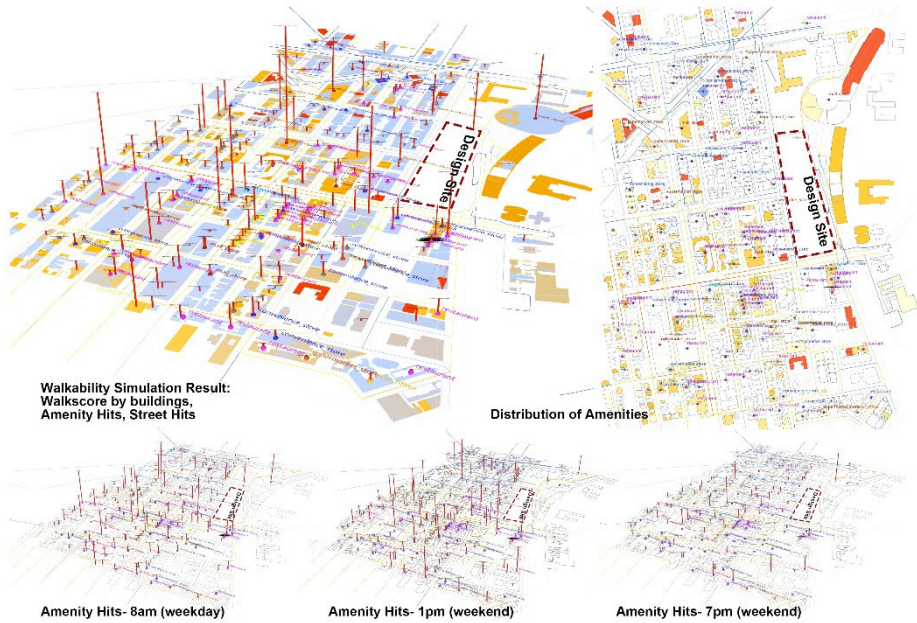


Figure 3. Mobility Simulation Results in Mong Kok

The simulation results indicate that although the northwest area, near the regeneration site, is close to the Mongkok East MTR station, this area received much less Amenity Hits and a lower building-level Walkscore, due to a lack of potential attraction and destination points, and poor connections in the pedestrian walking system. The planned introduction of new functions and routes in the regeneration area may have the potential to link the fragmented pedestrian paths and provide multiple new destinations for potential users.

### 3.3. GENERATIVE DESIGN TESTING

Based on the evaluation of the current condition, our study integrated walkability analysis within a generative urban design workflow to provide varied design scenarios for the regeneration site, create new linkages and enhance the walkability the area. The generative workflow in Grasshopper was set up as follows:

- Three input parameters were used, based on the results of the mobility model: Access/attractor points, Paths, and Skyline. Access/attractor points represent potential destinations outside the site that need to be linked through the new urban cluster. Paths represents the division pattern within the site and Skyline is a measure that ensures that views from both sides of the site are connected.
- The size of the urban cluster units was entered, resulting in a generative design result that divides the site into different modules according to the size of the unit.
- Three basic typologies of urban modules were defined, describing different forms

of tower-podium and open space massing.

- Using Interactive Evolutionary Algorithms (IEAs) in Biomorpher and Colibri, different combinations of urban modules and clusters were iterated based on the above inputs.
- Metrics for different design scenarios were exported to .json and .csv, including the Skyline, Path, Attractor, Attractor location, Total area, Open space area values, and the numbers of Typologies 1,2,3.

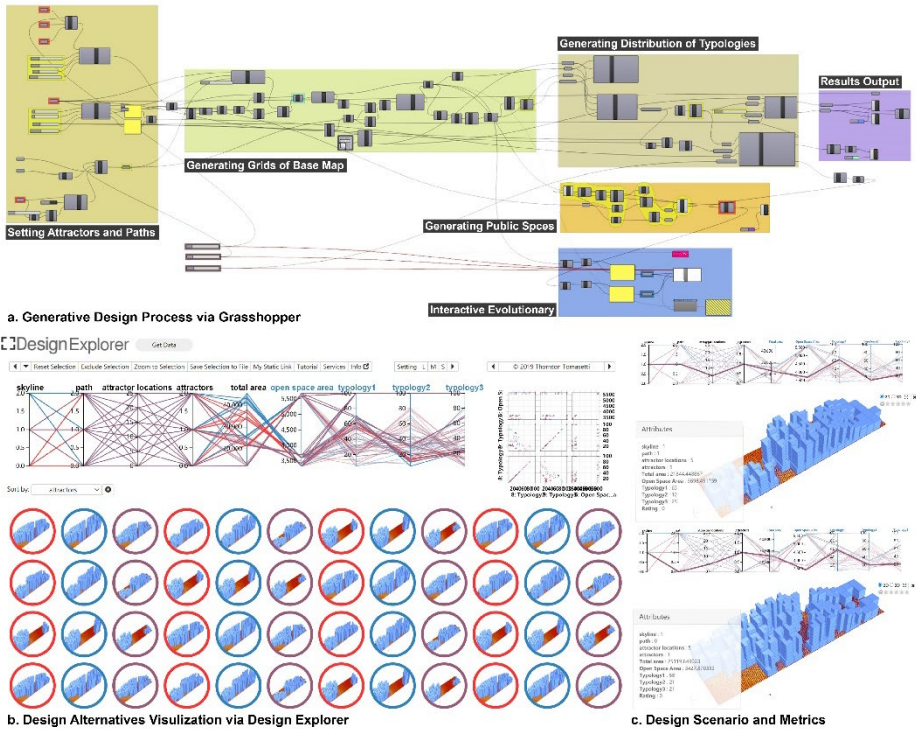


Figure 4. Generative Design Intervention and design examples in Design Explorer

### 3.4. COMPARATIVE EVALUATION OF DESIGN ALTERNATIVES

In its final stage, our study involved evaluating the selected design scenarios against the original mobility model of the case study site, to compare the walkability improvements within the design site and their impact on the surrounding area. Three chosen design scenarios all optimise the connectivity of the amenities and street network on both sides of the site, to improve the current disconnect of the Mong Kok East MTR station zone from the Mong Kok area. Scenario (a) is defined as a mixed-

used cluster, with different types of ground floor amenities and a high building density and less open space (3548 m<sup>2</sup>). Scenario (b) provides a different linkage pattern, where the northeast and southwest parts are linked to optimizing leisure needs. Scenario (b) also incorporates additional amenities like restaurants and bars. Scenario (c) uses the same linkage pattern as Scenario (a), but provides more public open space (4630 m<sup>2</sup>).

In this study, Scenario (a) (b) (c) were entered into the original mobility model in Mong Kok via Urbano. Figure 5 documents some of the results, showing that Scenario (a) has a higher average Amenity Hit (614.74), as it provides multiple potential destinations. In terms of Walkscore, the within-site Walkscore varied little across Scenario (a) (b) (c), however, when comparing the overall study area, Scenario (a) has the highest score (39.01), indicating that it is the most optimised for the accessibility of amenities over the wider Mong Kok district

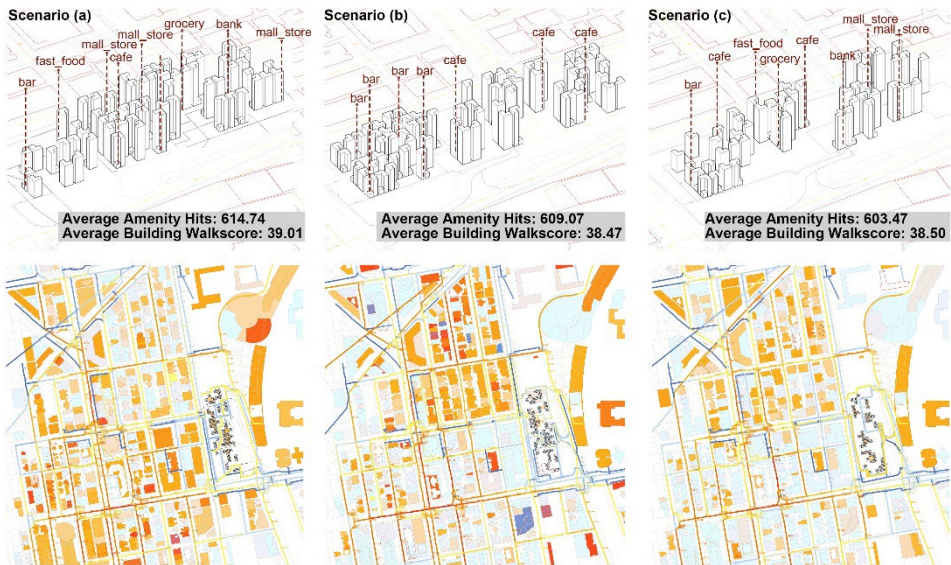


Figure 5. Comparative Result of Design Scenarios

#### 4. Discussion and Conclusion

Understanding the consequences of urban design decisions on urban inhabitants' mobility during the very early phases of urban design processes can assist in the development of better proposals that incorporate walkability comprehensively. The work presented in this paper revolves around a new workflow, that connects the simulation of a mobility model to the optimisation of generative urban design solutions, allowing for the quantitative evaluation of the optimisation results. This data-driven design method, which includes generative and parametric characteristics, allows designers to explore a broader range of potential solutions. In addition, as the impact of design outcomes on the surrounding environment can be reflected in real-time in the simulation results, designers can benefit from the involvement of a wider range of data

and metrics from these areas in the design solution-seeking process during a multi-stakeholder engagement process.

The simulation model of walkability conditions is a complicated assignment that can develop in line with the development of new forms of urban studies and a better understanding of urban districts as complex systems. With the ongoing efforts to improve urban data quality and accessibility, future studies could apply the mobility model in a wider context to test its validity. Moreover, although this work integrates factors beyond amenities in the process of generative design, the mobility model was simplified to convey pedestrians through the shortest path, and future research could combine with street visual quality elements to optimise the multi-indicators of the mobility model for assessing walkability. Our future work will focus on the expansion of the new methodology towards more comprehensive generative urban design applications, aimed at promoting human-centric design strategies for inclusive, supportive and sustainable urban environments.

## References

- Blecic, I., Cecchini, A., & Trunfio, G. A. (2017). Computer-aided drafting of urban designs for walkability. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10407 LNCS, 695–709.
- Brown, N. C., Jusiega, V., & Mueller, C. T. (2020). Implementing data-driven parametric building design with a flexible toolbox. *Automation in Construction*, 118.
- Dogan, T., Samaranyake, S., & Saraf, N. (2018). Urbano: A new tool to pro-mote mobility-aware urban design, active transportation modeling and access analysis for amenities and public transport. *Simulation Series*, 50(7), 275–282.
- Frank, L. D., Sallis, J. F., Conway, T. L., Chapman, J. E., Saelens, B. E., & Bachman, W. (2006). Many pathways from land use to health: Associations between neighborhood walkability and active transportation, body mass index, and air quality. *Journal of the American Planning Association*, 72(1), 75–87.
- Leyden, K. M. (2003). Social Capital and the Built Environment: The Importance of Walkable Neighborhoods. *American Journal of Public Health* (Vol. 93, Issue 9).
- Raman, S. (2010). Designing a liveable compact city physical form of city and social life in urban neighbourhoods. *Built Environment*, 36(1), 63–80.
- Saelens, B. E., Sallis, J. F., & Frank, L. D. (2003). Environmental correlates of walking and cycling: Findings from the transportation, urban design, and planning literatures. *Annals of Behavioral Medicine*, 25(2), 80–9. Lawrence Erlbaum Associates Inc.
- Sun, G., Webster, C., & Zhang, X. (2021). Connecting the city: A three-dimensional pedestrian network of Hong Kong. *Environment and Planning B: Urban Analytics and City Science*, 48(1), 60–75.
- Villani, C., & Talamini, G. (2019). Patterns of Stationary Activities in the Elevated Pedestrian Networks of High-Density Asian Cities: The case of Mong Kok, Hong Kong. *Environment-Behaviour Proceedings Journal*, 4(12), 321.
- Yang, Y., Samaranyake, S., & Dogan, T. (2020). Using Open Data to Derive Local Amenity Demand Patterns for Walkability Simulations and Amenity Utilization Analysis. *eCAADe* 37, 665–674. [https://doi.org/10.5151/proceedings-ecaadesigradi2019\\_627](https://doi.org/10.5151/proceedings-ecaadesigradi2019_627)
- Zhao, J., Sun, G., & Webster, C. (2021). Walkability scoring: Why and how does a three-dimensional pedestrian network matter? *Environment and Planning B: Urban Analytics and City Science*, 48(8), 2418–2435. <https://doi.org/10.1177/2399808320977871>

# ELEMENTAL MOTION IN SPATIAL INTERACTION (EMSD)

*A Framework for Understanding Space through Movement and Computer Vision*

MARGARET Z. ZHOU<sup>1</sup>, SHI YU CHEN<sup>2</sup> and JOSE L. GARCÍA DEL CASTILLO Y LÓPEZ<sup>3</sup>

<sup>1,2,3</sup> *Harvard University.*

<sup>1</sup>*Imargaretzhzhou@gmail.com, 0000-0003-1187-9025*

<sup>2</sup>*jessica.shiyu@gmail.com, 0000-0002-2826-3318*

<sup>3</sup>*personal@garciadelcastillo.es, 0000-0001-6117-1602*

**Abstract.** Spatial analysis and evaluation are becoming increasingly common as new technologies enable users, designers, and researchers to study spatial motion patterns without relying on manual notations for observations. While ideas related to motion and space have been studied in other fields such as industrial engineering, choreography, and computer science, the understanding of efficiency and quality in architectural spaces through motion has not been widely explored. This research applies techniques in computer vision to analyse human body motion in architectural spaces as a measure of experience and engagement. A taxonomy framework is proposed to categorize human motion components relevant to spatial interactions, for analysis through computer vision. A technical case study developed upon a machine-learning-aided model is used to test a selection of the proposed framework within domestic kitchen environments. This contribution adds further perspective to wider research explorations in the quality, inclusivity, engagement, and efficiency of architectural spaces through computer-aided tools.

**Keywords.** Pose Estimation; Spatial Evaluation; Architectural Usability; Motion Studies; Computer Vision; SDG 3; SDG 9.

## 1. Introduction

To experience architecture, one cannot sufficiently understand its spatial ambitions through viewing - space needs to be navigated, explored, and oriented to obtain an informed experience. Spatial perception and evaluation are considered highly subjective and difficult to measure, and existing qualitative studies rely on surveys and interviews to gather data. Feedback is communicated through verbal and emotional experiences that summarize the user's journey through space. However, how a user feels after occupying space is driven by numerous external and internal factors that vary frequently.

Building interaction manifests in the form of human movements that occupy space. Each motion is the output of a subconscious, intuitive decision that the human brain makes in a given moment. Although subjective opinions cannot be measured, actions and trends could be identified and used to extrapolate future decisions. The study of human behaviour and cognition within Architecture is an emerging field. Within this, human circulatory and spatial patterns are particularly important for understanding how space is seen, understood, and used.

Usability studies in fields such as human-computer interaction and product design have enabled designers to evaluate their solutions before being used widely; however, in architecture, this is only possible with simulations and studies from built precedents. Understanding motion trends computationally could help expand research strategies for data simulations, how space shapes movement patterns, and the psychology behind wellbeing, emotion, and comfort in the built environment. Identifying outliers and repetitive or arduous motions, paired with additional cognitive studies, could help pinpoint areas of circulation pressure, energy waste, and ease of use.

The role of computation in addressing environmental pressures of the decade is crucial. Research continues to advance academic understanding of spatial and environmental phenomena in a carbon-neutral way, specifically through computational simulations that can maximize efficiency and usability before build. This research positions itself most predominantly with Goal 3 and 9 of the United Nations Sustainable Development Goals. It aims to further understandings of how the built environment affects users to ensure healthy lives and promote wellbeing for all through both physical ease of use and mental health awareness. This further contributes to building resilient infrastructure that can learn from human responses and interaction to increase inclusivity for all user and body types, and promotes innovation through intelligent, adaptable buildings that put the user first throughout all stages of design, before it is built.

This research explores how the analysis of body motion in architectural spaces could be enhanced using computer vision techniques as a potential tool to depict human experience. A taxonomy framework of motion elements is proposed, designed to visually measure an individual's interactions within built environments. Building upon existing research in motion detection, this codification was developed primarily through iterations of observations, and literature analyses. We propose four informational groups that represent facets of motion relative to architectural space: State, Type, Goal, and Texture. A technical case study developed using the machine learning model PoseNet is used to test the viability of a selected portion of the framework within domestic kitchen environments. (Papandreou et al., 2018). Results from the research could help to access further opportunities in studying usability through motion within fields of universal design, usability, and health and wellbeing.

### 1.1. BACKGROUND

Earliest forms of academic motion study research emerged with the invention of photographic still capture by early cameras that allowed for tracking from one still frame to the other. Studies on motion tracking became popular in the early 20th century, capturing movements that were previously unseen to the human eye through chronophotographic work with animal locomotion such as horses running and human

movement (Muybridge, 1901), or by isolating movement from its context and depicting tracking graphically in order to find new patterns (Marey, 1895).

Point tracking methods, while helpful, have limitations: simple 2D data points are digestible for humans identifying patterns yet indecipherable to the machine. Classifications help to group types of movements together for identification and data usage. Frank and Lillian Gilbreth introduced the idea of motion efficiency, where “motion waste” can be saved by reducing ineffective motions (1919). This led to the creation of Therbligs, a taxonomy of primary movements comprised within common handheld tasks, used to identify costly motion within manufacturing and assembly. Gilbreth’s students furthered these studies into the domestic home by measuring space used by families when performing daily tasks with equipment and furniture. (Callaghan and Palmer, 1944).

Movement classifications also exist in fields such as modern dance. Choreographer Rudolf von Laban created the Laban Movement Analysis, which captures the differences in qualities of movement that would be missed in purely formal analyses (Newlove, 1993; Prinsloo et al., 2019). It has been used in applications such as work efficiency and physical rehabilitation (Ajili et al., 2017).

Recent developments in motion capture and tracking bring technologies that greatly advance accessibility, cost, and accuracy. Optical motion capture systems utilize networks of cameras and sensor markers to process high-resolution movement for use in entertainment, such as OptiTrack, and ergonomic studies (Nagymáté and Kiss, 2018; Bortolini et al., 2018). Motion tracking through devices such as HaLO Indoor Positioning bring new ideas to denoting circulation in indoor environments; however, studies are limited to the number of wearable devices available to participants and focus on the complete journeys of individual users (Hu and Park, 2017).

Within computer vision research, Bobick (1997) and Moeslund et al. (2001) provide differing frameworks and definitions for identifying actions such as hitting a baseball or playing football through image-based representation using static and dynamic recognition (Mohamed, 2015). More recent taxonomies, as reviewed by Aggarwal and Ryoo (2011) take different approaches to classifying motion and build upon the idea of identifying categories of “gestures” that are broadly applicable to many different types of actions. There are also public datasets available for machine learning models, such as the KTH and Weizmann datasets, which provide videos for general ‘actions’ such as walking or running (Schuldt et al., 2004; Blank et al., 2005).

In architecture, motion extends beyond efficiency and tracking, embodying a greater purpose of fulfilment, enjoyment, and pleasure. Studies into measuring these qualities have been initiated by the idea of enabling the machine to perceive space as a human designer (Peng, 2017). Hirschberg et al. (2006) discuss the topic of motion within the study of architecture but focus on the opportunity for motion to create form, while Fox and Polancic (2012) explore gestural interaction with intelligent architectural environments. While existing explorations into motion identification address different aspects of the problem space through the lens of industrial engineering, computer science, and design, applications to usability in architectural spaces have not been widely explored.

## 2. Framework

The following framework combines applications of more recent computational techniques to ideas about quality, engagement, and usage of architectural spaces in order to identify and understand motion.

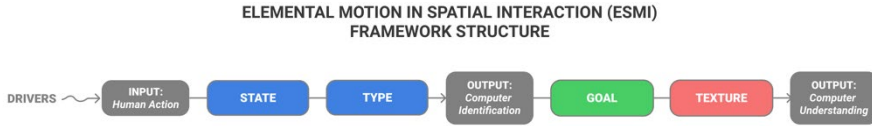


Figure 1. Framework Structure - Elemental Motion in Spatial Interaction ESMI

Elemental Motion in Spatial Interaction ESMI is a taxonomy of motion elements for identifying and understanding movements within built environments. Figure 1 describes the structure of the framework in which human action inputs could be identified and then further understood with computer vision. Drawing upon existing research, particularly the framework of Therbligs and those discussed by Aggarwal and Ryoo, it has been developed through observations, discussion, and literature analyses.

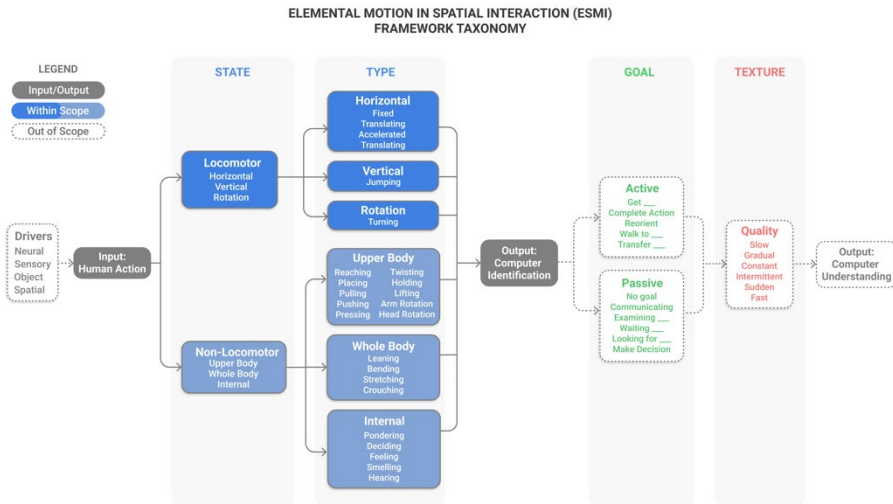


Figure 2. Full Proposed Taxonomy - Elemental Motion in Spatial Interaction ESMI

It proposes four taxonomy groups, *State* and *Type*, which allow an algorithm to firstly identify motion, and *Goal* and *Texture*, which add information for understanding the intent and quality of motion. Figure 2 shows the full framework including lists and hierarchy of identifiable data, and how they build on each other. Both outputs contribute layers of information to the quality of space; the former through data collection of positions of bodies within a given area and respective types of movement, the later through extrapolated data that attempts to understand emotional responses.



Drawing upon Laban's understanding of locomotor and non-locomotor movement, which distinguishes between movements that are done while the body is travelling from point to point versus those that use the axis of the body without travelling, the *State* of a motion is proposed here as one of the first identifications needed when observing motion occurrences (Newlove, 1993). Through determining whether the subject in focus is moving, the taxonomy can provide a general context of what space might be occupied next, as well as what types of movement could be completed. Area used by a person who is standing in the same place, is different from a person who is actively moving across a space.

*Type* is proposed as a breakdown of primary motion available to a person when using an architectural space. These elements can be layered on top of one another, but each has distinguishable conditions that do not overlap. Locomotor movement is split into horizontal, vertical, and rotational due to their different implications on subsequent, predicted, occupied space. For example, a 90-degree turn has a different prediction to a vertical jump. Non-locomotor is grouped into *Upper Body*, *Whole Body* and *Internal* due to their different implications on purposes. For instance, while a person can be reaching and leaning simultaneously, reaching is mutually exclusive to leaning, and one does not infer the other.

Through processing the *State* and *Type* of a motion, a motion identification could be inferred, noting how the space is being occupied over a period of time. Further contextual groups of information, *Goal* and *Quality*, could help an algorithm understand why it is being used and the impact on users. Both groups infer qualitative factors through physical measurements.

*Goal* is an identification that gives more insight into the purpose of the motion. Through observations, it was found that an observer would often make predictions to determine what movements might come next or the purpose of the motion initiation. This involves looking at previous trends and predicting future trends under a larger goal. These can be *active* or *passive*. *Active* goals engage with the surrounding spatial context through a physical process, such as walking to another point or transferring an object from point to point. *Passive* goals are typically engaged through an internal process that have physical motions as a by-product, such as tapping the floor while waiting for something. An algorithm can infer this through previous actions, direction, and trend assumptions.

The group of information *Texture* adds qualitative insight into the inference of how an individual might be feeling whilst moving. Observations showed evidence that the speed and style of motion can often relate to emotional states. Laban labels this as effort/dynamics – subtle movements that provide character. For example, a person who is angry is more likely to be tense and contracted. On the architectural scale, this translates into the speed of motion. A person walking at a faster pace is likely determined to achieve something and in a more focused mood. A person moving intermittently may be confused or struggling to find their next goal.

The output of this framework is, firstly, an identification of motion that can quantify how a space is being used over a period, and secondly, an understanding of response to the experience of a space. An example of how the framework could be applied to various scenarios is shown in Figure 3. This contributes to the field of occupancy and ethnographic studies, as well overall usability studies for how individuals respond to

specific architectural decisions within spatial environments, potentially revealing measurable trends through the data.

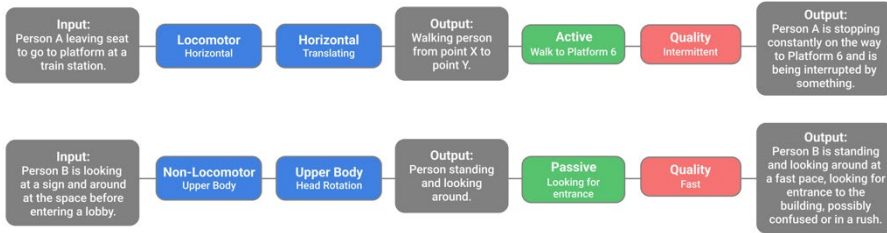


Figure 3. Framework applied to example scenarios

### 3. Case Study

#### 3.1. METHODOLOGY

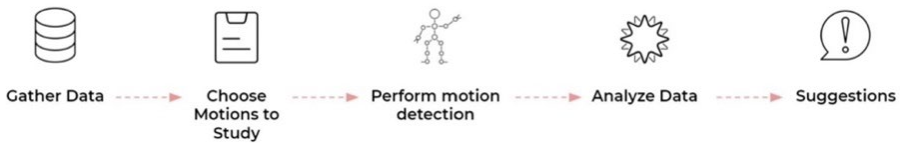


Figure 4. Methodology for Case Study

The goal of this case study is to develop an application of the proposed framework to initially assess its validity and applicability. Figure 4 shows the method process of the case study leading to a possible use case of the data which contributes to design suggestions. While all four taxonomy groups are discussed and proposed by the framework, this case study will focus primarily on State and Type, leading to Computer Identification. Inspired by the domestic motion studies done by Callaghan and Palmer (1944), household kitchens were selected as the target area of study. Two sets of video footage of kitchen usage by their inhabitants were collected. The first case, Kitchen 1, is a 25-minute recording of a 170 cm adult male making breakfast; the second case, Kitchen 2, is a 20-minute recording of a 163 cm adult female washing dishes.

Video footage was used for the purpose of obtaining a general understanding of the activities in the kitchen. In order to focus the scope of the case study, only EMSI *State* and *Type* groups were tested. To identify these, the footage was broken down into frames, a technique most closely related to the single-layered sequential approach (Aggarwal and Ryoo, 2011, p.14). Joint locations in each frame were detected using PoseNet (Papandreou et al., 2018), a machine learning model trained to estimate human skeleton poses. A custom algorithm was then developed to detect and measure the differences between frames to approximate a motion.

The algorithm detects for the following motions: translation, fixed, rotation, bending and reaching. Horizontal-Translation and Horizontal-Fixed *Types* are

considered as a binary set for calculation. To determine the fixed or translating state, the average location of points detected by the pose estimation framework to be above the hip are computed. The average position of pose estimation in each frame is subtracted from the average position of the frame prior to it. If the change is greater than a predetermined value, the frame is temporarily noted as a translational motion. To reduce detection errors, the final assignment of motions is determined based on the current and two previous detections of motion, which can be seen in Figure 5.

Rotation, Bending, Reaching are all identified with JavaScript code based on the distance and positioning changes between shoulder, nose, and wrist locations over the series of video frames, with each detection requiring a manually adjusted threshold factor to refine accuracy.

### 3.2. RESULTS

After detecting the motions in the footage, a file containing information of the motions detected, time of detection, and motion coordinates were generated for data analysis. Figure 5 shows an example of the data generated from the video motion analysis.

Type: Translation	Time	X,Y	Type: Rotation	Time	X, Y	Type: Reaching/Bending	Time	X, Y
translating	9.54	30.82, 292.85	turning	9.52	30.82, 292.85	reaching	9.52	43.14, 312.76
translating	9.56	30.82, 292.85	no turning	9.53	30.82, 292.85	reaching	9.53	43.14, 312.76
fixed	9.58	30.95, 292.21	no turning	9.54	30.82, 292.85	reaching	9.54	43.14, 312.76
fixed	9.59	30.80, 291.67	no turning	9.56	30.82, 292.85	reaching	9.56	43.14, 312.76
...	...	...	...	...	...	...	...	...

Figure 5. Example of first few lines of the exported file

A series of motion density distribution graphs, shown in Figure 6, were generated to identify the types and locations of motion occurrences within the context of the environment. In kitchen 1, the graphs indicate the user mostly occupied the L-shaped area at the back of the kitchen and stands between the stove and the sink at the back. Rotational motions occur throughout the active space, while Bending only occurs between the stove and sink, and Reaching occurs near the cabinets on the left. In kitchen 2, motions are mostly concentrated in front of the sink. There is a fair amount of rotation and bending motion in front of the sink. There are also some occurrences of Reaching motion near the top of the fridge.

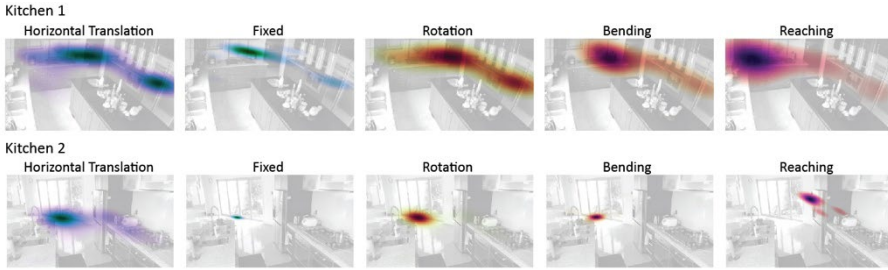


Figure 6. Graphs showing density and distribution of varying motion types in each kitchen space based on the activity recorded - Horizontal, Fixed, Rotation, Bending, Reaching

Design outcomes and suggestions can accommodate different action priorities and can change with varying case studies. For this work-in-progress case study, the data gathered is used to suggest areas of architectural rehabilitation to reduce reaching and/or reduce bending for users who have difficulties. Possible design suggestion areas to improve the user experience of the space in these two kitchens is to reduce bending and reaching motions by eliminating items stored in high or extremely low locations to avoid straining and inaccessibility. Figure 7 shows areas identified for possible architectural design adjustments in both kitchens.

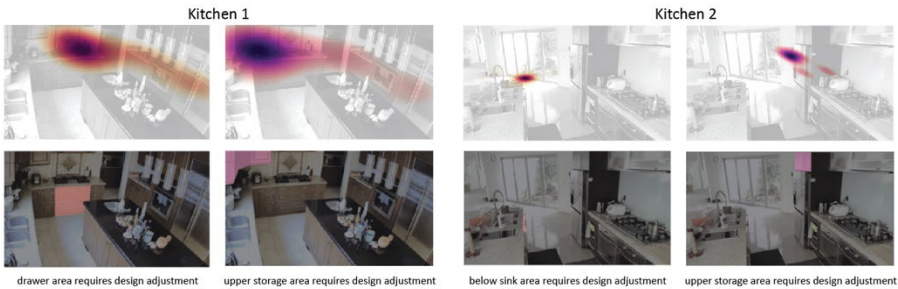


Figure 7. Suggested architectural design areas based on information from motion graphs

## 4. Discussion

### 4.1. LIMITATIONS

The framework is still under development, and further progress will address current limitations on the technical side. Factors such as lighting, and camera angles play a large role in the accuracy of the data. Experiments with two or more cameras located in different places in the same room would correct this limitation and help to provide more precise location data and motion information, particularly if one camera has any obstructed views from any furniture.

Potential privacy intrusions or ethical boundaries were considered throughout the research. While recorded videos with participants were used in the documented case study, unique private information related to participant identity, location and time is not important nor needed to conduct studies. The only information that is crucial is

joint identification, which is conducive to anonymity, and highly encouraged in any further work. Video capture is not needed in the real-time data collection process.

Future case studies would benefit from incorporating object and environment recognition, as well as more refined motion detection models, to further draw relationships between physical context and motions identified.

Furthermore, the framework could also be applied to building simulations if the input is changed from real-time video to a series of data points or recordings proposed by agent-based pedestrian simulation software.

While the framework considers motion specific to architectural interaction, it does not go into depth regarding every listed category. It is merely proposed as a starting point for which to think about how interaction can be measured within different spatial contexts and what categories would be important to study. It has the potential to be used for studying movement within fields of accessibility, architectural rehabilitation, simulation, renovation, as well as broader fields of health and wellbeing that are crucial to the UN Sustainable Development Goals, such as neuroimmunology and influence on homeostatic conditions, psychology, universal design, and usability studies for the built environment (Dougherty et al. 2018). The study indicates research opportunities for pose and motion estimation in aiding spatial evaluation.

## 4.2. CONCLUSION

The Elemental Motion in Spatial Interaction (ESMI) framework of using typologies of motion to analyse spatial experience in architecture proposes a new approach to evaluating an architectural space by breaking down human movement in an occupied space. This approach reduces inherent biases when studying existing spaces, providing insight into how architectural space is used and experienced by differing individuals, depending on the motion efforts exerted.

The case study demonstrates one possible workflow of implementing selected parts of the framework, using the aid of existing pose detection algorithms. This workflow is at an early stage with several limitations and many possibilities for improvement. Further development will include performing analysis on more than one person at a time, with an enhanced algorithm to remove lag and improve accuracy. This can be further developed to detect architectural elements in a video to establish a measurable relationship between motion and environment. While the current iteration of the framework has four predominant contextual categories, there are many more factors that could be considered and tested to determine relevance, reliability, and utility.

## References

- Aggarwal, J. K., & Ryoo, M. S. (2011). Human Activity Analysis: A Review. *ACM Computing Surveys*, 43(3). <https://doi.org/10.1145/1922649.1922653>
- Ajili, I., Malle, M., & Didier, J.-Y. (2017, September). Robust human action recognition system using Laban Movement Analysis. *Procedia Computer Science*. 21st International Conference on Knowledge-Based and Intelligent I, Marseille, France. <https://doi.org/10.1016/j.procs.2017.08.168>
- Blank, M., Gorelick, L., Shechtman, E., Irani, M., & Basri, R. (2005). Actions as space-time shapes. *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, 2, 1395-1402 Vol. 2. <https://doi.org/10.1109/ICCV.2005.28>

- Bobick, A. (1997). Movement, Activity and Action: The Role of Knowledge in the Perception of Motion. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 352, 1257–1265. <https://doi.org/10.1098/rstb.1997.0108>
- Bortolini, M., Gamberi, M., Pilati, F., & Regattieri, A. (2018). Automatic assessment of the ergonomic risk for manual manufacturing and assembly activities through optical motion capture technology. *Procedia CIRP*, 72, 81–86. <https://doi.org/10.1016/j.procir.2018.03.198>
- Callaghan, J., & Palmer, C. (1944). *Measuring space and motion*. John B. Pierce Foundation.
- Dougherty, B., & Arbib, M. (2013). The evolution of neuroscience for architecture: Introducing the special issue. *Intelligent Buildings International*, 5. <https://doi.org/10.1080/17508975.2013.818763>
- Fox, M., & Polancic, A. (2012). Conventions of Control: A Catalog of Gestures for Remotely Interacting with Dynamic Architectural Space. *ACADIA 12: Synthetic Digital Ecologies*.
- Gilbreth, F., & Gilbreth, L. (1917). *Applied motion study: A collection of papers on the efficient method to industrial preparedness*. Macmillan.
- Hirschberg, U., Sayegh, A., & Zedlacher, S. (2006). 3D motion tracking in architecture: Turning movement into form—Emerging uses of a new technology. *Communicating Space(s) 24th ECAADe Conference Proceedings*.
- Hu, Z., & Park, J. H. (2017). HalO [Indoor Positioning Mobile Platform]. *ACADIA 2017: DISCIPLINES & DISRUPTION [Proceedings of the 37th Annual Conference of the Association for Computer Aided Design in Architecture]*, 284–291.
- Marey, E.-J., & Pritchard, E. (1895). *Movement*. D.Appleton.
- Moeslund, T., & Granum, E. (2001). A Survey of Computer Vision-Based Human Motion Capture. *Computer Vision and Image Understanding*, 81, 231–268. <https://doi.org/10.1006/cviu.2000.0897>
- Mohamed, A. (2015). A Novice Guide towards Human Motion Analysis and Understanding. *CoRR*. <http://arxiv.org/abs/1509.01074>
- Muybridge, E. (1901). *The human figure in motion*. Chapman & Hall.
- Nagymate, G., & Kiss, R. (2018). Application of OptiTrack motion capture systems in human movement analysis: A systematic literature review. *Recent Innovations in Mechatronics*, 5. <https://doi.org/10.17667/riim.2018.1/13>
- Newlove, J. (1993). *Laban for Actors and Dancers: Putting Laban's Movement Theory into Practice—A step-by-step guide*. Routledge.
- Papandreou, G., Zhu, T., Chen, L.-C., Tompson, J., & Murphy, K. (2018). *PersonLab: Person Pose Estimation and Instance Segmentation with a Bottom-Up, Part-Based, Geometric Embedding Model*. Retrieved from <http://arxiv.org/abs/1803.08225>
- Peng, W., Zhang, F., & Nagakura, T. (2017). Machines' Perception of Space: Employing 3D Isovist Methods and a Convolutional Neural Network in Architectural Space Classification. *ACADIA 2017: DISCIPLINES & DISRUPTION [Proceedings of the 37th Annual Conference of the Association for Computer Aided Design in Architecture]*, 474–481.
- Prinsloo, T.-T., Munro, M., & Broodryk, C. (2019). The Efficacy of Laban Movement Analysis as a Framework for Observing and Analysing Space in 'Rosas danst Rosas'. *Research in Dance Education*, 20, 331–344.
- Schuldt, C., Laptev, I., & Caputo, B. (2004). Recognizing human actions: A local SVM approach. *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004*, 3, 32-36 Vol.3. <https://doi.org/10.1109/ICPR.2004.1334462>

# RETAIL COMMERCIAL SPACE CLUSTERING BASED ON POST-CARBON ERA CONTEXT

*A Case Study of Shanghai*

QINYU CUI<sup>1</sup>, SHUYU ZHANG<sup>2</sup> and YITING HUANG<sup>3</sup>

<sup>1,3</sup>*School of Architecture and Urban Planning, Shenzhen University, Guangdong, China.*

<sup>2</sup>*College of Architecture and Urban Planning, Tongji University, Shanghai, China.*

<sup>1</sup>*cuiqinyu2020@email.szu.edu.cn, 0000-0001-6687-9992*

<sup>2</sup>*hadyyu@gmail.com, 0000-0002-0459-4200*

<sup>3</sup>*906125150@qq.com, 0000-0001-5441-1612*

**Abstract.** In the post-carbon era, it has become a development and research trend on adjusting commercial locations to help achieve resource conservation by using big data. This paper uses multi-source urban data and machine learning to make reasonable evaluations and adjustments to commercial district planning. Many relevant factors are affecting urban commercial agglomeration, but how to select the appropriate ones among the many factors is a problem to be considered and studied, while there may be spatial differences in the strength of each influencing factor on commercial agglomeration. Therefore, this paper takes Shanghai, a city with a high economic and commercial development level in China, as an example and identifies the influencing factors through a literature review. Next, this paper uses the machine learning BORUTA algorithm of features selection to screen the influencing factors. It then uses multi-scale geographically weighted regression model (MGWR) to analyse the spatial heterogeneity of factors affecting retail spatial agglomeration. Finally, based on the background of the changing transportation modes and the unchanged social activities in the post-carbon era, the future spatial planning pattern of retail commercial space is discussed to provide particular suggestions for the future location adjustment of urban commerce.

**Keywords.** Business District Hierarchy; Agglomeration Effect; Spatial Variability; Multi-scale Geographically Weighted Regression Model; Machine Learning; Big Data Analysis; SDG 8; SDG 12.

## 1. Introduction

The post-carbon era is an ecologically harmonious, green, low-carbon, and sustainable society, and its main technical characteristics are digitization and intelligence (Rifkin,

2013). In this context, the transformation of the agglomeration characteristics of urban commercial space is an inevitable requirement to meet the needs of future human life. It is also an important guarantee for developing a city with high efficiency and low consumption.

Retail commercial space is the main place for human leisure activities. In the early days, empirical case studies were mainly carried out on the central place theory. With the development of space measurement technology, the combination of quantitative and qualitative research methods has been widely used. In quantitative research, effective selection based on influencing factors is very critical. The appropriateness of variable selection in previous studies was mostly tested by covariance and bivariate (Jia and Zhang, 2021). In addition, the spatial heterogeneity of influencing factors has always been a problem in the quantitative research of geospatial types. In urban quantitative research, in the areas of spatial regression models such as geography and economics, the fitting effect of linear regression models is usually inferior to that of geographically weighted regression (Zhang et al., 2019), mainly because of the uneven distribution of urban economic, social environment and the functional space of various entities in the city. Multi-scale geographic weighted regression models can give each relationship exclusive spatial bandwidth (Oshan et al., 2019), which can reflect the local, regional, or global spatial scale effects of different influencing factors and achieve a more accurate result model fitting effect.

In commercial space agglomeration research, although a lot of research has been conducted on the influencing factors of accumulation, there are few studies on the feature selection of influencing factors based on machine learning. In addition, quantitative research on urban commercial agglomeration based on geographically weighted regression is still extremely lacking. In the post-carbon era, people try to use clean energy as much as possible to reduce carbon emissions. This will change the way people travel, and there is a strong coupling between transportation stations and commercial clusters. In summary, in the context of the post-carbon era, with the changes in people's lifestyles, business agglomeration in new and old spaces has changed. This is also the purpose of this research.

Therefore, this paper selects Shanghai, which has relatively complete commercial development in China. Machine learning feature selection, multi-scale geographic weighted regression model, and spatial analysis technology are carried out on multi-source data to conduct in-depth analysis and research on the spatial heterogeneity of factors affecting commercial agglomeration in Shanghai.

## **2. Materials and methods**

### **2.1. STUDY AREA**

For this research, we selected the central district within the outer ring line of Shanghai, of around 660 km<sup>2</sup>, as the study area (Figure 1). From the perspective of industrial structure, Shanghai's tertiary industry accounts for the largest proportion of its GDP. As one of the main forces driving and supporting the development of Shanghai's tertiary industry, retail business reflects the development and expansion of Shanghai's business. Therefore, Shanghai is a suitable case city to study China's urban commercial



space structure.

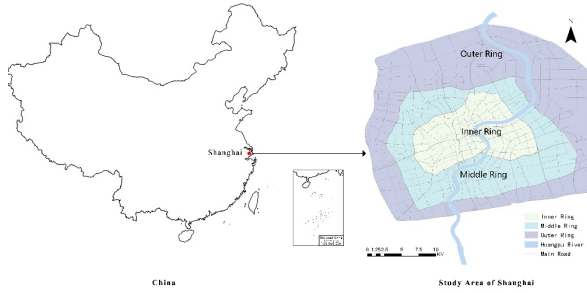


Figure 1. Location of Shanghai and the study area

2.2. VARIABLES AND DATA

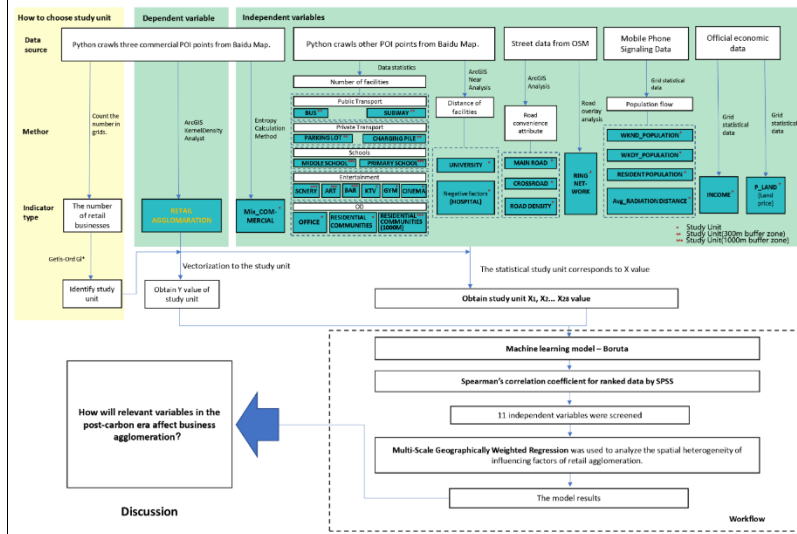


Figure 2. Indicator system and analytical framework

Agglomeration of retail spaces was considered the dependent variable, and a series of urban indicators were considered independent variables. The data sources and quantification methods of the variables are introduced below (Figure 2).

The dependent variable for this study was the accumulation of retail spaces, from now on referred to as “RETAIL AGGLOMERATION”. We collected data from the Baidu Map Open Platform (<http://lbsyun.baidu.com/>), crawling the names, commercial types, and spatial locations using Python programming language. Then we cleaned up the data, deleted the repeated and problematic data, and reclassified the data combined with the map to obtain the final data.

Based on the relevant research results at home and abroad and combined with research questions, this paper determines 27 variables affecting retail agglomeration in 4 types, 8 categories, among which the variables of 4 types are: transport properties,

supporting facilities, population flow, and others. 27 variables and their processing methods are detailed in figure 2.

### 2.3. METHODOLOGY

#### Step 1: Analysis of retail agglomeration by Kernel Density Estimation (KDE)

Commercial kernel density is a value used to evaluate commercial spatial density, influenced by neighborhood division. The Kernel density estimation (KDE) calculates the density of discrete nodes in different threshold ranges ( $h$ ) of the output grid cells of the research region (Fei *et al.*, 2019). A larger kernel density value indicates a stronger concentration—i.e., a higher degree of commercial space agglomeration. The kernel density at the center of the grid is the sum of the densities within the projection area of a kernel function  $K(\cdot)$ , as shown in Formula (1):

$$f^x = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) \quad (1)$$

where  $f^x$  is the kernel density at the center of the research region;  $h$  is the threshold value (i.e. the searching radius);  $n$  is the number of nodes in the threshold range;  $d$  is the dimension of the calculating data;  $K(\cdot)$  is a non-negative kernel function with the property of the probability density (Chen *et al.*, 2013).

#### Step 2: Analysis of Commercial Spatial Agglomeration Characteristics Using Moran's I and Getis-Ord $G_i^*$ algorithm

In our study, global Moran's I is used to evaluate the spatial clustering pattern of retail commercial space.

Next, we used the Getis-Ord  $G_i^*$  algorithm to measure retail commercial space distribution's hot or cold regions. When the  $G_i^*$  statistic is higher than the mathematical expectation and passes the hypothesis test, it is a hot spot, otherwise a cold spot.

Mathematically, Getis-Ord  $G_i^*$  is expressed as

$$G_i^* = \frac{\sum_{j=1}^n W_{i,j} x_j - \bar{x} \sum_{j=1}^n W_{i,j}}{\sum_j W_{i,j} x_j} \quad (2)$$

#### Step 3: Using machine learning BORUTA to select suitable variables and make correlation analysis

This study considers several variables that affect business accumulation, some of which have significant effects. In contrast, others have no significant effects, so references' point is needed to help distinguish the really important variables from the unimportant ones. We have applied the BORUTA algorithm that provides criteria for selecting relevant features to cope with this problem.

The BORUTA algorithm applies random forests for feature relevance estimation. This method mainly compares real predictive variables' importance with random shadow variables through statistical tests and multiple RF runs. A copy of each variable is added in each run, doubling the set of predictors. Variables with significantly larger or smaller importance values are declared important or unimportant, respectively. Then delete all unimportant variables and shadow variables, and repeat the previous steps until all variables are classified, or the pre-specified number of runs is performed.

After screening the significant variables by Boruta, the collinearity of independent and dependent variables was detected, and Spearman’s rank correlation coefficient was carried out by SPSS tool.

**Step 4: MGWR was used to analyse the spatial heterogeneity of variables of retail agglomeration**

Multi-scale Geographic Weighted Regression (MGWR) was proposed by Fotheringham (2017) and Yu et al. (2019); they improved the statistical inference, which is superior in replicating parameter surfaces with different levels of spatial heterogeneity and provides valuable information on the scale at which different processes operate. In this study, we use MGWR to analyze the spatial heterogeneity and scale difference of the influencing factors of commercial spatial agglomeration. Mathematically, MGWR is expressed as

$$y_i = \sum_{j=1}^k \beta_{bwj}(u_i, v_i) x_{ij} + \varepsilon_i \tag{3}$$

Where  $\beta_{bwj}$  represents the regression coefficient of local variables,  $bwj$  represents the bandwidth used by the regression coefficient of variable  $j$ ,  $(u_i, v_i)$  represents the geographical spatial coordinates of sample point  $i$ ,  $x_{ij}$  represents the observed value of variable  $j$  at sample point  $i$ , and  $\varepsilon_i$  represents the random disturbance term.

**3. Previous Analysis**

**3.1. RETAIL COMMERCIAL SPACE CLUSTERING EXTRACTION**

*Kernel Density and hot spot analysis*

After comparing and analyzing the four kernel density search radii of 400m, 800m, 1200m, and 1600m, respectively, the search radius was set to 400m for retail commercial kernel density analysis by combining the common Shanghai neighborhood size (Figure 3). Subsequently, by dividing neighborhoods and 400×400m grids by major roads as the study units, we made an exploratory attempt to use the Getis-Ord  $G_i^*$  index to explore the spatial distribution hierarchy of retail businesses in Shanghai. Since the area size of the neighborhoods in the study area varies greatly, the hot spot analysis effect is different from common sense, so the grid was finally adopted as the study unit for hot spot analysis (Figure 3).

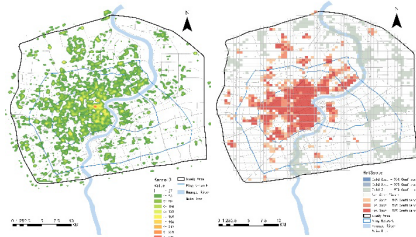


Figure 3. Retail commercial space clustering form map and hot spot distribution

### *Study unit extraction*

Spatial sets with Z-values greater than 1.65 are hotspot areas selected as the study units of this paper, and the kernel density values are extracted vectorized into the study units. Spatial visualization was performed in ArcGIS Pro 2.5 by the natural breakpoint method (Figure 4-a), and the spatial distribution of the extracted study units was roughly mostly clustered within the inner ring and scattered on the outer side. The high-value regions are also mostly concentrated in the inner ring, with some local high-value regions formed sporadically on the outer side.

In addition, since some of the variables in the previous paper (e.g., traffic and activity space) need to be considered under the 5/15-minute living circle scale, the study units were made into 300m and 900m buffers, and the values of the corresponding variables were counted separately (Figure 4-b/c/d).

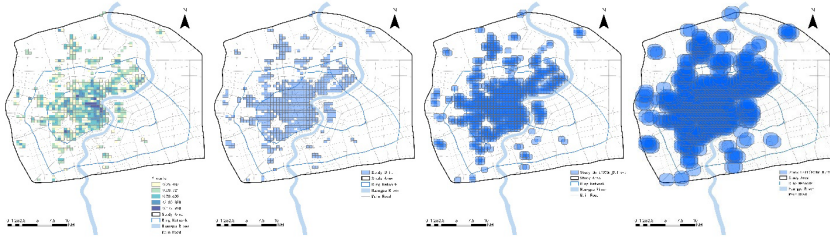


Figure 4. Study unit related features: (a) average kernel density value for each unit; and (b) study unit; and (c) study unit (300m buffer zone); and (d) study unit (900m buffer zone).

## 3.2. MODEL VARIABLE SELECTION

### *Machine learning BORUTA and Spearman correlation analysis*

According to the results of BORUTA variable screening, there are 11 definite acceptance variables, 2 tentative variables, and the rest are rejection variables (Figure 5).

This paper uses SPSS software to perform Spearman correlation analysis on the independent variables determined to be accepted and the dependent variable and visualize them. We can see that the correlation between the independent and dependent variables is strong, showing significance at the 1% statistical level. In addition, the correlation coefficients of the dependent variables are all less than 0.8 with each other, there is no multicollinearity, and all the variables determined to be accepted can be put into the model operations (Figure 6).

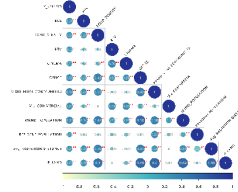
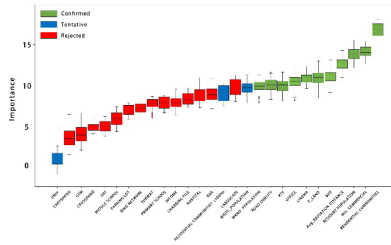


Figure 5. Feature selection importance graph using the BORUTA algorithm

Figure 6. Correlation coefficient heat map.

## 4. Result Analysis

### 4.1. MODEL BUILDING

Table 1. Model diagnostic information description

Model	Adj. R <sup>2</sup>	RSS	AICc	RSS Moran's I
OLR	0.442	359.305	1420.926	0.077***
GWR	0.563	283.667	1367.037	-0.017
MGWR	0.598	270.749	1339.839	-0.024

MGWR model has the highest fit, the smallest residual sum of squares and AICc, and the best overall fit among the three models (Table 1). In addition, the residuals of the OLR model had significant positive spatial autocorrelation. In contrast, the residuals of the GWR and MGWR models did not have significant spatial autocorrelation, indicating that the geographically weighted class of models considering spatial heterogeneity outperformed the global model in dealing with the spatial autocorrelation of residuals (Harris *et al.*, 2018). In summary, the Shanghai business agglomeration can be modeled by the influencing factors selected in this paper to build the MGWR model for regression analysis.

### 4.2. ANALYSIS OF SPATIAL HETEROGENEITY OF INFLUENCING FACTORS OF COMMERCIAL SPACE AGGLOMERATION

On the whole (Figure 7):

- 1) the population size formed by the residential and leisure behaviours of the population is the primary factor constituting retail commercial accumulation;
- 2) the impact of two different quantitative scales of social activities, movie watching and KTVs, on urban retail spatial agglomeration, showed complementary spatial distribution;
- 3) the intensity of the impact of land price, commercial mixing degree, public transportation and commercial buildings on commercial accumulation did not differ geographically and spatially, and the remaining relevant factors of The uneven spatial

distribution of the other relevant factors are the main reason for the differences in the spatial concentration of urban retail businesses;

4) In addition, the higher the density of the road network, the greater the accessibility of crowd activities, which is more conducive to the concentration of businesses.

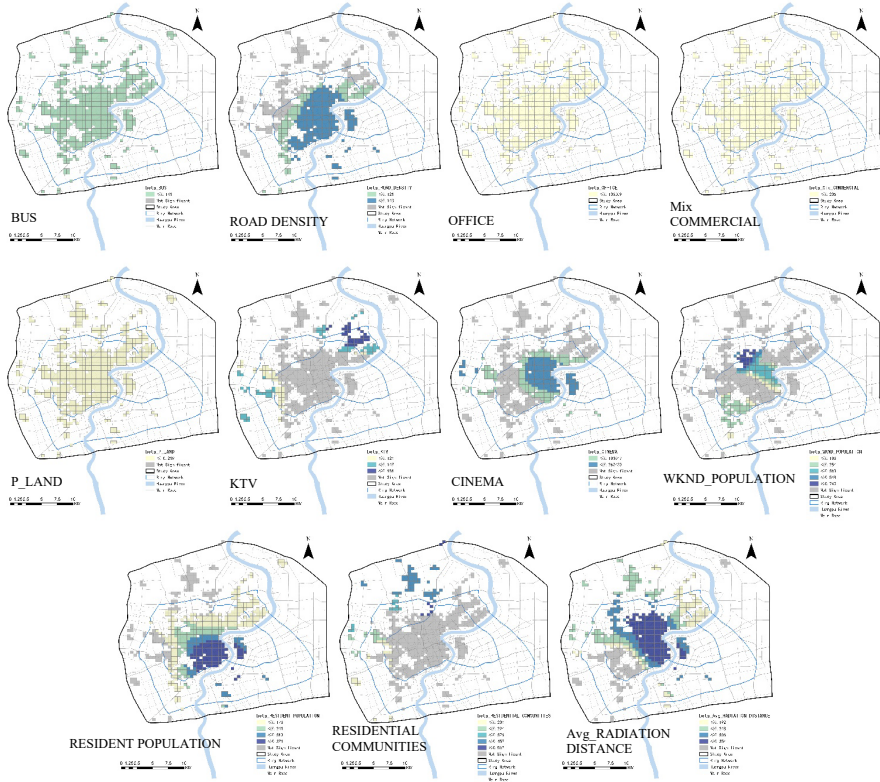


Figure 7. Spatial heterogeneity of factors influencing the accumulation of commercial spaces

## 5. Conclusion and Discussion

In this paper, we took retail commerce within the central city of Shanghai as an example. Based on machine learning BORUTA screening variables and regression analysis using the MGWR model, we determined the positive effects of relevant influencing factors on retail commercial accumulation and the reasons for spatial differences.

In the context of future urban development in the post-carbon era, some of the lifestyles of city dwellers will change. The transformation of new energy, the further development of high technology, the Internet of Things, the way urban residents travel, and their consumption habits bring changes, such as new energy vehicles, driverless technology, and online order consumption. The offline, face-to-face activities of urban residents are also a typical example. Despite the rapid development of technology in the past 20 years, it has not been greatly affected or replaced, and it is still an important

part of urban residents' lives.

Figure 8 shows that some retail businesses used to serve people who traveled in both directions from home or work to public transportation and clustered in this space. In the future, the clustering characteristics of retail commerce will favour the new public transportation space near the work area, such as the charging piles or parking lots of shared trams, and the realization of a 5-minute service circle centered on various types of transportation nodes.

Some factors affecting retail business agglomeration are inferred in chapter 4.2 of the paper. Based on this conclusion, the agglomeration characteristics of retail business in the post-carbon era are discussed.

1) The model analysis shows that the accumulation of retail business is closely related to the crowd's activities, and the current business agglomeration is still clustered around residential and entertainment areas.

2) Living, commuting, and work are the main components of a working day for urban residents. The areas from residence to transportation stations and from work to transportation stations are the key areas for small businesses to gather.

3) The transportation mode of commuting in the post-carbon era will change. For example, electric vehicles need to be replenished by charging piles, which may become new traffic sites. At present, the retail business layout in this part of the area is still blank.

4) The Covid-19 epidemic has changed people's consumption habits and promoted the new retail model of online shopping-offline distribution. In the future, some offline retail stores will be transformed into non-core links responsible for sorting, packaging, distribution, etc., and the rest will be transformed into experience centers, exhibition centers, and service centers. This business model transformation can ensure the upgrade of consumer experience and minimize operating costs.

5) The service space generated by large-scale retail business clusters is the main place for social and leisure activities at work and weekends. Its functions include high-end brand image display, mid-to-high-end catering consumption experience, and physical leisure space such as KTV and cinema. This space has a high degree of commercial mixing, rich commercial types, grades, and functions. The function and service radius will not change greatly and will remain in the future.

To sum up, commercial gatherings of different scales will undergo different changes, but commercial gatherings are always process-oriented by crowd activity. Figure 8 presents our summary and assumptions about the characteristics of retail business accumulation in the current and post-carbon era.

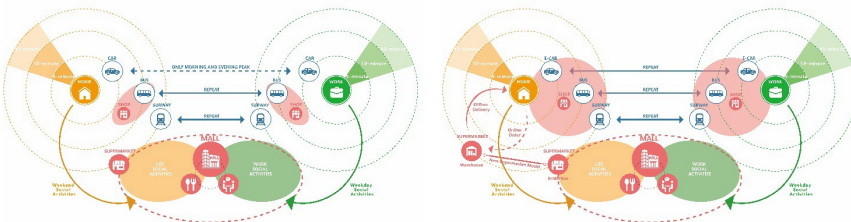


Figure 8. Comparison between the present and the post-carbon era: the characteristics of retail business agglomeration

## 6. Application and Production

The results of this paper will assist commercial real estate investment companies in evaluating the investment value and appreciation potential of projects in various business districts in Shanghai. This paper obtained data and platform support from Metrodata Technology Company to transform research into a platform product, as shown in figure 9 below.

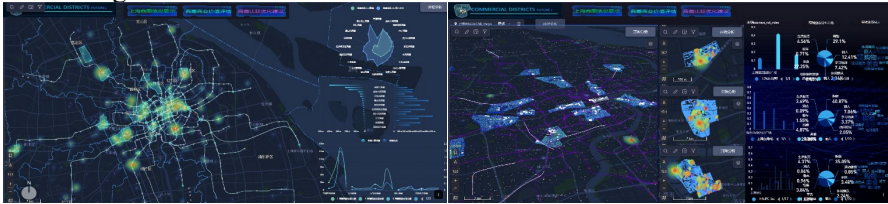


Figure 9. Platform support from Metrodata Technology Company

## Acknowledgements

Metrodata Technology Company supports the data and analysis platform of this paper. (<https://www.metrodata.cn/>)

## References

- Chen, C., Lin, C., Xiu, C. (2013). Distribution of centrality of traffic network and its relationship with economic density of tertiary industry in Shenyang. *Progress in Geography*, 32(11), 1612–1621.
- Fei, M. A., Fr, A., Kfy, B., Yg, A., Cz, A., Dan, G. A. (2019). The spatial coupling effect between urban public transport and commercial complexes: A network centrality perspective - ScienceDirect. *Sustainable Cities and Society*, 50, 101645-101645.
- Fotheringham, A. S., Yang, W., Kang, W. (2017). Multiscale Geographically Weighted Regression (MGWR). *Annals of the American Association of Geographers*, 107(6), 1247-1265.
- Harris, P., Fotheringham, A. S., Crespo, R., Charlton, M. (2018). Inference in multi-scale geographically weighted regression. *WILEY*, 43(3), 399.
- Jia, J., Zhang, X. (2021). A human-scale investigation into economic benefits of urban green and blue infrastructure based on big data and machine learning: A case study of Wuhan. *Journal of Cleaner Production*, 316, 128321.
- Oshan, T. M., Li, Z., Kang, W., Wolf, L. J., Fotheringham, A. S. (2019). MGWR: A Python Implementation of Multiscale Geographically Weighted Regression for Investigating Process Spatial Heterogeneity and Scale. *International Journal of Geo-Information*, 8(6), 269.
- Rifkin, J. (2013). The Third Industrial Revolution. *International Study Reference*, 6(1), 8-11.
- Yu H, Fotheringham A S, Li Z, et al. (2019). Inference in multi-scale geographically weighted regression. *Geographical Analysis*, 52(1), 87-106.
- Zhang, S., Wang, L., & Lu, F.. (2019). Exploring housing rent by mixed geographically weighted regression: a case study in Nanjing. *International Journal of Geo-Information*, 8(10), 431.



# DIGITAL TWIN-BASED RESILIENCE EVALUATION OF DISTRICT-SCALE ARCHETYPES

*A COVID-19 scenario case study using a university campus pilot*

PRADEEP ALVA<sup>1</sup>, MARTIN MOSTEIRO-ROMERO<sup>2</sup>, CLAYTON MILLER<sup>3</sup> and RUDI STOUFFS<sup>4</sup>

<sup>1,2,3,4</sup>*National University of Singapore*

<sup>1</sup>*akipaa@nus.edu.sg, 0000-0001-7997-1894*

<sup>2</sup>*mosteiro@nus.edu.sg, 0000-0001-7160-0298*

<sup>3</sup>*clayton@nus.edu.sg, 0000-0002-1186-4299*

<sup>4</sup>*stouffs@nus.edu.sg, 0000-0002-4200-5833*

**Abstract.** District-scale energy demand models can be powerful tools for understanding interactions in complex urban areas and optimising energy systems in new developments. The process of coupling characteristics of urban environments with simulation software to achieve accurate results is nascent. We developed a digital twin through a web map application for a 170ha district-scale university campus as a pilot. The impact on the built environment is simulated with pandemic (COVID-19) and climate change scenarios. The former can be observed through varying occupancy rates and average cooling loads in the buildings during the lockdown period. The digital twin dashboard was built with visualisations of the 3D campus, real-time data from sensors, energy demand simulation results from the City Energy Analyst (CEA) tool, and occupancy rates from WiFi data. The ongoing work focuses on formulating a resilience assessment metric to measure the robustness of buildings to these disruptions. This district-scale digital twin demonstration can help in facilities management and planning applications. The results show that the digital twin approach can support decarbonising initiatives for cities.

**Keywords.** Digital twin; City Information Modelling; Planning Support System; Energy Demand Model; SDG 11.

## 1. Introduction

A digital twin is a virtual imitation of physical processes that can be updated in real-time with their physical counterparts. In 2002, Michael Grieves introduced the concept of digital twins in production engineering. Explorations of digital twins took place rapidly in various scales and domains. However, the concept has widened and loosened to applications portraying digital simulation models, which relate to social and economic systems as well as the physical systems. Numerous scholars emphasise the

conflict in the definition of digital twins concerning a city or district (Batty, 2018; Kandt and Batty, 2021). Furthermore, digital twins that represent city's physical assets rarely include processes that define its social and economic function. However, Stoter et al. (2021) highlights the present consensus among researchers that urban digital twins should: (i) be based on 3D city models; (ii) contain objects with geometric and semantic information; (iii) contain real-time sensor data; and (iv) integrate a variety of analyses and simulations to be able to make the best design, planning and intervention decisions. Importantly, in order to support those decisions, all information should be presented to users (citizens, decision-makers, experts) in a user-friendly visualisation of the digital twin in a one-stop-shop dashboard. Similarly, Wildfire (2018) makes a useful distinction between digital twins that represent high-frequency and low-frequency cities. Low-frequency cities operate with long-term planning over years, decades, and centuries. But high-frequency cities operate by sensing in short time frames, second by second, minute by minute up to cycles of days and months.

A city- or district-scale digital twin needs to contain geospatial urban environment data models in time series. The urban environment data models include building information modelling (BIM) for individual buildings, landscape, mobility, climate, demography, and infrastructure systems; these models are also known as city information modelling (CIM) (Gil et al., 2011). CIM extends the use of geographic information systems (GIS) in urban planning as decision support tools through integration with computer-aided design (CAD) (Batty et al., 1998; Dave & Schmitt, 1994; Webster, 1993). The Open Geospatial Consortium (OGC) standard CityGML serves as one of the most comprehensive 3D geospatial data container formats for CIM (Gröger et al., 2008; Kolbe, 2009). The conversion of detailed architectural models stored in IFC to semantic 3D city models in CityGML is a topical subject (Liu et al., 2017; Noardo et al., 2020). Some works in the literature (Biljecki et al., 2021; Stouffs et al., 2018) try to achieve complete and near-lossless conversion from IFC to CityGML.

Coupling urban environment models with simulation software is a nascent approach to predict accurate results (Mosteiro-Romero et al., 2020). Additionally, district-scale energy demand models contribute to planning support systems by understanding complex interactions in urban areas. In our paper, we introduce a workflow to create a digital twin from an open-source energy simulation software, City Energy Analyst (CEA).

In the following section, we introduce our methodology with subsections: (i) digital twin framework and data modelling, demonstrating a pilot case study and introducing our framework to create a digital twin dashboard; (ii) creating inputs for different post-pandemic occupancy and climate change scenarios; and (iii) dashboard and visualisation, in which we explain the salient characteristics of this district-scale digital twin dashboard and its preparation. Subsequently, in the discussion section, we explain a resilience evaluation method for the disruption caused due to the pandemic. Finally, we conclude with suggestions for research investigating extensions to these ideas.

## 2. Methodology

As a pilot, we developed a digital twin through a web map application for a district-scale university campus of area 170ha with around 300 buildings. Initially, we

modelled the campus using City Energy Analyst (CEA), which extracts the building footprint and height information (LOD0) from OpenStreetMap (OSM). Subsequently, the 2D shapefiles of the buildings were extruded with OSM heights and then converted to CityGML and 3D tiles format for web streaming. At the same time, detailed IFC models (LOD3) of a few buildings on the campus were converted and merged with the rest of the model with an IFC to CityGML pipeline. The digital twin dashboard was built with visualisations of 3D campus, real-time data from sensors, energy demand simulation results from CEA, and occupancy rates from WiFi data.

2.1. DIGITAL TWIN FRAMEWORK AND DATA MODELING

Figure 1 shows the development workflow of the digital twin dashboard as a web map application. The process started with extracting the 2D input shapefiles from City Energy Analyst (CEA). CEA acquires the geometries and attributes directly from OpenStreetMap (OSM). This process generally takes many iterations to check the validation of the data. It is possible that the OSM attributes don't correlate to physical reality. Hence, building typology and the number of floors were specifically checked and updated as required.

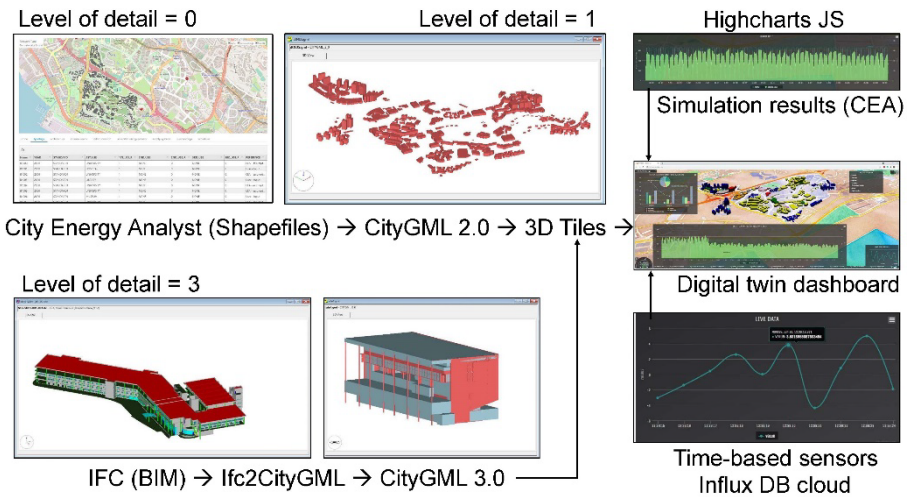


Figure 1. Framework for the district-scale digital twin dashboard as a web application

The campus has a variety of buildings, whose predominant usage is not limited to academia. Using CEA, buildings were classified into university, lab, office, residential, retail, restaurant, hospital, gym, museum, and library as per usage. Additionally, this step helped energy modelling within CEA and the results can be viewed in the inbuilt CEA dashboard (see Figures 2-4). The simulation results are generated in CSV file format within the project folder of CEA. The building shapefiles are then merged with the output simulation results. This merging procedure can be done either in QGIS or FME with the support of unique building IDs assigned in CEA.

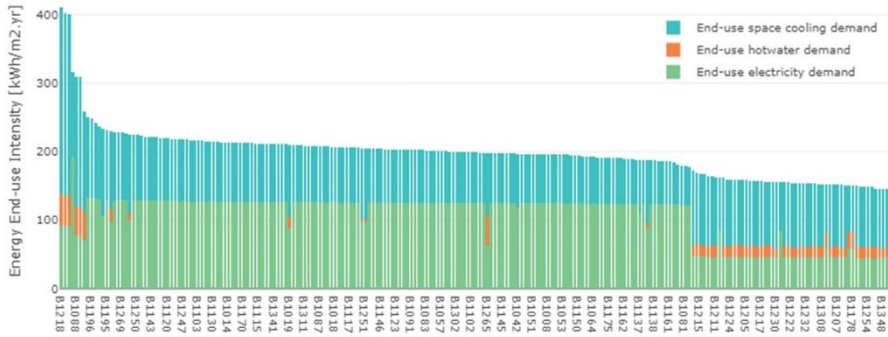


Figure 2. City Energy Analyst (CEA) - Energy end-use intensity for the district.

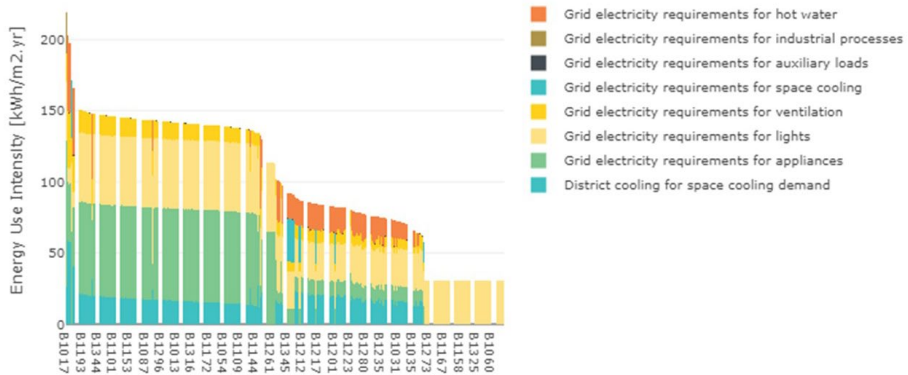


Figure 3. City Energy Analyst (CEA) - District Grid Energy final use intensity.

For instance, the energy demand CSV with total building electricity consumption and cooling load results were merged with the geometry. However, the geometry from CEA is two-dimensional (Level of detail = 0), basically simple polygons of building footprints. Thus, the polygons were imported to FME with a default shapefile reader and extruded to their building heights. Later, a writer was used to convert the 3D geometry into CityGML and Cesium 3D Tile format (Cesium GS, 2021). Detailed IFC-BIM models (LOD3) of a few buildings on campus were combined with the rest of the model (LOD1) using an IFC to CityGML work pipeline (Stouffs et al., 2018).

All the converted 3D tiles were used to create the web map application as the digital twin dashboard. In addition, time-based sensor data stored in the InfluxDB cloud (an open-source cloud-native serverless platform) was accessed through a robust API to display high-frequency building information. This real-time sensor data was acquired from weather stations installed in campus. These capture indoor temperature, humidity etc. at a 10-minute frequency. Finally, Highcharts (a multiplatform charting JavaScript library) was used to display the simulation results and live data interactively on the dashboard (Chaturvedi & Kolbe, 2019; Würstle et al., 2020).

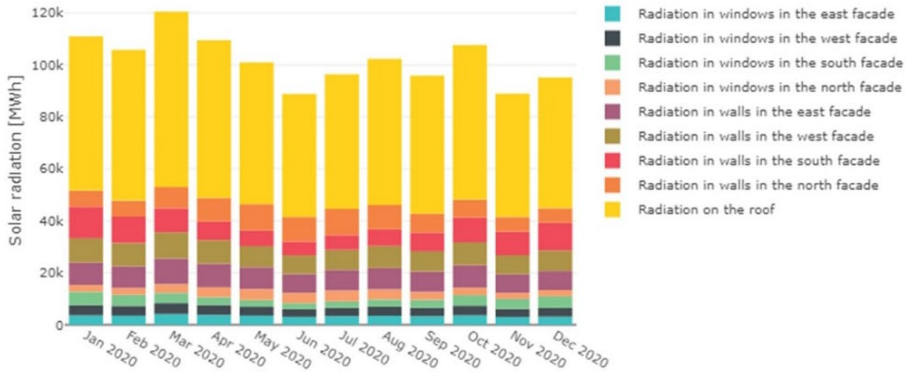


Figure 4. City Energy Analyst (CEA) - Monthly solar radiation simulated for the district.

## 2.2. CREATING SCENARIO INPUTS FOR OCCUPANT MODELLING

To test the digital twin's capability to display and compare different possible future trajectories for the case study area, scenarios on the impact on the built environment of the COVID-19 pandemic and climate change are analysed. For the former, we consider the long-term effects of increased remote working, while for the latter we analyse the long-term effects of changing climate patterns on space cooling demand.

Pre-pandemic occupancy patterns for the buildings in the case study area are derived from hourly data on the number of devices connected to the campus WiFi network at each hour of the year 2018, following an approach similar to Zhan and Chong (2021).

Two parameters are subsequently varied:

- Parameter 1 – Work from home ratio (WFH): The share of building occupants that work from home on any given day, vary between 0%, 25%, 50% and 75%.
- Parameter 2 – Operating building floor area ratio (OFA): The share of a buildings gross floor area is correspondingly varied between 100%, 75%, 50%, and 25% according to the share of occupants present in order to compare the effects of different building operation strategies on buildings' energy performance.

We create the input scenarios by varying these parameters and compare the resulting hourly space cooling demand simulated from 2020 to 2060 to a baseline of 0% WFH and 100% OFA (see Figure 5). In order to account for the long-term effects of climate change, future weather patterns are extracted from the Coupled Model Intercomparison Project (CMIP) and using the R package *epwshiftr* (Chong & Jia, 2021). Yearly weather files from 2020 to 2060 are then passed to the CEA simulation.

We then combine the results (.csv file) with the 3D model as explained in section 2.1. The metadata is further used to create a user-interactive visualisation in the dashboard for the respective decade and parameters.

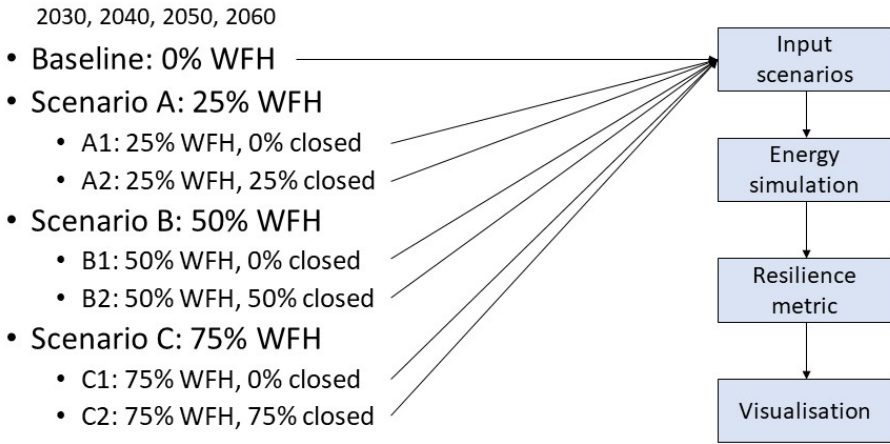


Figure 5. Input scenarios for different post-pandemic remote working shares and corresponding building occupancy

### 2.3. DASHBOARD AND VISUALISATION

The digital twin dashboard is set up using XAMPP (an open-source cross-platform web server solution stack package) for testing. The campus model in 3D Tile format and the metadata acquired from the CEA simulation result is used for styling in JavaScript. Further visualisation is customised within JavaScript as specified in Cesium 3D Tile format. We have created various scenes in the dashboard for users to interactively navigate (see Figure 6-8).

The dashboard display can switch between scenes thematically based on building typology, cooling loads, highlighting individual building typology (academic, residential, office etc.), thermal network of building cluster, energy-use intensity, and its ratio to occupancy. Tables 1-2 show the attributes assigned to each building in the campus - Unique building ID, gross floor area, occupancy, building height, number of floors, base simulation results (building electricity consumption, cooling load), WFH and OFA based cooling load results.

Name	GFA_m2	People	E_sys_MWh/y	EIntensity	QC_sys0_kW	height_ag	floors_ag
B1053	15759.433	375	1964.045	124.627	449.878	18	6
B1099	7549.806	217	960.958	127.282	354.140	18	6
B1114	20548.821	795	2549.615	124.076	541.343	27	9
B1104	9982.529	179	1253.797	125.599	267.545	18	6
B1047	31905.023	442	3956.811	124.018	2137.702	33	11

Table 1. Attributes of building - unique building ID, gross floor area, occupancy, building electricity consumption, Intensity, cooling load, building height and number of floors

Name	2030_0_OQC	2030_2_OQC	2030_5_OQC	2030_7_OQC	2040_0_OQC	2040_2_OQC	2040_5_OQC	2040_7_OQC
B1053	687.8	656.085	619.28	590.55	682.576	651.282	617.296	584.478
B1099	554.919	537.488	517.389	500.032	551.749	535.272	516.623	506.256
B1114	832.288	781.08	722.676	650.729	798.994	750.619	699.026	636.765
B1104	468.027	451.177	431.963	424.798	458.166	443.835	429.815	426.731
B1047	2140.195	2087.288	2031.877	2025.018	2128.446	2081.494	2030.752	2009.785

Table 2. Attributes of building - cooling loads for year 2030 with 0% WFH, 25% WFH, 50% WFH, 75% WFH, for year 2040 with 0% WFH, 25% WFH, and 50% WFH



Figure 6. Digital twin dashboard showing the building types scene



Figure 7. Digital twin dashboard showing the cooling load scene

Using the Highcharts-javascript library, the simulation results, the comparison, and real-time sensor data are displayed interactively on the dashboard. The Highcharts library has tools that can be used to create customised data visualizations. And its charting libraries work with back-end database or server stack. Highcharts offer wrappers for programming languages such as .Net, PHP, Python, R, Java, and frameworks such as Angular, Vue, and React (Mishra, 2020).

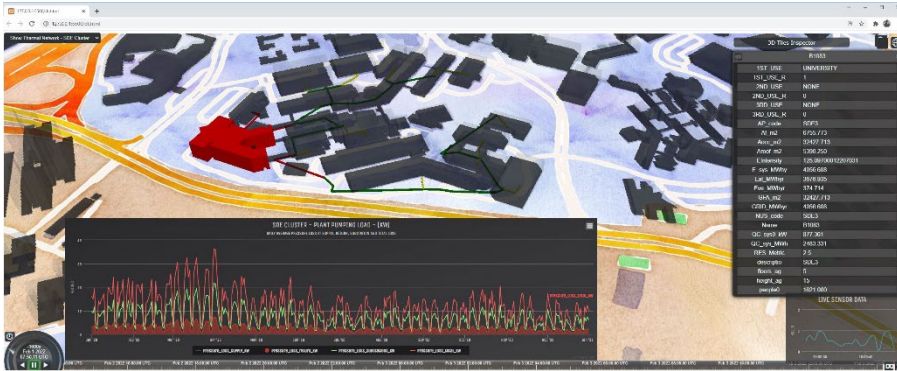


Figure 8. Digital twin dashboard showing the thermal network scene

### 3. Discussion

As a pilot, we developed a digital twin through a web map application for a district-scale university campus of area 170ha. Overall, the digital twin dashboard was built with visualisations of a 3D campus, real-time data from sensors, energy demand simulation results from CEA, and occupancy rates from WiFi data. The impact on the built environment is simulated with post-pandemic and climate change scenarios. This can be observed in various scenarios created through varying occupant rates and corresponding average cooling loads in the buildings. We create the input scenarios by varying WFH and OFA parameters and compare the resulting hourly space cooling demand simulated from 2020 to 2060 to a baseline of 0% WFH and 100% OFA. The users then get the opportunity to try different combinations of parameters to see corresponding results on the dashboard.

With this the created digital twin of built environment correspond to the social and economic systems during a pandemic. It helps to understand the consequences of such disruptions on the built environment and design resilience as a proactive measure. We believe the digital twin dashboard will be further used by urban designers, city planners, policymakers, facilities managers, and other stake holders. For instance, the changes of consumption patterns during and after pandemic can be approximately estimated with various parameter combinations. It may help in decisions regarding a new development within the district, as decision-makers can rely on the estimate of consumption.

A study on behavioural patterns of users and energy consumption in district scale is important for building sustainable future cities. This pilot study of digital twins in energy systems management on a district scale adds further to the discourse of creating large scale multi-variate models. And our future work will focus on formulating a resilience assessment metric to measure the robustness of buildings to these disruptions.

### 4. Conclusion

There is a need for conceptual and empirical work on digital twins that focuses on energy systems management on a district scale. In this paper, we identified issues related to building digital twins which can effectively contribute to the discussion. We



have developed a use case of a digital twin through a web map application of district scale. We describe our methodology for creating the dashboard using a framework. Correspondingly, input scenario variation for occupant modelling is explained. And our ongoing work focuses on formulating a resilience assessment metric to measure the robustness of buildings to these disruptions. This district-scale digital twin demonstration can help in facilities management and planning applications. The paper further contributes to the development of digital twin approaches that can support decarbonising initiatives for cities.

### Acknowledgements

This work is an outcome of the Future Resilient Systems project "Digital Twin-Enabled System Resilience" at the Singapore-ETH Centre (SEC) supported by the National Research Foundation, Prime Minister's Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme.

### References

- Batty, M. (2018). Digital twins. *Environment and Planning B: Urban Analytics and City Science*, 45(5), 817–820. <https://doi.org/10.1177/2399808318796416>
- Batty, M., Dodge, M., Jiang, B., & Hudson-Smith, A. (1998). *GIS and urban design*. UCL (University College London), Centre for Advanced Spatial Analysis (UCL): London.
- Biljecki, F., Lim, J., Crawford, J., Moraru, D., Tauscher, H., Konde, A., Adouane, K., Lawrence, S., Janssen, P., & Stouffs, R. (2021). Extending CityGML for IFC-sourced 3D city models. *Automation in Construction*, 121, 103440. <https://doi.org/10.1016/J.AUTCON.2020.103440>
- The CEA Team. (2021). *CityEnergyAnalyst v3.24.0 (v3.24.0)*. Zenodo. <https://doi.org/10.5281/zenodo.5037913>
- Chaturvedi, K., & Kolbe, T. H. (2019). Towards establishing cross-platform interoperability for sensors in smart cities. *Sensors (Switzerland)*, 19(3). <https://doi.org/10.3390/S19030562>
- Chong, A., & Jia, H. (2021). *GitHub - ideas-lab-nus/epwshiftr: Create future EnergyPlus Weather files using CMIP6 data*. <https://github.com/ideas-lab-nus/epwshiftr>
- Dave, B., & Schmitt, G. (1994). Information systems for urban analysis and design development. *Environment and Planning B: Planning and Design*, 21(1), 83–96.
- Gil, J., Almeida, J., & Duarte, J. (2011). The backbone of a City Information Model (CIM): Implementing a spatial data model for urban design. In *29th International Conference on Education and research in Computer Aided Architectural Design in Europe: Respecting fragile places, eCAADe 2011* (pp. 143-151). [http://papers.cumincad.org/cgi-bin/works/paper/ecaade2011\\_104](http://papers.cumincad.org/cgi-bin/works/paper/ecaade2011_104)
- Gröger, G., Kolbe, T. H., Czerwinski, A., & Nagel, C. (2008). *OpenGIS® City Geography Markup Language (CityGML) Encoding Standard. Version 1.0.0*. <https://doi.org/10.25607/OBP-635>
- Kandt, J., & Batty, M. (2021). Smart cities, big data and urban policy: Towards urban analytics for the long run. *Cities*, 109, 102992. <https://doi.org/10.1016/J.CITIES.2020.102992>
- Kolbe, T. H. (2009). *Representing and exchanging 3D city models with CityGML*. Lecture Notes in Geoinformation and Cartography, 15–31. [https://doi.org/10.1007/978-3-540-87395-2\\_2](https://doi.org/10.1007/978-3-540-87395-2_2)

- Liu, X., Wang, X., Wright, G., Cheng, J. C. P., Li, X., & Liu, R. (2017). A state-of-the-art review on the integration of Building Information Modeling (BIM) and Geographic Information System (GIS). *ISPRS International Journal of Geo-Information*, 6(2), 53.
- Mishra, S. (2020). *Practical highcharts with angular: Your essential guide to creating real-time dashboards* (1st 2020. ed.). Berkeley, CA: Apress.
- Mosteiro-Romero, M., Maiullari, D., Pijpers-van Esch, M., & Schlueter, A. (2020). An Integrated Microclimate-Energy Demand Simulation Method for the Assessment of Urban Districts. *Frontiers in Built Environment*, 6. <https://doi.org/10.3389/FBUIL.2020.553946>
- Noardo, F., Harrie, L., Ohori, K. A., Biljecki, F., Ellul, C., Krijnen, T., Eriksson, H., Guler, D., Hintz, D., Jadidi, M. A., Pla, M., Sanchez, S., Soini, V. P., Stouffs, R., Tekavec, J., & Stoter, J. (2020). Tools for BIM-GIS integration (IFC georeferencing and conversions): Results from the GeoBIM benchmark 2019. *ISPRS International Journal of Geo-Information*, 9(9). <https://doi.org/10.3390/IJGI9090502>
- Stoter, J., Ohori, K. A., & Noardo, F. (2021, October 4). *Digital Twins: A Comprehensive Solution or Hopeful Vision?* | *GIM International*. <https://www.gim-international.com/content/article/digital-twins-a-comprehensive-solution-or-hopeful-vision>
- Stouffs, R., Tauscher, H., & Biljecki, F. (2018). Achieving complete and near-lossless conversion from IFC to CityGML. *ISPRS International Journal of Geo-Information*, 7(9). <https://doi.org/10.3390/IJGI7090355>
- Webster, C. J. (1993). GIS and the scientific inputs to urban planning. Part 1: description. *Environment & Planning B: Planning & Design*, 20(6), 709–728. <https://doi.org/10.1068/B200709>
- Wildfire, C. (2018). *How can we spearhead city-scale digital twins?* <http://www.infrastructure-intelligence.com/article/may-2018/how-can-we-spearhead-city-scale-digital-twins>
- Wüstle, P., Santhanavanich, T., Padsala, R., & Coors, V. (2020). The Conception of an Urban Energy Dashboard Using 3D City Models. *Proceedings of the Eleventh ACM International Conference on Future Energy Systems*, 523–527. <https://doi.org/10.1145/3396851.3402650>
- Zhan, S., & Chong, A. (2021). Building occupancy and energy consumption: Case studies across building types. *Energy and Built Environment*, 2(2), 167–174. <https://doi.org/10.1016/J.ENBENV.2020.08.001>

# AN INTEGRATED PARAMETRIC GENERATION AND COMPUTATIONAL WORKFLOW TO SUPPORT SUSTAINABLE CITY PLANNING

*Evaluate decarbonisation scenarios in the city of Sheffield*

HANG XU<sup>1</sup> and TSUNG-HSIEN WANG<sup>2</sup>

<sup>1,2</sup>*School of Architecture, the University of Sheffield*

<sup>1</sup>*1815191192@qq.com, 0000-0002-4063-6295*

<sup>2</sup>*tsung-hsien.wang@sheffield.ac.uk, 0000-0002-4500-8415*

**Abstract.** To examine how efforts in the built environment can contribute to global climate change mitigation at the urban scale, urban building energy modelling (UBEM) is one of the research areas gaining increasing interest in recent years. However, limited studies systematically illustrate a comprehensive UBEM workflow for most architects and urban planners considering available public datasets, particularly at the early conceptual design phase. In current UBEM studies, major challenges arise from the lack of fine-grained measured urban data and incompatibility between software. To address these challenges and support future sustainable cities and communities, this paper proposed a streamlined computational workflow of UBEM to facilitate sustainable urban design development. Through a case study of Sheffield in the UK, this paper demonstrated an automated and standardised computational workflow that can test the decarbonisation potential in built environments by evaluating energy demand and supply scenarios at an urban scale. This workflow is envisaged to be applicable at various scales of an urban region given an appropriate geographic information system (GIS) dataset.

**Keywords.** Parametric Design Generation; Urban Sustainability; Urban Building Energy Modelling; Building Performance Simulation; Renewable Energy; Decarbonisation; SDG 11.

## 1. Introduction

Half of the world's population now lives in cities, and this is projected to increase to two-thirds by 2050 (UN, 2020). Cities also have a great potential to contribute to global carbon reduction if energy demand falls. To support future sustainable cities and communities, where there's a need for resource efficiency and climate adaptation in urban area, urban building energy modelling (UBEM) is becoming a popular method to develop sustainable strategies for the built environment through estimating urban scale energy demand loads.

Bohringer et al. (2007) listed two approaches to estimate energy use in the built environment: the top-down and bottom-up approaches. The accuracy and reliability of modelling energy use heavily depend on the availability and quality of energy-use data sets and explanatory variables (Abbasabadi and Ashayeri, 2019). Cheng et al. (2019) used a data-driven framework for estimating the building energy use intensity (EUI) in urban regions, supported by a geographic information system (GIS) integrated data. Devila et al. (2016) created an urban-scale energy model for Boston and evaluated PV efficiency in two scenarios, with existing geospatial datasets. Similarly, Alhamwi et al. (2017) developed a GIS-based model to investigate the deployment of renewable energy sources at the city level. Li (2018) listed a series of ArcGis plugins for urban energy analysis and renewable strategies. Krietemeyer and El Kontar (2019) investigated a method for integrating an UBEM with GIS for spatiotemporal visualisation and analysis.

However, existing UBEM tools, such as CitySim (Robinson et al. 2009), are usually built on an independent work platform. Some are developed as plugins that rely on GIS-processing software, such as ArcGis (Li, 2018). For most architects and urban planners, an additional and unfamiliar interface can bring inconvenience into a real work scenario. This gap exists in applying UBEM for sustainable planning within traditional architectural software such as Rhinoceros 3D.

Umi is one existing UBEM technology developed by MIT based on Rhinoceros 3D, which allows neighbourhood-scale simulation of energy demand, daylight and walk score in the building environment (Reinhart et al., 2013). Although useful, it only provides a limited simulation selection. Elk is Grasshopper 3D plugin that allows a wide range of city model generation and can be further developed into UBEM. However, only OSM file is acceptable in ELK, and this limitation brings a series of problems, such as missing geometry due to incomplete GIS data. Ferrando et al. (2020) further concluded that the challenge of UBEM remains on the availability of high-quality data, and in particular, new methods with more robustness need to be developed to tackle GIS-based city building footprints. This project proposes a streamlined and expandable workflow where GIS data is employed to create an urban-scale energy model and facilitate sustainable city planning.

## 2. Methods

As “The Green Heart of Great Britain”, Sheffield is selected as the case study to test the workflow. Rawal (2019) argues that the methodology of creating UBEM should depend on the need of the result, and the level of details that influences the accuracy of the result. For most UBEMs created by the traditional bottom-up approaches, buildings with similar usage, age, and size can be organised into the representative building system (archetype). An archetype database can be created based on local energy codes combined with measured or surveyed data (Chen and Hong, 2018). For UBEMs, the details of building energy profiles are typically generated based on archetypes.

In this study, the challenge arises from the balance between available archetypes data in the Sheffield region and the required details and accuracy needed for energy simulation. Hence, a customised scale-up methodology based on a bottom-up approach is proposed to create an urban energy model. The workflow consists of three Grasshopper 3D modules, as shown in Figure 1, including:

- **Model generation:** This module retrieves both geometry and non-geometry data, such as building types of a GIS model for mass building generation.
- **Performance simulation:** This module calculates the average energy usage intensity (EUI) of representative buildings to predict total energy consumption, including energy demand for heating, domestic hot water (DHW), cooling, lighting and equipment loads.
- **Renewable application:** This module aims at decarbonising the Sheffield region by different sustainable design strategies at an urban scale, including fabric efficiency and solar energy.

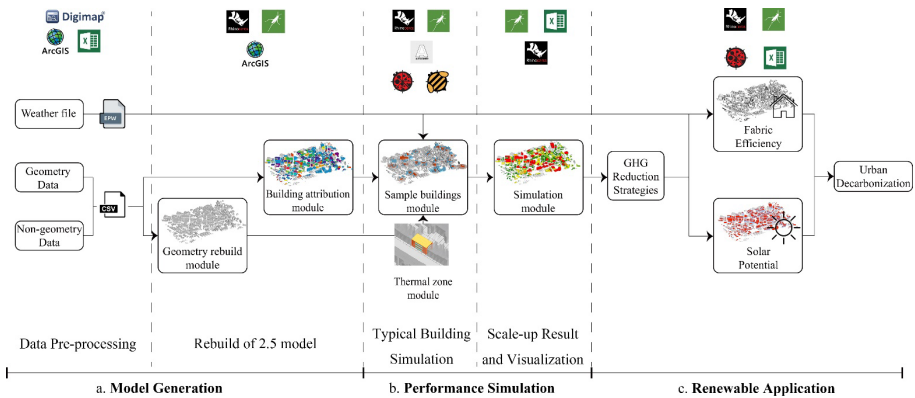


Figure 1. The workflow diagram of three modelling and evaluation modules

### 3. Programming and simulation

#### 3.1. MODEL GENERATION

##### 3.1.1. Rebuild of 2.5D model

In the UK, Digimap is one of the most powerful platforms for architects and urban planners to access to raw spatial data in a wide range of formats (EDINA, 2021). An integrated data package was created, using original data downloaded from Digimap.

A data-driven 2.5D urban model then was created in Rhinoceros 3D. Figure 2 shows that spatial data of grid SK38NE (5\*5km) collected from Digimap was employed to generate a 2.5D urban model. There are nine groups at the first level and 52 categories at the second level, ranging from accommodation to bus transport (EDINA, 2021). We consider the existing building types to synthesise buildings in the targeted city region for simulation. Twelve categories are chosen and compared against the CIBSE-TM46 Energy Benchmarks (Table 1). Figure 2(a) shows all 3D geometry of buildings in grid SK38NE. Figure 2(b) illustrates buildings with twelve

representative building types.

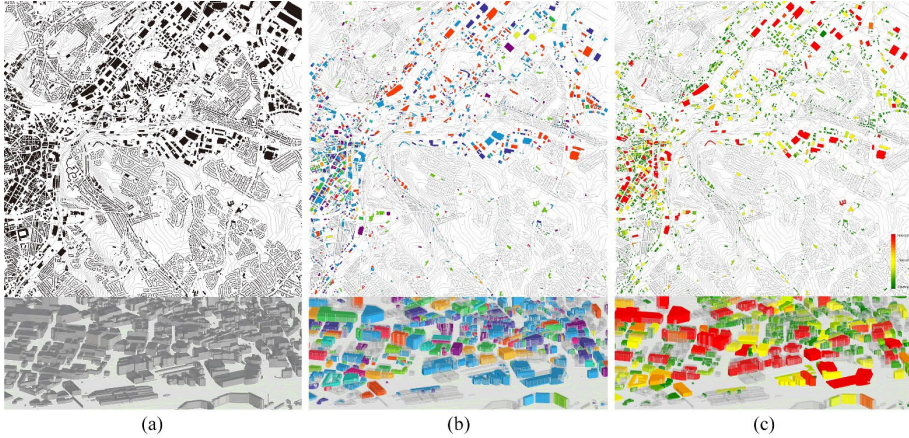


Figure 2. (a). Buildings with a GIS database from DigiMap (b); Buildings shaded by selected building types (c); Visualization map of total energy demand of selected buildings

Category	CIBSE Benchmark Category	Category Code	Operational Schedule	Service included
Accommodation	General accommodation	22	Non-continuous occupancy, often only used in evenings	Heating, lighting, cooling, laundry and drying rooms
Eating and Drinking	Restaurant	7	Wide variety of operational schedules, from selected portions of weekdays to 24/7 operation	Heating, lighting, cooling, food storage, heating of pre-prepared food
Commercial services	Large non-food shop	4	Typical week and weekend days	Heating, lighting, cooling, appliances for small number of employees
Attractions	Cultural activities	10	Daytime use, similar to office hours but more likely to be open in weekends	Heating, lighting, cooling, humidity control
Entertainment	Entertainment halls	11	Mainly in evenings, some daytime use. All days of week	Heating, lighting, cooling of main space
Sports	Dry sports and leisure facility	14	Ranges from occasional use to daily and evening	Heating, lighting and basic office equipment
Education	School and seasonal public buildings	17	Weekday usage for part of the year	Heating, lighting and basic office equipment, teaching equipment, computers
Health	Hospital; clinical and research	20	Continuous for the majority of the facility	All services
Public infrastructure	Public building with light usage	16	Intermittent	Heating, lighting and cooling
Manufacturing and production	Workshop	27	Generally working week but can be multi-shift	Industrial heating and lighting standards
Retail	General retail	3	Weekdays and early evenings, commonly part or all of weekend	Heating, lighting, cooling, appliances for small number of employees
Transport	Terminal	26	Daytime and evenings each day	Heating, lighting, cooling, baggage handlings

Table 1. Customised building attribution according to CIBSE-1TM46 Energy Benchmarks (Source: Field, 2008)

## 3.2. PERFORMANCE SIMULATION

### 3.2.1. Typical building simulation

The average EUI of 12 building types is employed to approximate the energy demand of the whole region. The simulation process includes a thermal zone module that can automatically generate thermal zones for simulation (Figure 3). Chen and Hong (2018) listed three common methods for thermal zoning. The first one, OneZone, can create

on general thermal zone for each floor based on the building footprint. The second method, Prototype, uses empirical data to approximate energy consumption by building type. The third one, AutoZone, automatically divides the building footprint into one, or multiple, core zones and perimeter zones. In this paper, we considered AutoZone to split core and perimeter zones to better simulate dynamic performance of urban buildings.

Specifically, when splitting zones, the perimeter zone typically consists of one-to-multiple fixed-depth spaces along the building boundary with exterior openings. The remaining interior region in the floor plan forms the core zone. Dogan and Reinhart (2016) developed Autozoner to construct perimeter and core zones from unknown interior spaces and deployed this tool as a plugin for automatic zoning in energy simulation. In this study, we create thermal zones for simulation using two-dimensional profiles generated from Autozoner. Building components including walls, floors, windows, and roofs are subsequently specified using Honeybee in grasshopper (Figure 3) to formulate valid energy models for estimating representative building EUI. We randomly selected ten sample buildings per building type to derive the representative EUI.

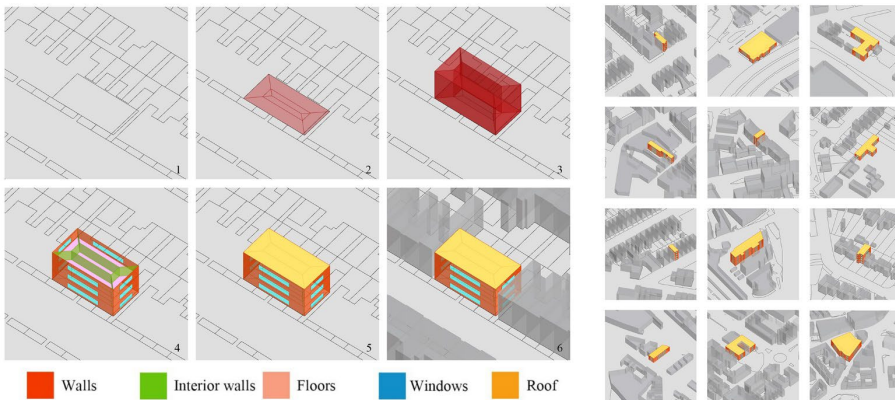


Figure 3. Test for thermal zone module to deal with different sample geometry

### 3.2.2. Scale-up results and visualisation

The final step in this session is to calculate the total energy demand for each building type according to the average EUI, and to generate a demand map to make the result easy to read (Figure 2c). Among 2760 buildings in this region (SK38NE), the ‘commercial services’ area is the largest (2433852.51 square meters), followed by ‘manufacturing and production’ and ‘retail’. Buildings used for ‘attractions’ only accounts for 2% of the total floor area. Total energy demand for 12 building types was calculated accordingly, as summarised in Figure 4.

## 4. Renewable applications

### 4.1. FABRIC EFFICIENCY

The EU Commission ‘2030 climate and energy framework’ sets up three key targets for 2030: at least 40% cut in greenhouse gas emissions (from 1990 level), at least 32% share from renewable energy, and 32.5% improvement in energy efficiency (EC, 2020). Sheffield city region also proposed a target named ‘Net-zero CO2 emission by 2040’, including several goals for the built environment. Among these, quantitative goals are set for building fabric as ‘65000 cavity walls insulated, and 119000 solid walls insulated by 2040’ (SCR, 2021). These goals are turned into three replaceable parameters in the integrated workflow of this study with three specific U-values at different times (1980s, 2010s and 2040s). Figure 5 illustrates the resulting CO2 emission and EUI change of 12 building types.

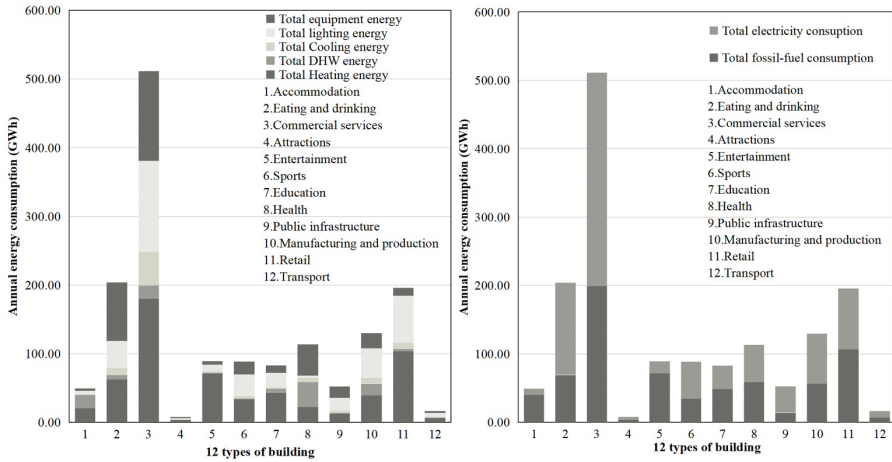


Figure 4. Annual energy demand of 12 types in grid SK38NE

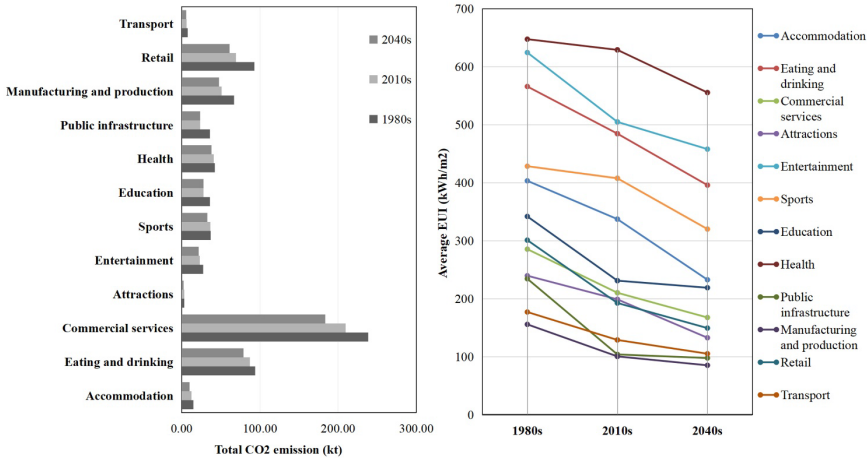


Figure 5. (Left) Total CO2 emission changes of 12 building types from 1980-2040; (Right) EUI change of 12 types from 1980-2040 in grid SK38NE.

4.2. SOLAR POTENTIAL



Two typical days (21 July and 31 December) were selected as the testing scenarios to study the specific energy demand in the Sheffield region. As observed, a peak of energy consumption exists at 7:00 am on a typical winter day, and for a typical summer day, energy demand for all types is higher in the afternoon and climb to the peak at 5:00 pm. In winter, the demand for fossil-fuel is higher all the time, whereas, in summer there is more consumption for cooling (Figure 6).

In Figure 7, simulated building loads with and without Photovoltaic Thermal (PVT) panels generation in this region are shown in dark and light grey for both scenarios. The result of heating demand on 31 December shows a similar profile with that of total energy demand, indicating that fossil-fuel is largely consumed in winter. At 12:00 pm on 31 December, heating provided by PVT can reduce almost half of fuel demand at that time, whereas a bit less for electricity, with a decrease of 25%. On a typical summer day, the heating generated is much higher than demand, and the profile shows zero fuel demand from 5:00 am to 5:00 pm. PVT also has a good electricity generation performance in summer, stupendously reducing electricity demand by 92% at 3:00 pm and slightly less than 30% in peak hours. For equivalent CO2 emission, approximate 50 % of the drop is observed at 12: 00 pm on 31 December, whereas more obvious on 21 July, with a maximum reduction of 87% at 3:00 pm.

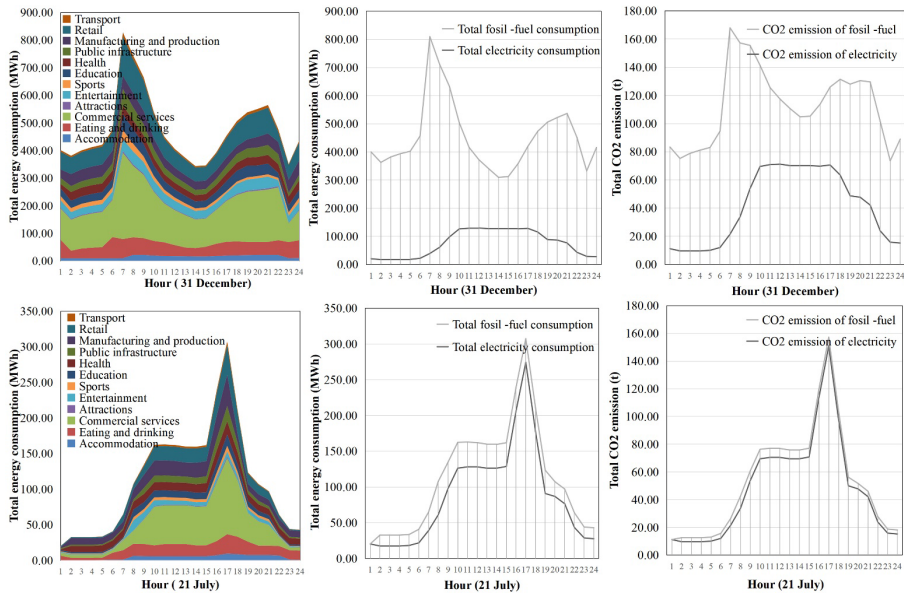


Figure 6. Daily energy demand and equivalent CO2 emission in two scenarios

## 5. Discussion of the Results

- With the implementation of ‘Scale-up methodology’, many steps and modules are optimised with minimal impact on accuracy to save running time. The total operation time of the simulation with 3 different U-value is approximately 6 hours for 120 sample buildings in the selected Sheffield region (30s-60s for one sample

building). For fabric efficiency, improving U-value can significantly decrease the CO<sub>2</sub> intensity of fossil-fuel for all twelve types of building, among which “Accommodation” will have the most reduction of 53% in the 2040s (compared with 1980s). EUI reduction for most types reached a threshold in 2010, indicating that in the future, the benefits of improving fabric performance will not be as efficient as before.

- As tested in Sheffield scenarios, solar energy in this region can significantly improve renewable energy provision. Peak energy generation and carbon reduction are observed at the time of day when solar radiation is at its strongest (44% carbon reduction at 12:00 pm on 31 December and 87% carbon reduction at 3:00 pm on 21 July).

The results indicate that the building environment can significantly impact global carbon emissions. Yet, it requires the concerted efforts of all industries to achieve the ultimate goals for future sustainable cities and communities.

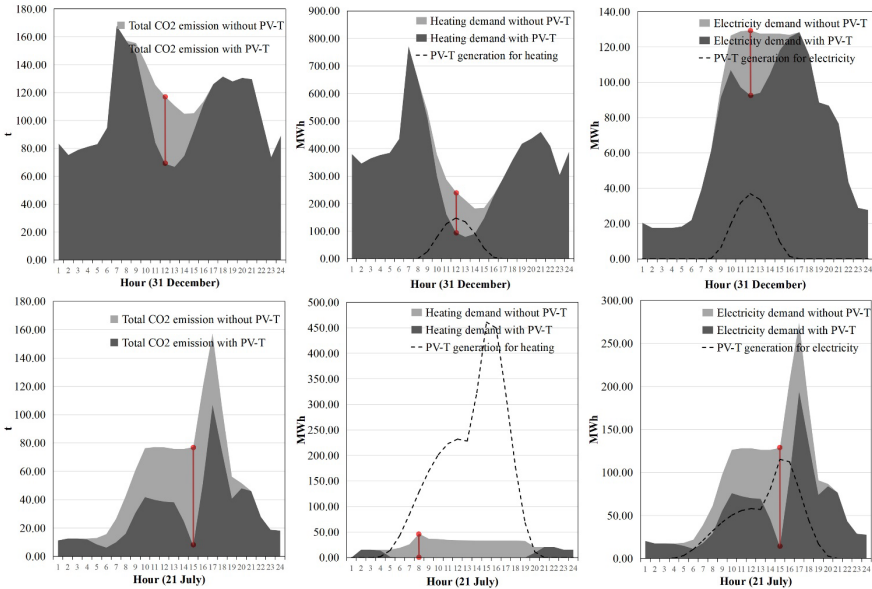


Figure 7. PVT's contribution for decarbonising and energy demand reduction in two scenarios

## 6. Conclusion

This study shows an integrated GIS-based workflow to support sustainable cities and communities from model generation to sustainable design applications at an urban scale. Compared with other UBEMs, it offers advantageous flexibility in resolving information interoperation between various software tools and data formats. With more high-quality sample buildings, the accuracy of simulation results can be easily improved based on the proposed workflow. Scaling up this workflow for any region in the UK can also be achieved given the appropriate input data. Figure 8 shows the implementation of this workflow in a wider range (grid SK38NE, SK38NW, SK39SE,

SK39SW). Further study is to continue smooth integration and expansion with versatile renewable strategies to pursue a more equitable and sustainable built environment at various spatiotemporal scales.

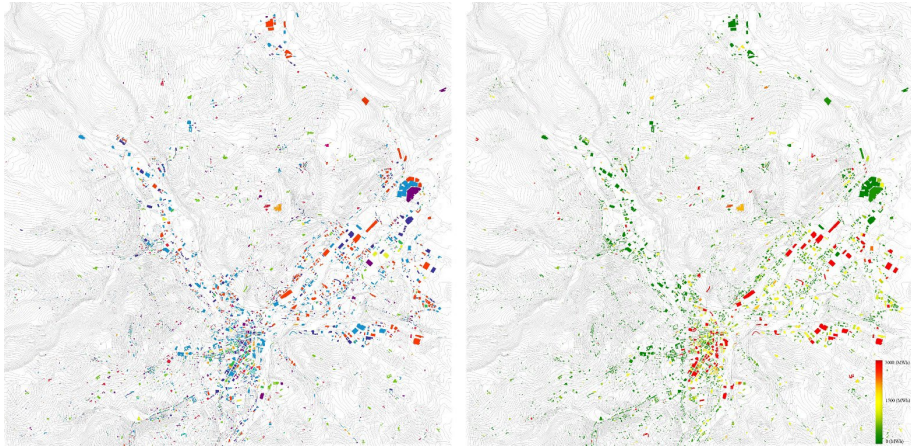


Figure 8. (Left) Building attribution of 4 grids and (Right) energy demand map of 4 grids

## References

- Abbasabadi, N. and Ashayeri, M. (2019). Urban energy use modelling methods and tools: A review and an outlook. *Building and Environment*, 161, 106270. <https://doi.org/10.1016/j.buildenv.2019.106270>
- Alhamwi, A., Medjroubi, W., Vogt, T. and Agert, C. (2017). GIS-based urban energy systems models and tools: Introducing a model for the optimisation of flexibilisation technologies in urban areas. *Applied Energy*, 191, 1-9. <https://doi.org/10.1016/j.apenergy.2017.01.048>
- Chen, Y. and Hong, T. (2018). Impacts of building geometry modelling methods on the simulation results of urban building energy models. *Applied Energy*, 215, 717-735. <https://doi.org/10.1016/j.apenergy.2018.02.073>
- Chen, Y., Hong, T., Luo, X. and Hooper, B. (2019). Development of city buildings dataset for urban building energy modelling. *Energy and Buildings*, 183, 252-265. <https://doi.org/10.1016/j.enbuild.2018.11.008>
- Davila, C.C., Reinhart, C.F., and Bemis, J.L. (2016). Modelling Boston: A workflow for the efficient generation and maintenance of urban building energy models from existing geospatial datasets. *Energy*, 117, 237-250. <https://doi.org/10.1016/j.energy.2016.10.057>
- Dogan, T. and Reinhart, C. (2017). Shoeboxer: An algorithm for abstracted rapid multi-zone urban building energy model generation and simulation. *Energy and Buildings*, 140, 140-153. <https://doi.org/10.1016/j.enbuild.2017.01.030>
- EC. (2021). *2030 Climate Target Plan*. European Commission. Retrieved June 1, 2021, from [https://ec.europa.eu/clima/eu-action/european-green-deal/2030-climate-target-plan\\_en](https://ec.europa.eu/clima/eu-action/european-green-deal/2030-climate-target-plan_en)
- EDINA. (2021). *Digimap*. The University of Edinburgh. Retrieved November 29, 2021, from <https://digimap.edina.ac.uk/>
- Ferrando, M., Causone, F., Hong, T. and Chen, Y. (2020). Urban building energy modeling (UBEM) tools: A state-of-the-art review of bottom-up physics-based approaches. *Sustainable Cities and Society*, 62, 102408. <https://doi.org/10.1016/j.scs.2020.102408>
- Field, J. (2008). *Energy Benchmarks CIBSE TM46: 2008*. CIBSE Publications.

- Krietemeyer, B., & El Kontar, R. (2019, April). A method for integrating an UBEM with GIS for spatiotemporal visualization and analysis. In *Proceedings of the 10th Annual Symposium on Simulation for Architecture and Urban Design Conference* (pp. 87-94). *Simulation for Architecture and Urban Design (SimAUD)*.
- Li, C. (2018). 2.09-GIS for urban energy analysis. *Earth Systems and Environmental Sciences; Elsevier: Oxford, UK*, 187-195.
- Rawal, R., Fennell, P., Ruyssevelt, P. and Poola, V. (2019). Determining the most appropriate form of Urban Building Energy Simulation Model for the city of Ahmedabad. In *Building Simulation 2019*(16). International Building Performance Simulation Association (IBPSA).
- Reinhart, C., Dogan, T., Jakubiec, J. A., Rakha, T., & Sang, A. (2013, August). Umi-an urban simulation environment for building energy use, daylighting, and walkability. In *13th Conference of International Building Performance Simulation Association, Chambéry, France* (Vol. 1, pp. 476-483). International Building Performance Simulation Association (IBPSA).
- Robinson, D., Haldi, F., Leroux, P., Perez, D., Rasheed, A., & Wilke, U. (2009). CitySim: Comprehensive micro-simulation of resource flows for sustainable urban planning. In *Proceedings of the Eleventh International IBPSA Conference* (No. CONF, pp. 1083-1090). International Building Performance Simulation Association (IBPSA).
- SCR. (2021). *SHEFFIELD CITY REGION DRAFT ENERGY STRATEGY*. Sheffield City Region. Retrieved June 1, 2021, from <https://governance.southyorkshire-ca.gov.uk/documents/s2220/Appendix%201%20Final%20Draft%20SCR%20Energy%20Strategy.pdf>
- UN. (2021). *Population*. United Nations. Retrieved June 1, 2021, from <https://www.un.org/en/global-issues/population>

# ENVIRONMENTAL PERFORMANCE ASSESSMENT

## *The Optimisation of High-Rises in Vienna*

SERJOSCHA DUERING<sup>1</sup>, THERESA FINK<sup>2</sup>, ANGELOS CHRONIS<sup>3</sup> and REINHARD KOENIG<sup>4</sup>

<sup>1,2,3,4</sup>*AIT Austrian Institute of Technology.*

<sup>4</sup>*Bauhaus Universität Weimar.*

<sup>1</sup>*serjoscha.duering@ait.ac.at* <sup>2</sup>*theresa.fink@ait.ac.at*

<sup>3</sup>*angelos.chronis@ait.ac.at* <sup>4</sup>*reinhard.koenig@ait.ac.at*

**Abstract.** Our cities are facing different kinds of challenges - in parallel to the urban transformation and densification, climate targets and objectives of decision-makers are on the daily agenda of planning. Therefore, the planning of new neighbourhoods and buildings in high-density areas is complex in many ways. It requires intelligent processes that automate specific aspects of planning and thus enable impact-oriented planning in the early phases. The impacts on environment, economy and society have to be considered for a sustainable planning result in order to make responsible decisions. The objective of this paper is to explore pathways towards a framework for the environmental performance assessment and the optimisation of high-rise buildings with a particular focus on processing large amounts of data in order to derive actionable insights. A development area in the urban centre of Vienna serves as case study to exemplify the potential of automated model generation and applying ML algorithm to accelerate simulation time and extend the design space of possible solutions. As a result, the generated designs are screened on the basis of their performance using a Design Space Exploration approach. The potential for optimisation is evaluated in terms of their environmental impact on the immediate environment.

**Keywords.** Simulation; Prediction and Evaluation; Machine Learning; Computational Modelling; Digital Design; High-Rise; SDG 11, SDG 13.

## 1. Introduction

To tackle the challenges of sustainable cities and climate action, environmental performance of planning decisions needs to be considered as they carry a significant potential for improving the quality of urban space. This holds special relevance for high-rises, as they have a major impact on their surroundings: for instance, in terms of shading and wind circulation.

An informed planning environment should make the impacts on the environment transparent and assess them in order to be used as a decision-making support for climate-resilient planning. Yet, environmental impact assessment is seldomly applied in practice within the design process. If so, at the very end of a project, where it is too late and costly to conduct major design revisions. This is caused by long design evaluation cycles and high costs due to the complexity of simulation engines. Moreover, simulation times of hours or even days make it infeasible to systematically test and optimize designs, not to speak about conducting design-space exploration. Therefore, we propose this framework for environmental assessment and optimisation of high-rise planning in an urban context.

## 2. State of the art

### 2.1 COMPUTATIONAL URBAN DESIGN

The historical evolution of Computational Design was discussed by Caetano et. al whereas the first definition by Moretti (1971) describes parametric design with “the relationships between the dimensions” (Caetano et al., 2019). The potential of using computational design in the urban planning context was investigated and its use as an adequate tool was explored (Mukkavaara et al., 2020). The approach of computational design methods has been used more and more often, especially in the early design phase by architects and planners, when there is still a larger possibility space of different designs (Caetano et. al, 2020)

A commonly used software environment is Rhino 3D (McNeel) and its native plug-in Grasshopper. This set-up contains a wide range of freely available open-source plug-ins for simulating and evaluating design scenarios. In addition to the simulation of the statics of buildings, environmental simulation can also be carried out. The interfaces for GIS data import (eg. .geojson) enable the georeferencing of projects and the integration of open available urban databases.

### 2.2 PERFORMANCE BASED DESIGN / PREDICTION

The consideration of thermal comfort becomes a criterion for public life, especially in densely populated urban areas (Zhang, Liu, 2021). For this reason, the performance-based design approach has been studied and applied on different scales for the analysis of climatic and energetic aspects, especially in recent years. In particular, the combination of Parametric Design and Design Space exploration can provide insights for impact-oriented planning (Fink et al., 2020; Abdulmawla et al., 2018).

Physic based simulations of Computational Fluid Dynamics (CFD) are time and resource consuming and thus usually only carried out for a few shortlisted or final building designs. However, novel machine learning algorithms were successfully applied to the domain of CFD simulations. Although less accurate, they are several orders of magnitude faster compared to their traditional counterparts - by a factor 5000 to 15,000 with an accuracy of 80 to 95% compared to physic-based simulations (see Chronis et al., 2020; Duering et al., 2020).

### 3. Methodology

#### 3.1 OVERVIEW

The fundamental idea of this framework is to create a parametric model to evaluate and optimise many design scenarios for their environmental impact and potentially derive insights on systematic relationships between form and environmental performance. Hence the generation of 5.000 variations with 8 building types at different heights and rotations was carried out within the software Rhinoceros 3D and its plug-in Grasshopper. Each variant was evaluated on its impact on wind comfort, solar exposure of buildings and shading of open spaces and the surrounding buildings. A key challenge arising from this approach is to structure and make sense of the thousands of data points in order to arrive to actionable insights. A significant part of this research was therefore devoted to developing a strong design and performance space exploration and analysis dashboard.

For the purpose of this experiment, we choose a mixed-use site in Vienna, Austria and formulated three main design objectives:

- First, the shading impact of the design on the context should be reduced to the minimum
- Second, a high thermal comfort is aimed on the open space in the surrounding of the buildings
- Third, tower residents and neighbours shall perceive a sufficient amount of sun

Given the selected impact assessment models (see fig. 1), these objectives translate into the following strategies of 1) reduce the amount of added shading to surrounding public spaces and buildings 2) aim for low wind speeds in public spaces. This particularly holds relevance in the case of high-rise developments where downwind and wind tunnel effects can lead to discomfort and even affect the safety of pedestrians (Reiter, 2010). Finally (3), The facades of the towers shall be exposed to higher levels of solar radiation.

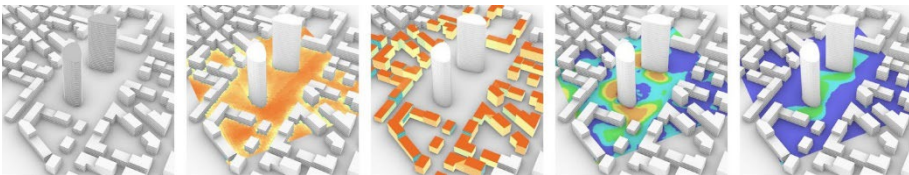


Figure 1. 3D model and simulation results (from left to right: Massing model; shading of public spaces and buildings; wind comfort for a stormy and average day).

### 3.1.1. Key performance indicators

Next, we defined seven Key Performance Indicators (KPI) to measure the level of success of the mentioned goals and strategies, namely:

- shadingGround - percentage change of shaded areas compared to the status-quo without towers
- shadingContext - percentage change of shaded facades of surrounding buildings compared to the status-quo
- shadingT1/T2 - percentage of the façade surfaces of tower one and two that receives solar radiation of more than 600 kwh per year
- S1\_dangerous - percentage of public spaces that are exposed to wind speeds  $>15\text{m/s}$  at a stormy day (wind speed of  $20\text{m/s}$ , from  $270^\circ$ )
- S1\_uncomfortable - percentage of public spaces that are exposed to wind speeds  $10 - 15\text{m/s}$  at a stormy day (wind speed of  $20\text{m/s}$ , from  $270^\circ$ )
- S0\_SittingLong - percentage of public spaces that are exposed to wind speeds of less than  $2.5\text{m/s}$  at an average day (wind speed of  $5\text{m/s}$ , from  $135^\circ$ )

## 3.2 HIGH-RISE BUILDINGS

The 3D models for the high-rise configuration of the case study were generated using eight building types (see fig 2). Their shape and relation to the site was further altered by the two parameters height and rotation. The parametric set-up was developed in Grasshopper. By optimising the shape, the improvement of the quality of stay in the outdoor area can be achieved. The two high-rise buildings were varied accordingly throughout design scenarios, while the context buildings always remained the same. The defined input parameters were stored in a database for the 5.000 scenarios and were used for the individual iterations using the colibri plug-in.

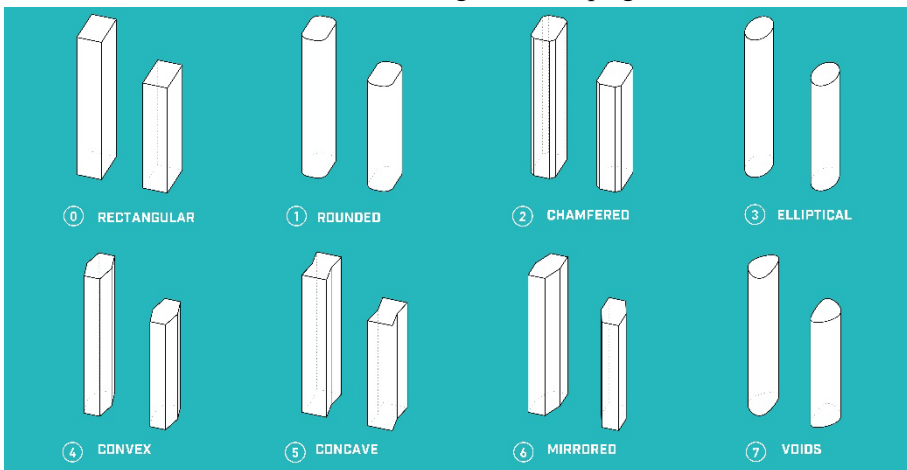


Figure 2. Building shapes that were used for the scenario generation.



### 3.3 PERFORMANCE SIMULATION

As a next step, the generated design scenarios were evaluated based on their environmental performance. For the evaluation of the individual designs, an automated process for the simulation of wind and sun was set up.

#### 3.3.1. *Wind Comfort*

The wind speed at pedestrian level is strongly dependent on the building configurations and the orientation of streets. The general conditions are formed by the local climate, with both the shape and orientation of buildings showing a significant effect on wind conditions (Notado et al., 2021; Gunawardena et al., 2017). In this study we use a model developed by the City Intelligence Lab (AIT), that was particularly trained to predict pedestrian wind comfort on the urban scale (more specific: the wind speeds experienced on a 1.75m height, given wind speed and direction). Machine learning algorithms were used to predict wind comfort based on wind speed and direction. The model speeds up the process by the factor of 10.000, while sustaining an accuracy of 80 - 95% in comparison to traditional Computational Fluid Dynamics (CFD) simulations (see Kabošová et al., 2021; Galanos et al., 2019). It was trained on 5.500 cases from cities around the globe using the industry standard OpenFoam.

The application of a trained neural networks allows to predict the wind comfort at a 1.75m height based on the wind speed and wind direction as inputs. This prediction is carried out within a few seconds per design variant and supports the evaluation of a great design space.

#### 3.3.2. *Solar Radiation*

Another driving factor of the environmental evaluation of design options is the solar intake of building facades. The shading of the open space and the context buildings is evaluated to ensure a high quality for the direct neighbourhood of the high-rise buildings. Solar radiation was simulated using Ladybug tools in the Rhino 3D/Grasshopper software environment. Both the solar radiation on the open spaces around the high-rise buildings and the radiation on the facades in the immediate context were calculated (see fig. 3). The analysis grid was defined with a size of 4 meter and

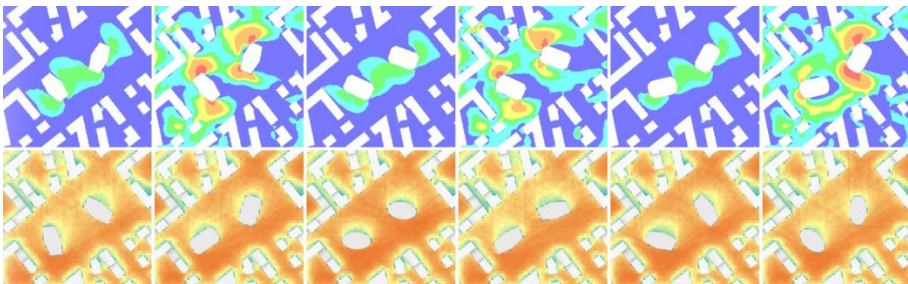


Figure 3. Overview of the solar radiation and wind comfort analyses.

the simulations were carried out on basis of weather data for the inner city of Vienna

and evaluated for the period of one year. The evaluation of the solar irradiation was determined in the percentage change in the shading of the open spaces and façades.

### 3.4 KPI-BASED EVALUATION OF SCENARIOS

For the evaluation of the simulation results we developed an interactive dashboard to analyse and visualize data (see fig. 4). Its main functions it to 1) group or sub-sample data (eg. through a parallel coordinate plot or selections on dimensionality reduced plotted data) and 2) derive insights by describing, comparing and correlating different sub-samples with one another.

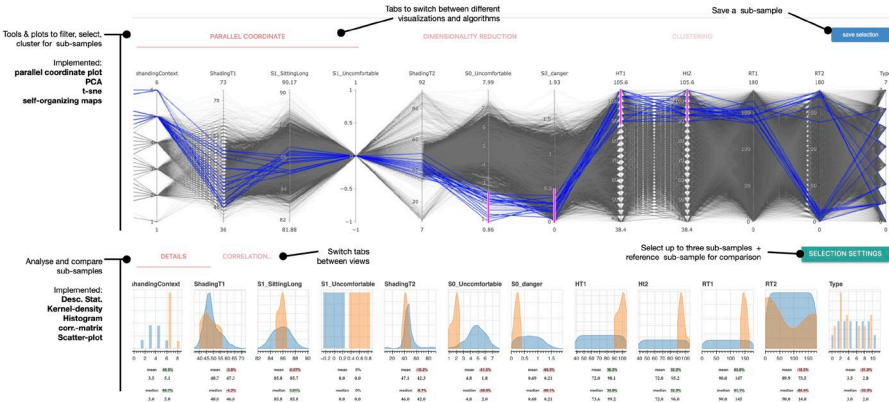


Figure 4. Interactive dashboard for design exploration. The upper section offers multiple ways to define sub-samples in combination with manual selections on plots: parallel coordinate plot, principal-component analysis, t-sne, self-organizing maps and k-means clustering. The lower section shows key statistics and visualizations of selected sub-samples including histograms, kernel-density plots, correlation-matrices and scatter plots.

For instance, by filtering for all design variants with good wind comfort or comparing the eight tower shapes and their KPI performance to one another. The dashboard was implemented with Python and JavaScript, heavily relying on the visualization library Bokeh.

Figure 5 shows an overview of the complete dataset including distribution plots for each KPI and design parameter along with descriptive statistics. All KPIs show high variances across the design variants indicating the significance of conducting form optimization. For instance, values for shadingContext differ between 1 and 6% (increased shading compared to the status-quo without towers) and S0\_danger and S0\_uncomfortable ranging between 0 and 1.9% and 0.9 to 8 which translates to several hundreds of square-meters of difference. For instance, values for shadingContext differ between 1 and 6% and S0\_danger and S0\_uncomfortable ranging between 0 and 1.9% and 0.9 to 8% which translates to several hundreds of square meters of additional uncomfortable or even problematic areas on the site. Moreover, most KPI show clear patterns of being normally distributed. For S0\_danger this is less so the case, indicating a relative even distribution of values between 0 and 1.2% followed by a sharp decrease in design variants with values over 1.5%.



Figure 5. Overview of the complete dataset. Top: all data points plotted in 2D using the t-sne algorithm. Lower section: Descriptive statistics and value distributions for the current selection made on the plot.

A look on the correlation matrix (simple linear regression between all columns of the data table) mostly reveals expected results such as high positive correlations between building heights and shading (H1, H2, shadingContext and shadingGround). Moreover, the plot suggests a high negative correlation between the RT1 (rotation of tower 1) and S0\_uncomfortable (wind comfort), a closer inspection of the scatterplot (fig. 6) however illustrates a non-linear underlying distribution that would be in need of transformation measures in order to provide meaningful results. This case highlights one limitation of the current state of this analysis tool.

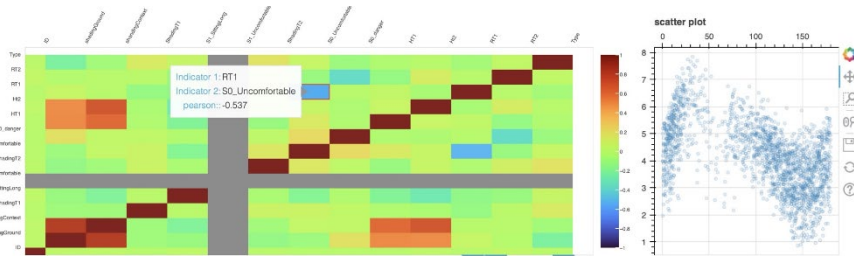


Figure 6. Correlation matrix and scatterplot of RT1 and S0\_uncomfortable.

To better understand the data and potential relationships between form parameter and environmental performance, we created 10 sub-samples following two approaches: Filter for form and filter for environmental performance. First, we created a selection for each of the eight building types. Next, we filtered for all design variants that were among the top 30% in each of the KPI, among the top 30% of shading related KPI and among the top 30% of wind-comfort related KPI (see figure 7 for more detailed explanation and numbers).

Looking at the upper part of the table, tower type 3, 4 and 6 can be found considerably more frequently (ca. 27%, 29%, 23%) in the top performing sample of shading related KPI. Type 4 (10%) and especially 5 (65%) are most frequently found in the sample with the best wind comfort and type 4 (24%), 5 (21%) and 6 (32%) in the top 30% sample across all KPI. Besides, (not surprisingly) tower heights showed to have significant impact on both shading and wind comfort while rotation had an especially strong impact on the latter.

The lower section of the table draws a more nuanced picture, as it depicts the difference in the mean for each KPI by tower type compared to the full sample, containing all design variants. It illustrates the difficulty to clearly declare any "winning" or "optimal" design variant, especially when there is no clear preference or weighing across the different KPI. For instance, if the goal of reducing dangerous wind speeds is valued more important than reducing the amounts of uncomfortable areas during a storm day, tower type 7 (void) tends to perform better (by 13%) compared to type 5 (4.4%). A similar pattern for type 7 can be observed for shading patters (or solar intake) to the tower's facades (performs worse than other types) and shadow casting on its surrounding (strong performance).

Moreover, it is important to mention that we were looking at mean values so far and not individual design variants. Thus, well performing variants could also exist as outliers within the dataset. In Fact, the single best performing variant was found within tower type 7.

Sub-Samples	T0	T1	T2	T3	T4	T5	T6	T7
Top performance selection (% of tower types)	5.9.	11.8.	5.9.	0	23.5.	20.6.	32.4	0
Top performance selection: shading (% of tower types)	0	14.3.	0	27	28.6.	0	22.6.	7.1.
Top performance selection: wind comfort (% of tower types)	7.5.	2.5.	7.5.	0	10	65	7.5.	0
KPI	T0	T1	T2	T3	T4	T5	T6	T7
S0_danger (mean)	2,9 %	0,9	-1,7	2,2	-8,6	4,4	-5,2	13
S0_Uncomfortable (mean)	8 %	-2	0,4	-8	-0,78	6,7	1,9	-4,8
S1_SittingLong (mean)	0.7%	-0,2	0,2	-0,8	0,16	1	0,1	-0,7
ShadingT1 (mean)	0	2,1	0	0,5	0,25	0,3	0,3	-10
ShadingT2 (mean)	-2	4,5	0	-0,6	1	2,8	0,5	-3,8
ShadingContext (mean)	-12 %	-0,1	-3,5 %	12	2,7	-10	-0,4	11,2
ShadingGround (mean)	-11	-0,5	-4 %	10	3,3	-8	0,3	9,6

Figure 7. Comparing sub-samples. Top section: subsample of design variants that are under the best 30% of all KPI, shading related KPI and wind comfort related KPI. The values depict the share of each tower type among the sample in %. Lower section: Tower type (as sub-sample) and difference of mean to the mean of the overall sample in %, for each KPI. Positive means better. Cells are color-coded for better readability.

#### 4. Research Results

In this paper we presented a parametric model to test through 5.000 design variants for a hypothetical high-rise development in the city of Vienna. Next to a range of eight basic footprint types (eg. Rectangular, round), the towers were altered in their height and rotation. Each variant was evaluated upon wind comfort, solar radiation and shadow casting using Vienna's climate and weather data. An interactive data dashboard was used to analyse the results of this experiment, with the goal to support design decision making and potentially uncover generalizable insights.

We found significant differences in the performance across design variants in all seven KPI. In an extreme case this would mean that by altering rotation and tower shape, the share of public spaces subject to low wind comfort (S1\_uncomfortable) could be reduced from 7.8% to 1.3%.

A closer look into the data showed that tower shapes 4,5 and 6 (Convex, Concave, Mirrored, see fig. 2) tend to deliver especially good performance across the studied KPI most reliably. However, this does not mean even better, or best performing design variant can't be found within the other studied shapes. For instance, shape type 7 (void) had very strong performing variants especially regarding wind comfort during stormy days (S0\_dangerous) and in terms of shading of the surrounding.

All design parameter (shape, rotation, and height) showed clear impact on performance, however given the manifold possible combinations and the interaction effects between the two towers (height and rotation were altered individually) and their context makes it hard to derive clear and linear design directions. Instead, this suggests that projects need to be approach in a case-by-case fashion.

Nevertheless, having an automatically generated database of project specific design variants provides the benefit of being able to derive design directions (eg. which combination of tower rotations tend to work best) and the necessary data to evaluate eg. manual designs in terms of the performance potential of a site (how many percentages of the generated designs does one outperform?). In order to capture the full potential of such data, a close integration of capable and, preferable interactive, multi-criteria evaluation tools as well as rapid enough performance simulations are needed. Two of which this paper hopes to have contributed to.

#### 5. Future Prospects

The presented framework can be flexibly extended by KPIs in the fields of energy, mobility and economy to facilitate the decision making and negotiation processes of planners and stakeholders. Due to the tremendous time savings in applying machine learning to simulation, more holistic impact assessments can be conducted, thus facilitating more informed decision-making. Furthermore, the impact of planning decisions could be integrated with movement patterns and socio-demographic data.

As for the next step we still see great potential to extend the capabilities of the design space exploration tool and most importantly to improve its integration with real-time and data driven design sessions. To combine the benefits of large, pre-generated databases with real-time experiments and discussions among stakeholders.

## References

- Abdulmawla, A., Schneider, S. and Bielik, M. (2018) Integrated Data Analysis for Parametric Design Environment, *Proceedings of eCAADe 2018*, Lodz (pp. 319- 326).
- Caetano, L. Santos, and A. Leitão. (2020) Computational design in architecture: Defining parametric, generative, and algorithmic design. *Frontiers of Architectural Research*. 9(2), 287–300. <https://doi.org/10.1016/j.foar.2019.12.008>.
- Duering, S., Chronis, A., and Koenig, R. (2020) Optimizing Urban Systems: Integrated optimization of spatial configurations. (p. 8). *Symposium on Simulation for Architecture and Urban Design 2020* (SIMAUD)
- Fink, T., Vuckovic, M., Petkova, A. (2021) KPI-driven parametric Design of Urban Systems. In 26th International Conference of the Association for Computer-Aided Architectural Design Research in Asia 2 (pp 579-588). *The Association for Computer-Aided Architectural Design Research in Asia* (CAADRIA).
- Elshani, D., Koenig, R., Duering, S., Schneider, S., Chronis, A. (2021) Measuring sustainability and urban data operationalization. In 26th International Conference of the Association for Computer-Aided Architectural Design Research in Asia 2 (pp 407-416). *The Association for Computer-Aided Architectural Design Research in Asia* (CAADRIA).
- Kabošová, L., Chronis, A., Galanos, T., Katunský, D. (2021) Leveraging urban configurations for achieving wind comfort in cities. *SIGraDi 2021*. <https://www.researchgate.net/publication/356893387>.
- Galanos, T. Chronis, A. Vesely, O. Aichinger, A. & Koenig, R. (2019) Best of both worlds – using computational design and deep learning for real-time urban performance evaluation. *1st International Conference on Optimization Driven Architectural Design*.
- Gunawardena, T., Fernando, S., Mendis, P. and Waduge, B. (2017) WIND ANALYSIS AND DESIGN OF TALL BUILDINGS, THE STATE OF THE ART. In *8th International Conference on Structural Engineering and Construction Management*, 2017 (pp. 10).
- Isola, P., Zhu, J., Zhou, T., & Efros, A.A. (2016). Image-to-Image Translation with Conditional Adversarial Networks. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 5967-5976.
- Mukkavaara, J.; Sandberg, M. (2020). Architectural Design Exploration Using Generative Design: Framework Development and Case Study of a Residential Block. *Buildings*, 10, 201. <https://doi.org/10.3390/buildings10110201>.
- Notado, C., Yogiama, C., Tracy K. (2021) Towards Wind-Induced Architectural Systematization. In 26th International Conference of the Association for Computer-Aided Architectural Design Research in Asia 2 (pp 447-456). *The Association for Computer-Aided Architectural Design Research in Asia* (CAADRIA).
- Reiter, S. (2010) Assessing wind comfort in urban planning. *Environment and Planning B: Planning and Design*. 37, (pp 857-873).
- Zhang, Y.; Liu, C. (2021) Digital Simulation for Buildings' Outdoor Thermal Comfort in Urban Neighborhoods. *Buildings 2021*, (11) 541. <https://doi.org/10.3390/buildings11110541>.

# AUTOMATED SEMANTIC SWOT ANALYSIS FOR CITY PLANNING TARGETS

*Data-driven solar energy potential evaluations for building plots in Singapore*

AYDA GRISIUTE<sup>1</sup>, ZHONGMING SHI<sup>2</sup>, ARKADIUSZ CHADZYNSKI<sup>3</sup>, HEIDI SILVENNOINEN<sup>4</sup>, AUREL VON RICHTHOFEN<sup>5</sup> and PIETER HERTHOGS<sup>6</sup>

<sup>1,2,4,5,6</sup> *Singapore-ETH Centre, Future Cities Lab Global Programme, CREATE campus, 1 CREATE Way, #06-01 CREATE Tower, Singapore 138602.*

<sup>3</sup> *CARES, Cambridge Centre for Advanced Research and Education in Singapore, Singapore.*

<sup>5</sup> *Arup, Berlin, Germany.*

<sup>1</sup>*ayda.grisiute@sec.ethz.ch, 0000-0001-7328-847X*

<sup>2,6</sup>*{shi, herthogs}@arch.ethz.ch, 0000-0002-3744-1649, 0000-0003-2562-8374*

<sup>3</sup>*arkadiusz.chadzynski@cares.cam.ac.uk, 0000-0001-8084-9474*

<sup>4</sup>*heidi.silvennoinen@sec.ethz.ch*

<sup>5</sup>*Aurel.von-Richthofen@arup.com, 0000-0002-8712-1301*

**Abstract.** Singapore's urban planning and management is cross-domain in nature and need to be assessed using multi-domain indicators — such as SDGs. However, urban planning processes are often confronted with data interoperability issues. In this paper, we demonstrate how a Semantic Web Technology-based approach combined with a SWOT analysis framework can be used to develop an architecture for automated multi-domain evaluations of SDG-related planning targets. This paper describes an automated process of storing heterogeneous data in a semantic data store, deriving planning metrics and integrating a SWOT framework for the multi-domain evaluation of on-site solar energy potential across plots in Singapore. Our goal is to form the basis for a more comprehensive planning support tool that is based on a reciprocal relationship between innovations in SWT and a versatile SWOT framework. The presented approach has many potential applications beyond the presented energy potential evaluation.

**Keywords.** Semantic Web; Knowledge Graphs; SWOT Analysis; Energy-driven Urban Design; SDG 7; SDG 11.

## 1. Introduction

In Singapore, urban planning and management are done by multiple government

agencies, implying that it is cross-domain in nature. This nature impacts how we measure, evaluate and understand the challenges the city is facing, such as climate change, the liveability of cities, or urban energy demand (Richthofen, 2022). Sustainable Development Goals (SDGs) is an effort to address these challenges and can support planners in comparing, monitoring and measuring the progress of planning strategies through the evaluation of diverse indicators (Massaro et al., 2020). However, such actions require the integration of heterogeneous data and flexible frameworks that enable a comprehensive evaluation of diverse SDG-related planning metrics. While relevant data have been made available through active digitalisation and improved storage capacities (Winkelhake, 2018), existing urban data formats and modelling techniques — such as GIS — remain relatively static and siloed. They lack standardised exchange formats, making it difficult to cross-reference, analyse or track changes (Chadzynski et al., 2021) and impede the integration of (big) data in planning tasks. Therefore, sustainable digitisation and innovation in city planning and knowledge management are essential for achieving SDGs (Massaro et al., 2020).

There is a need for more holistic analyses relating to SDGs that enable the evaluation of cross-domain urban indicators and metrics. Semantic Web Technologies (SWT) is a promising approach to address data fragmentation and improve interoperability by linking data and making it more connected and discoverable through Knowledge Graph (KG) platforms (Chadzynski et al., 2021). In addition, SWOT analysis framework (Strengths, Weaknesses, Opportunities and Threats) has proven to be a versatile and effective tool for strategic urban planning and the evaluation of cross-domain metrics for particular planning targets.

In this paper, we introduce an approach to automate the evaluation of multi-domain targets of city planning using KG, enabling flexible interaction with heterogeneous data and addressing interoperability issues hindering core aspects of sustainable digitalisation in city planning. We present the first steps towards a more comprehensive decision support tool, integrating a dynamic geospatial KG platform with the SWOT framework, resulting in automated multi-domain evaluations of SDG-related planning targets. The remainder of the paper is structured as follows. Section 2 presents the broader research scope and background, how SWOT framework and SWTs address the need for holistic urban analysis, and our use case — a SWOT analysis assessing on-site solar energy potential of plots across Singapore. Section 3 describes our methods for storing heterogeneous data in a KG, deriving planning metrics, and integrating it into the SWOT analysis. Section 4 presents the SWOT analysis results and potential benefits for planners, while Section 5 discusses the main contributions of SWOT and SWT integration, discovered limitations and future work.

## **2. Background**

### **2.1. CITIES KNOWLEDGE GRAPH**

Our work was done as a part of the Cities Knowledge Graph (CKG) research project — an example of a Semantic City Planning System (Richthofen et al., 2022). CKG is a pilot for a comprehensive knowledge management platform that applies SWT and KGs to address interoperability challenges between different urban knowledge domains in order to digitally support the synthesis actions at the core of city planning (e.g. building



consensus among stakeholders), hence addressing SDG 11 — Sustainable Cities and Communities — targets.

The CKG uses SWT to store data on the Web, link it using ontologies, and write rules for handling data across applications. Ontologies account for shared formal vocabularies of domain concepts, instances, and relations (Akroyd et al., 2021) and enable computers to infer semantic relationships between heterogeneous data. KGs express data as a directed graph with concepts or instances as nodes and relations as edges, and can be used to formalise domain knowledge, such as urban energy, mobility or planning (Shi, et al., 2021), into semantic models. Finally, KGs hosted in graph databases, often referred to as triple stores with query endpoints, enable exploration and interaction with such semantic models by executing SPARQL query statements (W3C, 2013).

The CKG is a subsystem of a broader collaborative research effort to develop a general, all-encompassing and dynamic knowledge graph called The World Avatar (TWA) — TWA consists of multi-domain knowledge representation and an ecosystem of autonomous computational agents operating on it (Akroyd et al., 2021). The agents can support planning tasks by dynamically reconfiguring the KG architecture and importing, exporting, updating and analysing urban data.

## 2.2. SWOT ANALYSIS FRAMEWORK

A SWOT analysis is a cognitive process defining and evaluating the interrelations between different factors of an entity or phenomenon across four categorical descriptors: Strengths, Weaknesses, Opportunities, and Threats (Learned et al., 1965). Ghazinoory et al. (2011) associate Strengths and Weaknesses with internal factors — variables that are a part of the entity, and Opportunities and Threats with external factors — variables that are external but can still influence the entity. In the context of future-oriented tasks such as planning, we also consider the temporal dimension in SWOT where SW represent present conditions and OT future ones.

Although originally used in the business field, this framework has recently been rigorously applied in quantitative indicator-based evaluations in the planning domain (Comino & Ferretti, 2016; Ervural et al., 2018; White et al., 2015). In these examples, indicators are categorised into four descriptors based on stakeholder workshops or literature reviews and prioritised using hierarchical weighting methods by domain experts. Additionally, the SWOT framework is also suited for geospatial analyses. Camino and Ferretti (2016) integrated GIS in their indicator-based SWOT analysis for parks to create thematic maps visualising the distribution and extent of vulnerable, threatened or valuable park areas, informing regional park management strategies for optimal environment preservation.

## 2.3. USE CASE — ON-SITE SOLAR ENERGY POTENTIAL

In this paper, we combine the SWOT framework with our CKG to automatically analyse the on-site solar potential of building plots. As cities are one of the main energy consumers worldwide, such an analysis responds to the need to consider clean energy systems already during the master-planning stages and is an established planning target (Shi et al., 2017) for addressing SDG 7 — Affordable and Clean Energy.

An evaluation of a plot's solar energy potential is a transdisciplinary effort and involves a variety of cross-domain metrics and factors, covering building typology, land use, urban form, cooling demand and energy integration costs resulting in a variety of methods (Shi et al., 2021). Using methods that only evaluate individual metrics is not enough to assess solar energy potential due to interaction effects between features like site coverage, land-use ratio and orientation (Shi et al., 2021) and accounting for these interactions is difficult. In comparison, our approach expresses these interactions through linked data.



*Figure 1 shows four chosen Key Growth Areas in Singapore, that aims to showcase integrated master-planning would benefit from solar energy potential considerations at early planning phases.*

This paper builds on previous work (Grisiute et al., 2021), presenting an improved version of that automated SWOT analysis (improved metrics and score aggregation), now applied to diverse regions in Singapore. Singapore's 2019 master plan identifies several regions as Key Growth Areas (URA, 2021), with the aim to showcase how integrated master planning and technology could help create more liveable and sustainable cities in line with SDGs, including the integration of clean energy. Solar energy is the most feasible renewable energy source in Singapore, and will hence require the installation of more photovoltaic panels (Energy Market Authority, 2017). In order to do so, solar energy potential studies at early planning and urban design stages in these regions could prove to be very beneficial. As an experimental use case, we tested our SWT-based SWOT analysis in four of Singapore's Key Growth Areas: Punggol Digital District, Central Business District, Paya Lebar Air Base District and Jurong Lake District (Figure 1).

### 3. Methodology

#### 3.1. STORING DATA IN THE CKG

We used open datasets from Singaporean government agencies (Data.gov.sg, 2021) and other planning-related datasets covering plot form, built form and land-use characteristics. First, we transformed these heterogeneous data to CityGML format with the Feature Manipulation Engine (FME). Second, we used our current CKG architecture — described by Chadzynski et al. (2021) — to transform flat geospatial

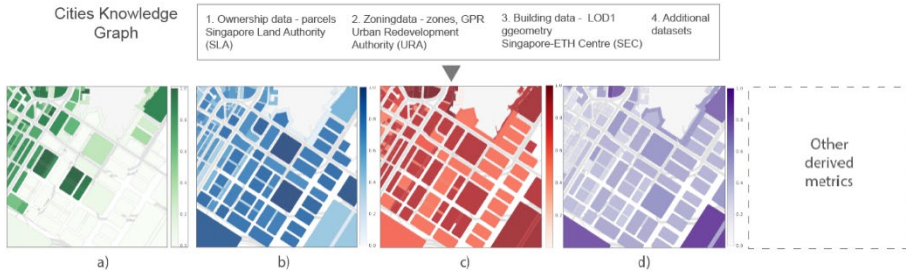


Figure 2 shows relevant metrics derived from datasets stored in the KG. Thematic gradient maps visualise each metric’s value from highest (dark) to lowest (light): a) site coverage indicates to what extent the plot is covered by roof surface, which in Singapore is the most efficient for PV installation; b) elongation as measure for compactness that allows for efficient deployment of PV infrastructure; c) shading factor indicates how much buildings on plot are taller than its surroundings, which has direct effect on inter-intra shading; d) orientation metric indicates how much each plot deviates from recommended orientation for façade PV.

CityGML files into semantic triples and store them in the CKG. This graph database was structured using an ontology of the CityGML 2.0 standard called OntoCityGML (Chadzynski et al., 2021). Finally, the geospatial processing capabilities of the Blazegraph graph database instance used as the triple store enabled us to perform geospatial queries (Blazegraph, 2020).

### 3.2. DERIVED METRICS

Our SWOT analysis required metrics that measure particular characteristics of on-site solar energy potential as inputs. We derived these by executing manifold combinations of SPARQL queries against the CKG to retrieve each dataset’s geospatial information (polygon geometries) and semantic information (zoning type, Gross Plot Ratio (GPR), building height). We linked the semantic information across datasets by comparing topological relations between the retrieved geometries. We then derived composite metrics that have a varying impact on on-site solar energy potential (Shi et al., 2021), like site coverage, land-use ratio, inter-intra shading factor and retrieved other relevant metrics, such as plot area, elongation and orientation, directly from the geometries (Figure 2).

### 3.3. SWOT ANALYSIS FRAMEWORK INTEGRATION

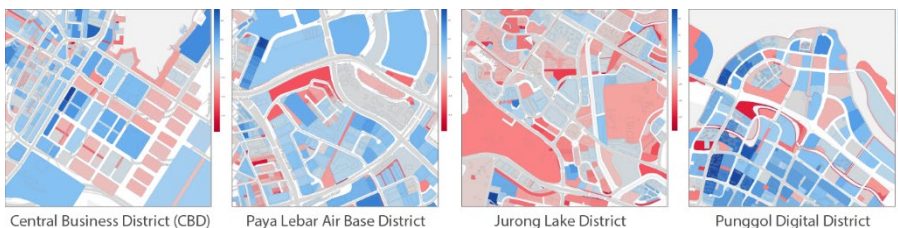
We included five of these metrics — site coverage, land-use ratio, plot orientation, elongation and shading factor — in our proof of concept evaluation of the solar energy potential of large numbers of building plots. As site coverage, plot orientation and elongation define the entity and are considered internal factors, their values were categorised as Strengths or Weaknesses (see also Section 2.2). Meanwhile, land-use ratio and shading factor values were linked to Opportunities and Threats, as these metrics define external context, but can still influence the plot. Refer to the cited papers for details on how each metric is defined.

- Site coverage | Extent to which the plot area is covered by roof surface, which is the most efficient location for PV installation in Singapore. Thus a high site coverage

scores higher as it tends to increase roof PV possibilities (Shi et al., 2021).

- Land-use ratio | The average of the deviations from a recommended ratio of commercial and residential land uses of 30 to 70% respectively within the city block (Shi et al., 2021). Plots that are in city blocks with less deviation from the recommendation score higher.
- Plot orientation | The deviation from preferable east-west orientation for efficient facade PV installation (Shi et al., 2021). It is modulated by the plot area as an orientation has a stronger impact on plots with a small area, limiting design possibilities. Plots with lower deviation score higher.
- Plot elongation | Is expressed as the ratio between two circle diameters related to the plot: diameter of a circle with the same area as the plot and the diameter of the smallest bounding circle (Shi et al., 2020). Elongation is again modulated by the plot area (for the same reason as above). Smaller elongation tends to increase compactness allowing for efficient deployment of PV infrastructure.
- Shading factor | The difference between plot building height and average building height within a given radius. Plots with tall buildings have larger facade areas for PV installation and less risk for inter-intra shading (Shi et al., 2021) and thus score higher.

For the proof of concept, we normalised individual metric scores to a [0-1] range so metrics can be combined easily, with each metric contributing equally to the combined assessment. In reality, the importance or weight of different metrics varies; determining these variations is an operations research task that falls outside the scope of this paper. The metric scores were normalised between 0.0 and 1.0, where 1.0 represents the maximum value in the dataset and 0.0 is the minimum value. By default, we assigned normalised individual metric scores lower than 0.5 to W or T and scores higher than 0.5 to S or O. In this way, each metric was assigned to the SWOT categories for each plot. We computed a single S W O and T value per plot by aggregating results. Finally, analysis results were visualised in thematic gradient maps (Figure 3), presenting the geospatial distribution of the overall SWOT score, which is the difference between aggregated SO and WT, implying that in a positive overall score SO outweigh WT, and vice versa for a negative total score.



*Figure 3 shows the geospatial distribution of overall SWOT analysis score across four chosen regions. Overall SWOT score is the difference between aggregated SO and WT. If positive, SO outweigh WT and are coloured in blue shades and, if negative, WT outweigh SO and are coloured in red shades. These maps show that our automated SWOT framework is consistent across cases with different topological features, development phases and information coverage.*

4. Results

We applied our automated SWOT analysis to every plot in the aforementioned four Key Growth Areas in Singapore. The spatial distribution of overall SWOT scores in gradient maps (Figure 3) was instrumentally used to explore the solar energy potential of different plots. For example, parks, public spaces and small vacant plots scored lower across all locations due to the interaction between low site coverage and unsuitable land-use ratio that results in limited PV installation possibilities. Plots with east-west orientation and close to parks or public spaces scored higher due to interaction between desirable orientation, site coverage and shading effects.

To explore the relative contributions of each descriptor and individual metrics to the total score, we further investigated one region — Central Business District (CBD) (Figure 4b). From the thematic maps of four descriptors (Figure 4c) and a selected



Figure 4 shows one of the four regions SWOT analysis results: a) presents the score composition of four example plots; b) visualises the distribution of total SWOT analysis results; c) maps visualise the distributions of separate descriptors. Concentrations of S and O primarily consist of developed plots with high-rise buildings, while W and T cluster in areas that are not yet developed; d) displays a bar chart with SWOT results for all plots, showing that most plots balance negative and positive factors and that only very few plots are highly suitable or highly unsuitable for PV installation.

sample of plots (Figure 4a), we see that the interaction between orientation and elongation metrics (driven by a non-uniform grid urban structure) were main contributors to Strengths and Weaknesses determining varying plot performance. Plots that border water bodies or are used as pedestrian passages scored particularly low due to high elongation. The absence of buildings in this developing area had two implications: site coverage additionally contributed to the plot's Weaknesses, while shading factor contributed to Opportunities of those plots that are already built up. The majority of the plots have commercial land-use types which usually have higher energy use intensity, resulting in low land-use ratio scores and contributing to Threats. Finally, in our results, the majority of plots featured a balance between negative and positive factors, resulting in moderate solar energy potential and only very few plots were highly suitable or highly unsuitable for PV installation (Figure 4d) - at least in this unweighted multi-criteria assessment.

## 5. Discussion and Future Work

This paper presented a novel approach that integrates SWT, spatial analysis based on derived metrics, and a SWOT framework in order to support decision making in city planning. The benefits of this integration have been demonstrated through an experimental use case that shows how this method could be used to inform planners in SDG-related planning tasks. We emphasise that the main aim of this paper is to introduce a proof-of-concept of our automation approach, thus the type of evaluation and the specific scores resulting from our evaluation should be considered illustrative examples to demonstrate our KG-based automated SWOT analysis. The presented results demonstrate two main innovations of our approach.

First, planners could use generated maps to indicate potential plots for PV installations at the early master-planning level. The consistency of generated maps in Figure 3 shows that the automated SWOT framework can work across areas with varying topological features, information coverage or development stages. Thus, planners can apply this framework to a variety of sites and cases and even use it for the evaluation of on-site solar potential over time, as the features of an area change.

Second, the innovative value of the automated proof-of-concept approach stems from the ability to effectively combine heterogeneous linked data (derived metrics) with multidisciplinary expert knowledge (criteria evaluation and classification) into a single intuitive model (SWOT) — an ability that has been officially acknowledged as being inherently necessary for the successful implementation of SDGs (Massaro et al. 2020). The reciprocal relationship between metrics and automated evaluation opens new possibilities for SWOT analysis applications and provides a well-known interface (at least for urban planners) to interact with large semantic data stores, supporting decision-making. For instance, evaluating land use (e.g. mixed-use) developments in an automated way with more granulated linked data could facilitate and even augment land use classification tasks (Shi et al., 2021). Moreover, the malleable KG data architecture is designed for data to be updated, expanded or created with relative ease; such data architecture supports planning scenario generation and exploration, enabling the simulation of parallel worlds (Eibeck et al., 2020). We demonstrated that SWT and KGs enable planners to analyse cities through multi-domain indicators, as derived metrics used in the SWOT analysis would be challenging to develop based on

individual datasets.

Despite the presented strengths, the work also has limitations. As the main aim of this specific SWOT analysis of solar energy potential is to illustrate our automation approach, the evaluation could be made more robust. The presented SWOT is limited by the choice of metrics, their uniform weighting, and different levels of granularity, and should therefore be considered a rough gauge for on-site solar energy potential assessment, primarily suitable for early-stage planning considerations. For instance, the low granularity of open zoning data poses challenges for computing an accurate land-use ratio. To make evaluation results more robust and informative, uniform metric weighting should be replaced by collaboratively determined and case-specific weights. The presented uniformly weighted assessment model could also prove useful in the expert-based experiments that would be needed to elucidate weights and even value functions. We also observed a known shortcoming of SWOT analyses — some metrics can be categorised as several descriptors. Cited SWOT analysis examples addressed this issue and could be used to inform further refinement of S W O and T categorisation. Additionally, while the current SWT and KGs integration with SWOT analysis framework requires human interaction, it could in principle be performed by a dynamic multi-agent system (Zhou et al., 2019).

We anticipate several directions for further work. Deriving diverse metrics requires the integration of additional datasets and urban simulation tools from different domains into the KG. We are also researching how city planning goals and SDGs can be structured ontologically, to develop more robust evaluation approaches.

## Acknowledgements

This research is supported by the National Research Foundation, Prime Minister's Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme.

## References

- Akroyd, J., Harper, Z., Souter, D. N. L., Farazi, F., Bhave, A., Mosbach, S., & Kraft, M. (2021). Universal Digital Twin—Land Use. *c4e-Preprint Series*. Cambridge.
- Akroyd, J., Mosbach, S., Bhave, A., & Kraft, M. (2021). Universal Digital Twin—A Dynamic Knowledge Graph. *Data-Centric Engineering*, 2.
- Blazegraph. (2020). Geospatial support in Blazegraph. Retrieved November 13, 2021, from <https://github.com/blazegraph/database/wiki/GeoSpatial>
- Chadzynski, A., Krdzavac, N., Farazi, F., Lim, M. Q., Li, S., Grisiute, A., Herthogs, P., et al. (2021). Semantic 3D City Database—An enabler for a dynamic geospatial knowledge graph. *Energy and AI*, 6, 100106.
- Chadzynski, A., Li, S., Grisiute, A., Farazi, F., Lindberg, C., Mosbach, S., Herthogs, P., et al. (2021). Semantic 3D City Agents—An intelligent automation for Dynamic Geospatial Knowledge Graphs. *c4e-Preprint Series*. Cambridge.
- Comino, E., & Ferretti, V. (2016). Indicators-based spatial SWOT analysis: Supporting the strategic planning and management of complex territorial systems. *Ecological Indicators*, 60, 1104–1117.
- Data.gov.sg. (2021). *Dataset List*. Retrieved November 13, 2021, from <https://data.gov.sg/search?groups%3Denvironment>

- Eibeck, A., Chadzynski, A., Lim, M. Q., Aditya, K., Ong, L., Devanand, A., Karmakar, G., et al. (2020). A Parallel World Framework for scenario analysis in knowledge graphs. *Data-Centric Engineering*, 1.
- Energy Market Authority. (2017). Renewable Energy—Overview. Retrieved February 4, 2017, from [https://www.ema.gov.sg/Renewable\\_Energy\\_Overview.aspx](https://www.ema.gov.sg/Renewable_Energy_Overview.aspx)
- Ervural, B. C., Zaim, S., Demirel, O. F., Aydin, Z., & Delen, D. (2018). An ANP and fuzzy TOPSIS-based SWOT analysis for Turkey's energy planning. *Renewable and Sustainable Energy Reviews*, 82, 1538–1550.
- Ghazinoory, S., Abdi, M., & Azadegan-Mehr, M. (2011). SWOT methodology: A state-of-the-art review for the past, a framework for the future. *Journal of Business Economics and Management*, 12(1), 24–48.
- Grisiute, A., Silvennoinen, H., Shi, Z., Chadzynski, A., Li, S., Lim, M. Q., Sielker, F., et al. (2021). Creating multi-domain urban planning indicators using a knowledge graph: A district energy use case in Singapore. *Presented at the International Conference on Evolving Cities*, Southampton, UK.
- Learned, E. P., Christensen, C., Andrews, R. S., & Guth, D. (1965). *Business policy: Text and cases*. Irwin Publishers.
- Massaro, E., Athanassiadis, A., Psyllidis, A., & Binder, C. R. (2020). Ontology-Based Integration of Urban Sustainability Indicators. In C. R. Binder, E. Massaro, & R. Wyss (Eds.), *Sustainability Assessment of Urban Systems* (pp. 332–350). Cambridge University Press.
- Richthofen, A., Herthogs, P., Kraft, M., & Cairns, S. (2022). Semantic City Planning Systems (SCPS): A Literature Review. *Journal of Planning Literature*, 08854122211068526.
- Shi, Z., Fonseca, J. A., & Schlueter, A. (2017). A review of simulation-based urban form generation and optimization for energy-driven urban design. *Building and Environment*, 121, 119–129.
- Shi, Z., Fonseca, J. A., & Schlueter, A. (2021). A parametric method using vernacular urban block typologies for investigating interactions between solar energy use and urban design. *Renewable Energy*, 165, 823–841.
- Shi, Z., Herthogs, P., Li, S., Chadzynski, A., Lim, M. Q., Richthofen, A. V., Cairns, S., et al. (2021). Land Use Type Allocation Informed by Urban Energy Performance: A Use Case for a Semantic-Web Approach to Master Planning. *The 26th International Conference of the Association for Computer-Aided Architectural Design Research in Asia, CAADRIA 2021* (Vol. 2, pp. 679–688).
- Shi, Z., Hsieh, S., Fonseca, J. A., & Schlueter, A. (2020). Street grids for efficient district cooling systems in high-density cities. *Sustainable Cities and Society*, 60, 102224.
- URA. (2021). Urban Transformations. Retrieved November 13, 2021, from <https://www.ura.gov.sg/Corporate/Planning/Master-Plan/Urban-Transformations>
- W3C. (2013). SPARQL Query Language for RDF. Retrieved November 14, 2021, from <https://www.w3.org/TR/rdf-sparql-query/>
- White, T. H., de Melo Barros, Y., Develey, P. F., Llerandi-Román, I. C., Monsegur-Rivera, O. A., & Trujillo-Pinto, A. M. (2015). Improving reintroduction planning and implementation through quantitative SWOT analysis. *Journal for Nature Conservation*, 28, 149–159.
- Winkelhake, U. (2018). *The Digital Transformation of the Automotive Industry*. Springer International Publishing.
- Zhou, X., Eibeck, A., Lim, M. Q., Krdzavac, N. B., & Kraft, M. (2019). An agent composition framework for the J-Park Simulator—A knowledge graph for the process industry. *Computers & Chemical Engineering*, 130, 106577



# QUANTIFYING THE COHERENCE AND DIVERGENCE OF PLANNED, VISUAL AND PERCEIVED STREETS GREENING TO INFORM ECOLOGICAL URBAN PLANNING

QING YANG<sup>1</sup>, CHU-FAN CAO<sup>2</sup>, HAI-MIAO LI<sup>3</sup>, WAI-SHAN QIU<sup>4</sup>,  
WEN-JING LI<sup>5</sup>, and DAN LUO<sup>6</sup>

<sup>1,2,3,6</sup> *The University of Queensland.*

<sup>4</sup> *Department of City and Regional Planning, Cornell University.*

<sup>5</sup> *Center for Spatial Information Science, The University of Tokyo.*

*lyangqingbob@gmail.com, 0000-0002-1758-7712*

<sup>2</sup> *chufan.cao@uqconnect.edu.au, 0000-0002-3976-3393*

<sup>3</sup> *haimiao.li@uqconnect.edu.au, 0000-0003-2548-1819*

<sup>4</sup> *wq43@cornell.edu, 0000-0001-6461-7243*

<sup>5</sup> *liwenjing@csis.u-tokyo.ac.jp, 0000-0003-2590-5837*

<sup>6</sup> *d.luo@uq.edu.au, 0000-0003-1760-0451*

**Abstract.** This research attempts to combine the fields of urban planning, urban design and cognitive psychology, and propose three corresponding evaluation indicators for urban ecology, and further explore the coherence and divergence between them. This research defines land vegetation coverage, visibility of street green vegetation, and people's green perception as planned green, visual green and perceived green. Specifically, the three measures (i.e., planned, visual and perceived) refer to objectively extracting park lands and canopy areas from land use data, objectively extracting green pixels from street views, and subjectively collected through visual surveys. This study hypothesizes that there could exist large variation between the three measures, which would provide distinct implications for city planners. To test our hypothesis, this study selects Brisbane as the research area, effectively using computer deep learning, data visualization and mathematical statistics methods to achieve an accurate description of the three sets of data, and proposes a comprehensive evaluation of the urban ecological theory system. The results show the credibility and scope of application of the three types of greening, and quantitatively proposed and tested the relevant theories of urban design.

**Keywords.** Urban Green Space; Urban Ecology; Street View Image; Green Perception; Subjective Measure; SDG 3; SDG 11; SDG 13.

## 1. Introduction

Urban ecology refers to the relationship between living things and their surroundings in the context of urban environments and significantly affects the urban culture and aesthetic experience. (Mostafavi and Doherty, 2016) Therefore, urban designers have

long been committed to promoting urban ecology to enhance the health and well-being of the urban population (Breuste et al., 1998; Gaston, 2010). However, there is no agreement on the consistent connotation or measurement of ecology. Urban green spaces (UGSs) are often used as a proxy for the social benefits of greening and its role in sustainable urban development. In this regard, the greening ratio obtained from satellite images and land use GIS data is used in many studies (Kabisch et al., 2015). While the green ratio works as a satisfactory proxy for the 'planned green', emerging studies indicate that a more precise definition of ecology is needed for analyses at street-scale. For instance, Ma et al. (2021) used Google Street View images rather than satellite imagery to measure human-scale and eye-level 'visual green'. In addition, with increased interdisciplinary communication, scholars have begun to study the positive effects of green space on human psychology (i.e. perceived green), including stress relief, improved mental health, and increased well-being (Parsons, 1991). For example, Reid Ewing (Ewing and Handy, 2009) recorded video clips of 48 commercial streets in the United States and asked experts to rate their imageability, enclosure, human scale, transparency, and complexity. They developed a statistical model of the physical characteristics of the city and people's subjective perceptions.

However, there has been less discussion regarding the coherence and divergence between the three main types of greening. These are: firstly, the 'planned green', reflecting the proportion of planned green space and objective ecology; secondly, 'visual green', representing the proportion of vegetation within the line of sight; and thirdly, 'perceived green', capturing people's subjective perception of the environment. We hypothesise that these three types of green are interrelated but have different impacts on urban ecology, which have not been addressed adequately by the current urban ecology research. Table 1 describes some related research.

*Table 1. Summary of Related Works*

Type	Existing studies	Definition & Method	Main findings
Planned	1.(Li et al.,2015a) 2.(Breuste et al., 1998; Gaston, 2010) 3.(Mostafavi and Doherty, 2016)	Planned green is defined as the distribution of vegetation cover (including trees, shrubs, lawns, and other forms of plants). Researchers usually use satellite imagery to check the green distribution manually or make vegetation cover maps through photoshop.	These studies show that the evaluation of urban planning or street greening can be effectively carried out by obtaining the Pearson correlation coefficient between the Green View Index and vegetation coverage.
Visual	1.(Li et al.,2015a) 2.(Li et al.,2015b)	Visual green is defined as the visibility of green vegetation from a human perspective. The vegetation amount is determined by extracting the pixels of green vegetation from the Google Street View image using the vegetation classification algorithm. Results are further manually checked, and non-vegetation features that were mistaken for vegetation are deleted.	Google Street View can be effectively used as a street-level greening assessment. The visibility of green vegetation is positively correlated with human safety perception. Plants above the horizontal line have a stronger correlation with safety than those below the horizontal line.
Perceived	1.(Qiu et al.,2021) 2.(Ma et al.,2021) 3.(Ewing and Handy, 2009)	Perceived green is defined as people's feelings about the quality of urban green space (including forest, grass, greenbelts). The street view imagery (SVI) collected from the sample area is rated by crowdsourcing for the four perceptual qualities and used as training labels. Computer vision segmentation extracts physical features from SVI as explanatory variables. Finally, the machine learning model is trained to make predictions.	Through cross-over research, the framework can effectively predict perception from SVI data consistently. The study found several groups of related factors and then reached conclusions that are consistent or different from traditional urban theories. The study also believes that there is polarisation and uneven development in Shanghai's Pudong city.

This study attempts to synthesise theories from various fields, including urban planning, urban design, and cognitive psychology, through statistical analysis of satellite images, street images, and subjective perceptions of urban residents about the spatial environment. We combine this synthesis with image recognition and extensive data analysis methods and propose quantitative definitions of the three types of green systems.

It is known that urban ecology is an extremely important sustainable factor. Effective ecological designs can improve people’s well-being and health. This research proposes typology studies for physical space by defining and quantifying three urban green indicators, which provides design reference with specific data for the consistency of objective ecological feature and subjective ecological experiences. Therefore, it satisfies the following goals: Good Health and Well-being, Sustainable Cities and Communities, and Climate Action.

## 2. Methodology

### 2.1. METHODS OF OBTAINING PLANNED GREEN, VISUAL GREEN, AND PERCEPTUAL GREEN DATA

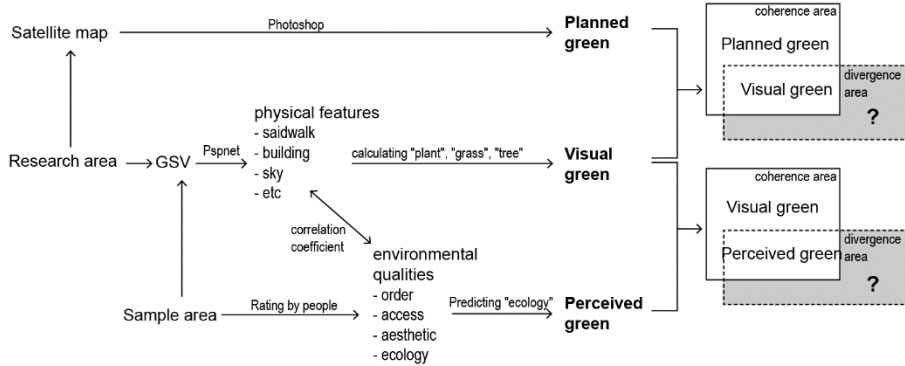


Figure 1. Flow chart for obtaining three evaluation indicators

Figure 1 shows how to obtain the three types of greening. Planned green and visual green are the physical characteristics of the street and are objective data obtained through analysis. By extracting the green pixels in the urban satellite image, we obtained the proportion of green pixels in the study area as the basis for planned green, reflecting the objective distribution of vegetation in the urban area. Planned green objectively measures green from an 'overlook' perspective. The formula can be expressed as the vertical projection area of the greening/urban land area.

For visual green, this research uses the pyramid scene parsing network (PsPnet) algorithm to identify objects in street view photos, such as trees, buildings, and cars and calculates the proportion in each photo. The proportion of pixels occupied by all vegetation types was the basis for our visual green. Visual green objectively measures green from human eye perspectives. The formula can be expressed as a green pixel/total street view imagery (SVI) pixel.

Perceptual green refers to people's subjective ratings of ecological experiences. We collected 500 street-view photos of Berlin city streets as training data. The photos were then randomly divided into pairs. Visitors from the online questionnaire platform were invited to score based on seven spatial environmental qualities, such as choosing which one is more ecological or orderly. The photos were ranked by scores based on the Microsoft TrueSkill algorithm. Then, based on the 38 street physical characteristics obtained from the PsPnet, the Pearson correlation coefficient was used to establish a correlation model with seven spatial environmental qualities. (For example, the data revealed that the enclosure strongly positively correlates with the sky, tree, and building areas.) After that, we collected more than 16,000 Google Street View photos in 22 areas of Brisbane as a database, used a statistical model to score these photos, and finally screened four representative areas for comparative research. Perceptual green subjectively measures green from human perceptual perspectives.

## 2.2. DATA ANALYSIS METHOD

### 2.2.1. The Relationship Between Planned Green and Visual Green

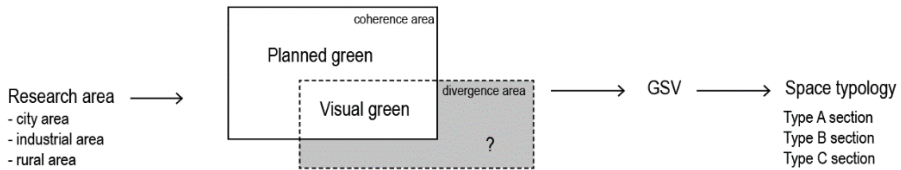


Figure 2. Analysis Method of Planned Green and Perceptual Green

Planned green and visual green behaved consistently in most areas. However, they diverged in some space typologies. These areas have a large amount of green vegetation from the urban planning perspective, but this was not reflected in the pedestrians' perspectives. This discrepancy was due to the view blockage caused by the inappropriate spatial organisation. The analysis method is shown in Figure 2. Firstly, we found the nodes with low visual green scores; then, we compared the distribution map of planned green to find places where the two values were different and returned to the street view to find the problem. Finally, we tried to classify and summarise these street scenes.

### 2.2.2. Relationship Between Visual Green and Other Related Factors and Perception of Green

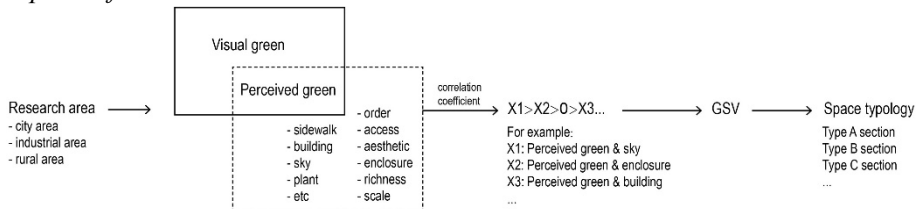


Figure 3. Analysis of Various Factors and Perceived Green

Figure3 shows the analysis process of this part. This part of the research was not limited to the analysis of visual green but included other factors related to perceived green. This was because visual green is just one of several factors affecting perceived green. Too much attention to visual green could distort the 'whole picture'. In addition, we divided city areas into different types, such as city, industrial, and rural areas, an approach that makes the study of urban phenomena more comprehensive and targeted.

### 3. Result

It is worth noting that some street view photos reflect the perspective from the roadway, which is inconsistent with the pedestrian perspective we studied. Therefore, this study refers to viewpoints close to the sidewalk and attempts to draw the road section accurately to ensure reliable analysis.

#### 3.1. COMPARING PLANNED GREEN AND VISUAL GREEN

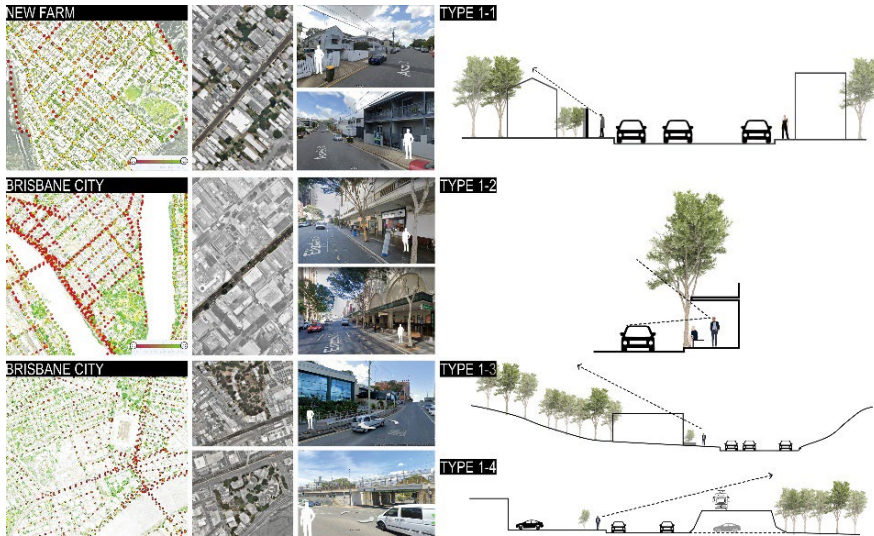


Figure 4. Four types of spaces where planned green and visual green do not match

As shown in Figure 4, the distribution of planned green and visual green is consistent in most areas, but there are still inconsistencies. By comparing planned green and visual green in most areas of Brisbane, the researchers discovered four common space typologies which lead to inconsistencies between planned green and visual green.

The New Farm (Figure 4) is a residential area close to the city centre. These sites have relatively higher housing prices and building densities than other suburban areas. Many owners make use of the open space in the plot as much as possible. As a result, buildings occupy a large number of front and back yards and reduce green areas. Through the analysis of Type 1-1, it was found that, although vegetation could be seen in the satellite image, most of the greenery was concentrated in the private backyard. The front yard was occupied by houses and could not accommodate trees. Instead, only shrubs could be planted, and these shrubs were hidden behind the fence. This pattern

leads to almost no green street view. This observation is influenced by social, political, and economic factors, such as excessive marketisation and privatisation, while the government's control of public space is reduced. At the same time, the high rate of return on capital and changes in the demographic structure affect land use.

As the most prosperous business district, Brisbane City has a very high building density. The subtropical climatic conditions require the building to include shading structures. As shown in Type 1-2, once the canopy of the street tree is too high, it will be obscured by the overhanging sunshade, which means only the trunk is visible to the pedestrian. If the sunshade could be shortened and the tree canopy lowered and used for shading, pedestrians would obtain a better green experience.

Milton is a relatively small business district that runs east to west beside Milton Road. Type 1-3 shows buildings along the street blocking the greenery. Satellite images show ample green spaces on the street corner, but people cannot see them because of the compact commercial stores along the street. In addition, the shop also occupies the landscaping space, therefore, only small shrubs can be grown. If the transparency of the ground floor could be increased and some spaces are left to accommodate trees, the ecological level of this area may be effectively enhanced. The situation in Type 1-4 is commonly found on pedestrian paths along railroads. Because such roads are often several metres above the ground, the sightline on the other side of the railway is blocked. In Brisbane, this situation is not very common, and green vegetation along the ramp is usually in place.

### 3.2. COMPARE PERCEIVED GREEN AND ITS RELATED FACTORS

This section examines several factors related to perceptual green in four different areas of Brisbane. Bardon and Red Hill are two suburban residential areas with dense vegetation and low building density. The City is located in the centre of Brisbane, facing the river on three sides, and has a grid of urban morphology and high building density. Finally, Woolloongabba is located on the south-eastern side of Brisbane City and is an industrial area. It is generally low in building density, with wide roads and large infrastructure, but lacks greenery.

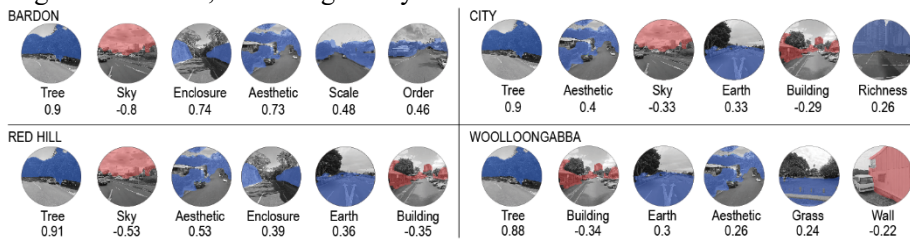


Figure 5. Factors related to perceived green in the four areas

These correlation coefficients can be seen in Figure 5. The data show that visual green and perceived green are highly positively correlated in these four areas. Therefore, visual green can be used to describe the ecological level of a city. However, from the urban design perspective, visual green is not the only factor that affects perceived green. The chart also shows the other factors related to perceptual greenery.

To further explore the significance of these data, the study selected two more

characteristic factors, order and enclosure, and analysed their specific performance in different areas in the physical space

### 3.2.1. Order and Perceived Green

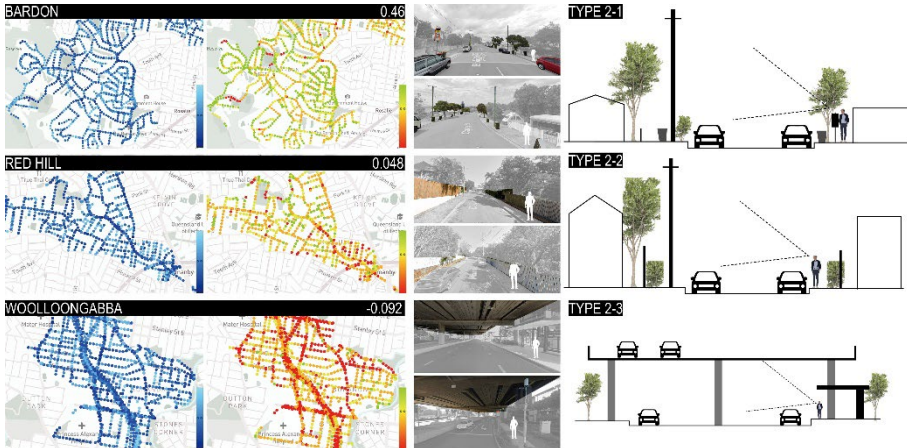


Figure 6. Spatial Analysis of Order and Perceived Green

Figure 6 clearly shows the research stages: first looking for relevant data and then returning to the street view to explore the problem. The data show that in Bardon, ecology and order have a strong positive correlation, with a correlation coefficient of 0.46. This generally means that the more orderly the space, the better the green perception in this area. The reverse is also true. The more chaotic the space, the worse the perception of green. To explore the potential causality behind these two related factors, it is necessary to analyse specific street scenes. Here, the two sets of data were first visualised. Then, by observing and selecting several nodes with low ecology and order scores and by summarising and analysing the characteristics of these nodes, we tried to discover how order potentially affects ecology. The study found that these street scenes all have some things in common: 1. Place trash cans along the street; 2. Park vehicles on the side of the road; 3. Telephone poles and signage; 4. Patchy vegetation. As shown in Type 2-1, many similar street patterns can be found in this area. Therefore, it may be inferred that this mode of spatial organisation will make people feel chaotic and potentially reduce their perception of green.

It is worth noting that, compared to Bardon, the data from Red Hill, which is also a suburb, show that there is a very weak positive correlation between ecology and order, with a correlation coefficient of only 0.048. In other words, the order of the streets here is not related to people's perceptions of green. If we make assumptions based on the conclusions drawn in Bardon that the lack of order in Red Hill still affects the perception of green, then there should be two other possible situations: the streets lack order but have a good green perception, or have order but lack green perception. Therefore, to find out how ecology and order run counter to each other, we looked for nodes with a larger difference in scores between the two. For example, the ecology score shown in the Figure6 was low, but the order score was high. It turns out that the

urban space here is generally organised by an orderly frame, which can be a continuous hedge, fence, building, or public facility. In other words, even if there is no single tree, it can be very orderly. Therefore, such a neat road no longer reflects the correlation between order and green perception.

A similar situation also occurs in Woolloongabba and City, which are industrial and downtown areas, respectively. There is a weak negative correlation between order and green perception, with a correlation coefficient of  $-0.062$  in City and  $-0.092$  in Woolloongabba. The photo shows the view under a viaduct in Woolloongabba and on the side of a stadium. This scene has similar characteristics; the artificial objects form an order, but the plants are squeezed aside.

These four areas show that order potentially affects green perception. Where there is a lack of planning, the order and perceptual greenness are more relevant, and the more artificial the place, and an irrelevant or even negative correlation between the two is likely.

### 3.2.2. Human Scale and Perceived Green

The definition of a human scale is not consistent. Generally speaking, if the various dimensions of the environment are close to the dimensions of a human being, we consider it to be a human-scale space. Road width, length, building height, width and other similar factors affect the sense of scale. Some details, such as the complexity of building components or decorative patterns and the number of pieces of street furniture, will also affect the sense of scale. In addition, even in high-density areas, if the design of the ground floor is sufficiently sophisticated, it can compensate for the human scale. According to Henry Arnold (1993), street trees can also adjust the scale of tall buildings and wide streets. This is because the canopy formed by leaves and branches allows people to experience a smaller space within a larger volume.



Figure 7. Spatial Analysis of Human Scale and Perceived Green

Figure 7 shows how the human scale and perceived green are related to the physical environment. In Bardon, the correlation coefficient between the human scale and



ecology was 0.48. This score shows that in this area, the closer the green elements are to the human scale, the better the people's green perception. The study also visualised the two sets of data and found places with higher scores for the ecology and human scale for analysis. It turns out that most of the roads at these nodes are relatively narrow, and the vegetation further forms an enclosure, shrinking the boundary. The tree canopy obscures the distant scenery and sky, thus forming a smaller space. The lush vegetation contributes a very high perceptual green score and forms a common road section in the suburbs. In City, the correlation coefficient between the human scale and ecology was 0.24. By comparison, it was found that although the high building density forms an oppressive frontage, the places with lush tree canopies can often get closer to the human scale.

It is worth noting that, in Woolloongabba, the correlation coefficient between the human scale and ecology was only 0.047. Through observation, it was found that most of the roads in this area are spacious, and the vegetation can only appear far away on both sides. Whether these trees are dense or sparse may affect green perception, but they cannot enclose the space. This is why even on a road where the sense of scale is always the same, the perception of green may be different.

Based on the analysis of these three regions, the higher correlation between human scale and ecology appears because tree crowns often provide human-scale space while forming a larger green interface. Therefore, these two do not seem to have a causal relationship. However, from another perspective, using vegetation to limit the space, especially increasing the density of the tree canopy, will be an effective means to promote green. This is because the vegetation can shrink the horizontal and vertical spatial scales in the street view.

#### **4. Discussion**

In general, although these three green scores have different meanings, they all have a high degree of credibility. In contrast, planned greens are more suitable for evaluating the overall ecology of a particular area owing to the simple calculations involved. However, other scores may be required when the research is specific to the street level. (For example, the four types of spaces proposed in this study.) In this regard, the visual green can replace the planned green for a more accurate ecological evaluation. (In different regions, the correlation coefficient between visual green and perceptual green may reach approximately 0.9.) Planned green and visual green are both objective physical quantities used to speculate on the ecological nature of the city, while perceptual green is the most realistic simulation of human reactions. However, this last score is the most difficult to obtain. Therefore, it is not necessary to use it to evaluate urban ecology (because visual green is reliable enough), but it can be used as a reliable measure of subjective feelings. By combining various influencing factors in the city, urban designers should put forward more quantitative hypotheses or test the various hypotheses already been proposed. For example, this study illustrates the connection between green perception and space syntax. Therefore, spatial typology research can intuitively reveal the comprehensive factors affecting people's green perception and provide urban designers with more objective and targeted urban design strategies.

Our study has several limitations. First, the small sample size and the lack of a detailed description of people's perceptions may lead to doubts about the scientific

validity. In the future, the sample size can be expanded, or the sample area can be classified. In addition, the evaluation criteria for subjective feelings can be refined. Researchers can use only suburban samples to predict the results of other suburbs.

## References

- Breuste, J., Feldmann, H., & Uhlmann, O. (1998). *Urban ecology*. Springer-Verlag.
- Ewing, R., & Handy, S. (2009). Measuring the unmeasurable: Urban design qualities related to walkability. *Journal of Urban Design*, 14(1), 65–84. <https://doi.org/10.1080/13574800802451155>
- Gaston, K. J. (2010). *Urban ecology*. Cambridge University Press.
- Kabisch, N., Qureshi, S., & Haase, D. (2015). Human–environment interactions in urban green spaces — a systematic review of contemporary issues and prospects for future research. *Environmental Impact Assessment Review*, 50, 25–34. <https://doi.org/10.1016/j.eiar.2014.08.007>
- Li, X., Zhang, C., & Li, W. (2015). Does the Visibility of Greenery Increase Perceived Safety in Urban Areas? Evidence from the Place Pulse 1.0 Dataset. *ISPRS International Journal of Geo-Information*, 4(3), 1166–1183. <https://doi.org/10.3390/ijgi4031166>
- Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q., & Zhang, W. (2015). Assessing street-level urban greenery using Google Street View and a modified green view index. *Urban Forestry & Urban Greening*, 14(3), 675–685. <https://doi.org/10.1016/j.ufug.2015.06.006>
- Ma, X., Ma, C., Wu, C., Xi, Y., Yang, R., Peng, N., Zhang, C., & Ren, F. (2021). Measuring human perceptions of streetscapes to better inform urban renewal: A perspective of scene semantic parsing. *Cities*, 110, 103086. <https://doi.org/10.1016/j.cities.2020.103086>
- Mostafavi, M., & Doherty, G. (2016). *Ecological urbanism* (Revised edition.). Lars Müller Publishers.
- Parsons, R. (1991). The potential influences of environmental perception on human health. *Journal of Environmental Psychology*, 11, 1–23.
- Qiu, W., Li, W., Liu, X., & Huang, X. Subjectively Measured Streetscape Perceptions to Inform Urban Design Strategies for Shanghai. *ISPRS International Journal of Geo-Information*. 10, 493. <https://doi.org/10.3390/ijgi10080493>

# THE EFFECT OF PATH ENVIRONMENT ON PEDESTRIANS' ROUTE SELECTION: A CASE STUDY OF UNIVERSITY OF CINCINNATI

JING TIAN<sup>1</sup>, MING TANG<sup>2</sup> and JULIAN WANG<sup>3</sup>

<sup>1</sup>University of Minnesota, <sup>2</sup>University of Cincinnati, <sup>3</sup>Penn State University

<sup>1</sup>tian0162@umn.edu, 0000-0002-2404-4097

<sup>2</sup>tangmg@ucmail.uc.edu

<sup>3</sup>jqw5965@psu.edu

**Abstract.** The present study on the influence of the path environment on pedestrians' route selection is mostly concentrated on the urban level while rarely discussed from the architectural level. Taking the University of Cincinnati (Ohio, US) as an example, this study aims to investigate whether the difference in the environmental settings of the route will affect pedestrians' walking experiences and future route selection, with the ultimate goal of ascertaining the underlying relationship between the route environments and the user behavior in the process of route selection and implementation. This study selected three routes from the Langsam library to the CEAS library. The research methods included data analytics, questionnaires, and comparative analysis. Firstly, through surveys and an E4 wristband, psychological and physiological data were collected. Secondly, Analysis of Variance (ANOVA) was used to examine whether there was a significant difference in pedestrians' walking experience among the three routes. Thirdly, through the analysis of questionnaires, the factors that play an important role in pedestrians' route selection were determined. It can be concluded that the three routes with different environmental settings bring a different experience to participants. More specifically, the level of comfort and openness of the route significantly affects the route selection of pedestrians, while the degree of fatigue during walking does not. To sum up, for the transition space from outdoor to indoor, the factors affecting pedestrian route selection include the route's degree of comfort and openness.

**Keywords.** Path Environment; Route Selection; Pedestrian; Data Analysis; Sustainable Built Environment; SDG 11.

## 1. Introduction

Many scholars have begun exploring the relationship between the pedestrian environment and human behavior. It contributes to building more walkable cities once we know how the environment affects pedestrians' route selection or sensory experience. Cepolina et al. use the scale of pedestrian comfort to assess whether a

facility's design impacts human comfort (Cepolina, Menichini, & Rojas, 2018). Guo et al. evaluates the environmental situation by analyzing the path selection of pedestrians (Guo & Loo, 2013). Isaacs explores how the environment of a path, such as the width and height of the elevation along the road, affects people's time perception during the walking journey (Isaacs, 2001). Miller and Laurie interview and observe the experience of pedestrian accessibility in high-density urban areas (Miller & Buys, 2013). Calvert explores urban walking experience holistic and multi-faceted experience (Calvert, 2020). Barros et al. try to determine which factors interfere with the choices people make about modes or transport of walking paths (Barros, Martínez, & Viegas, 2015). Hollmann reviews "how individual pedestrian behavior and the pedestrians' environment usage are realized in current pedestrian behavior simulation models has been undertaken" (Hollmann, 2015). Martínez and Ana present a Structural Equations model to assess pedestrian environment satisfaction (Martínez & Barros, 2014). Guo proposes a new method (based on path choice) to investigate the causal effect of the pedestrian environment on the utility of walking (Guo, 2009).

Also, more and more cities are increasingly considering the impact of path environment on the pedestrian's walking experience to improve the image and liveliness of a city. Many cities issued a corresponding urban street environment design manual to guide urban planners and urban designers to improve the community environment, comfort, and walkability. The design guidelines for street and sidewalk construction and retrofit design is also adopted by the cities such as Pennsylvania, Burlington, Vermont, Tacoma, Washington, Minneapolis, and Phoenix.

However, the present study on the influence of the path environment on pedestrians' route selection is mainly concentrated on the urban level while rarely discussed from the architectural level. Few studies have investigated the transition space from outdoor to indoor, and whether environment settings between the outdoor and indoor space affect pedestrians' walking experiences and route selections.

In this case, this study aims to investigate whether the difference in the environmental settings of each route would affect pedestrians' walking experiences and future route selection. Researcher conducts an experiment and collect pedestrians' psychological and physiological data while they are walking from the Langsam library to the CEAS library at the University of Cincinnati (Ohio, US). Then the collected data was analyzed to see whether path environmental settings affect pedestrians' experience and route selection. Three routes with different environmental settings from the Langsam library to the CEAS library are selected as the research testbeds. In this pilot study, a total of nine students participated the experiment. Their physiological data was obtained through an E4 wristband while the psychological data was collected through a questionnaire. Finally, through analysis of the questionnaire, data shows that the factors including comfort and openness of the transition space will affect pedestrians' route selection. However, the limitation of the study is the small sample size. Unfortunately, due to Covid-19, the researcher's ability to collect a larger research sample was restricted. In the future, there would be more generalizable and convincing results if larger sample size is used. It is hoped that this pilot study can ascertain the underlying relationship between the route environment and user behaviors in route selection and implementation.

**2. Experiment Design**

**2.1. SITE SELECTION**

The experiment site is located at the UC campus. For research purposes, two points were considered while selecting the routes: 1) the routes should share the same start point and endpoint; 2) environmental settings among the routes have significantly different characteristics. Based on the selection standards above, the gate of the Langsam library (outdoor) is set as the start point and the gate of the CEAS library (indoor) as the endpoint. Three routes connect these two points. These three routes have apparent characteristic differences in the setting of the roof garden, elevator, stairs, art installations, plaza, etc (Figure 1).

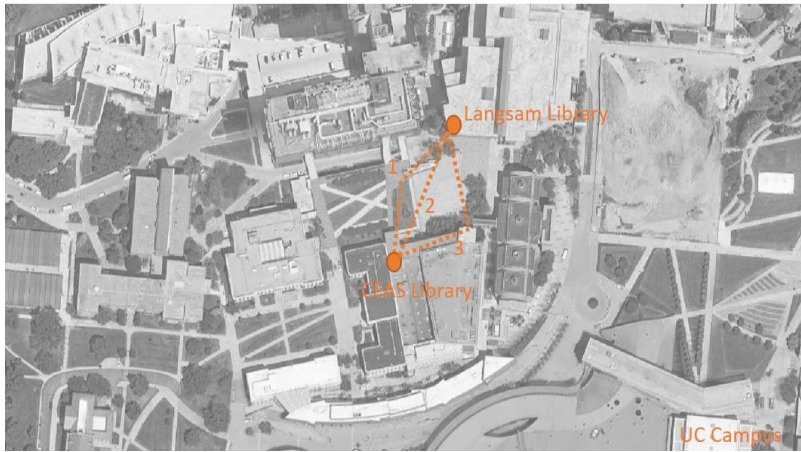


Figure 1: Site Location (Source: Author)

**2.2. ROUTE ANALYSIS**

The three routes from the Langsam library to the CEAS library have different outdoor and indoor space ratios, the proportion of stairs, and environmental settings (including plants and artwork) (Table 1). Route 1 has a total length of almost 155 m and the ratio of outdoor and indoor space is 2:1. The percentage of greenery and stairs are 28% and 40%, respectively (Figure 2). Route 2 has a total length of almost 157 m and the ratio of outdoor and indoor space is 1:2. The percentage of greenery and stairs are 3% and 42.7% (Figure 3). For route 3, the length is around 135 m and the ratio of outdoor and indoor space is 1:1. The percentage of greenery and stair is 14.8% and 2% (Figure 4).

Table 1: Routes Comparison (Source: Author)

Route No.	Length (≈)	Ratio of Outdoor and Indoor Space (≈)	Percentage of Greenery (≈)	Percentage of Stairs (≈)

Route 1	155m	2:1	28%	40%
Route 2	157m	1:2	3%	42.7%
Route 3	135m	1:1	14.8%	2%

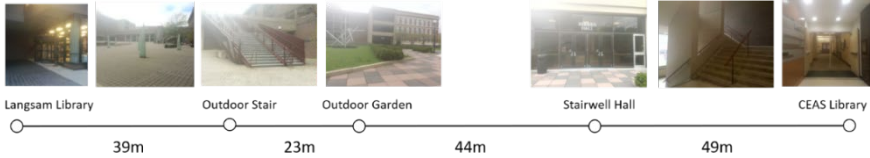


Figure 2: Overview of Route 1 (Source: Author)



Figure 3: Overview of Route 2 (Source: Author)

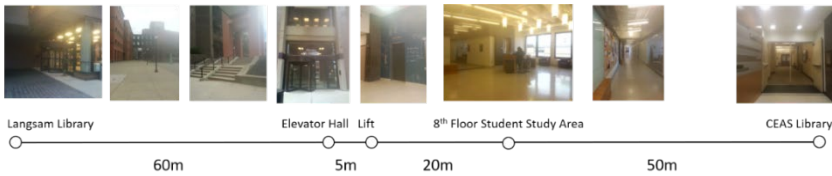


Figure 4: Overview of Route 3 (Source: Author)

Besides the physical indicators, we also consider the environmental factors of each route that may cause different psychological feelings. The route environment settings should afford the pedestrian with security or protection from injury in terms of the safety factors. In this study, the illumination and people flow are used to evaluate the safety of each route. There is a difference in illumination setting and pedestrian flow for these three routes. Route 1 offers greater security because it has much more outdoor space full of natural light, including a plaza, outdoor staircase, and roof garden. Even the staircase area is filled with natural light due to the huge glass windows. Meanwhile, a considerable crowd on the plaza and roof garden increase the route safety. Route 2 offers lower security. Besides the plaza, the whole route is very tortuous and dim due to a fire safety door and closed stairwell. Most places along the way are installed with artificial illumination. Very few people walk this route, which further weakens the safety of the route. Route 3 offers relatively higher security compared to route 2. Only the elevator and corridor place need artificial illumination and the remaining area of the route is full of natural light. There is many people on the route because of the existence of an elevator hall and study area, which helps enhance the route's safety.

In terms of continuous variables, it is about the possibility of constant involvement in various activities along the route. The types and diversity of events allowed in each route are different. Diversity activities allow for an array of visual possibilities. Due to

the plaza and roof garden, pedestrians may encounter all kinds of events while walking on route 1, such as a party on the plaza, photography at the art installation, chatting with friends, resting beneath the shade of trees, reading, etc. For route 2, the richness level of route activities is relatively low. There are a few spaces on the route that allow for any activities. Also, very few people walk on the route. Route 3 is easy for people to run into friends because of the elevator hall and study lobby. The richness level of route activities is relatively high. Students can carry out various activities in the study lobby.

### 2.3. EXPERIMENT SETUP

The experiment was conducted at the UC in the summer. Participants were asked to walk these three routes sometime from 12pm to 8pm on a sunny day. There was a total of 9 participants and 3 for each route. To avoid affecting the experimental results due to the different expectations and requirements of different participants (Faculty/Staff/Student) for the path environment, all participants in this experiment were students.

All participants need to walk from the Langsam library to the CEAS library through a different route. During the walking process, all of them need to wear an E4 wristband to record their changes in physiological data. The E4 is a medical-grade wearable device allowing us to collect real-time physiological data and conduct in-depth analysis and visualization ("E4 Wristband," n.a.). The data collected by E4 Wristband includes Electrodermal Activity (EDA), Blood Volume Pulse (BVP), Accelerometer (ACCE), and Skin Surface Temperature (TEMP). The data recorded through the E4 Wristband is synchronized to the E4 cloud storage and can be retrieved anytime.

After arriving at the destination, the participants were asked to complete a questionnaire similar to the Likert Scale. The questionnaire helped the researcher collect the psychological data regarding pedestrians' feeling and experience while walking the route. The questions included mainly care about the following indicators: comfort level, openness level of environment setting, fatigue level while walking, possibility to choose the same route later, ventilation quality, visual quality, temperature, walking time, etc. At the end of the questionnaire, participants could give some advice on the improvement of route environmental settings, which was optional. The whole questionnaire took an average of 2-3 minutes to complete. In addition, the researcher followed the participant and recorded the time when the participant arrived at a specific node of each route by a stopwatch. The timestamps helped to study whether the physiological characteristics vary with the environmental settings.

## 3. Analysis and Discussion

### 3.1. DATA ANALYSIS

Based on the analysis, pedestrians have a different experience while walking on route 3 compared to the other two routes. People feel more satisfied with route 3. In other words, the various environmental settings of different routes trigger changes in the physiological characteristics of the participants, which further cause the different experiences of the three routes. However, more analysis of physiological data is still

needed to increase the credibility of the conclusion.

3.1.1. *Identify the Indicator and Analysis Method*

For the indicator, the Electrodermal Activity (EDA) of all participants is for data analysis. EDA is used to measure the constantly fluctuating change at the surface of the skin. It will arise when the skin receives innervating signals from the brain. For example, if a person experiences emotional activation, the brain will send signals to the skin to used increase the level of sweating. We can measure how pedestrian feel while walking on each route based on the indicator.

For the analysis method, the Analysis of Variance (ANOVA) is used to determine whether differences between groups are statistically significant by comparing their means ("Understanding Analysis of Variance (ANOVA) and the F-test," 2016). First, the ANOVA tests whether there are noticeable differences among these three data sets in terms of EDA. Then, the F-test is used to compare every two groups of data to analyze whether there is a significant difference in experience between the two routes. "The null hypothesis for ANOVA is that the mean (average value of the dependent variable) is the same for all groups."(Creech, n.d.) If the P-value from the F-test is more than 0.05, the result supports the null hypothesis at the 95% confidence level. If the P-value is less than 0.05, then the result rejects the null hypothesis. This means there is a significant difference between these data sets, further indicating the differences in environmental settings of each route cause physiological characteristics change.

Table below is the mean of each participant's EDA (Table 2). ANOVA is used to examine these groups of data. Before the ANOVA test, all of these groups of data show normal distribution through the Shapiro-Wilk Normality test. The P-value of the three sets of data is 0.488, 0.976, and 0.28 respectively. All of them are greater than 0.05, indicating that the data shows a normal distribution with a 95% confidence level (Table 3).

Table 2: EDA Mean of Each Participant (Source: Author)

	Partici- pant	Route 1	Route 2	Route 3
EDA Mean ( $\mu$ s)	A	0.583	1.294	0.856
	B	0.11	1.576	3.543
	C	0.459	1.438	0.34

Table 3: Normality Test Result (Source: Author)

Shapiro–Wilk Normality Test		
Data: R1 EDA	Data: R2 EDA	Data: R3 EDA
W =0.92986, P-Value = 0.4881	W =0.99985, P-Value = 0.9765	W =0.8672, P-Value = 0.2876
Route 1	Route 2	Route 3



Then, according to the Test of Homogeneity of Variances, three data groups meet the requirement of homogeneity. The result of Fligner-Killeen Test, which is used to test the homogeneity of the variance, shows that the P-value is 0.4191 ( $>0.05$ ). This result indicates that the variance in EDA mean of each participant is statistically significantly the same for these three routes (Table 4).

Table 4: Result for Fligner-Killeen Test of Homogeneity of Variance (Source: Author)

<b>Fligner-Killeen Test of Homogeneity of Variance</b>
Data: Values by ind
Fligner-Killeen: med Chi-squared = 1.7395, df = 2, P-Value = 0.4191

### 3.1.2. Variance Comparison Among the Three Routes

The ANOVA result suggests that the participants' experience in the three routes are not completely significantly different. The P-value, by the ANOVA test, is 0.34 ( $>0.05$ ), which cannot reject the null hypothesis (Table 5). That is, the three sets of data are very similar, indicating the three routes give people a similar experience. Thus, we need to further explore the difference between every two groups of data.

Table 5: ANOVA Result for Three Groups (Source: Author)

	Df	Sum Sq	Mean Sq	F Value	Pr(>F)
Ind	2	2.557	1.278	1.263	0.349
Residuals	6	6.075	1.012		

Through F-test analysis to examine the difference between every two groups, route 3 brings different experience to people compared to routes 1 and 2 (Table 6). When analyzing the Route1 and 2, the 0.496 P-value ( $> 0.05$ ) suggests that pedestrians' experiences on these two routes are quite similar. Through analysis of Route 3 verse Routes 1 and 2 respectively, the 0.039 and 0.013 P-value ( $<0.05$ ) indicates Route 3 provides a different experience to people compared to the other two routes.

Table 6: P-value of F-test for Every Two Groups (Source: Author)

P-value	Route1_EDA	Route2_EDA	Route3_EDA
Route 1_EDA	/	0.4969	0.0398
Route 2_EDA	0.4969	/	0.0134
Route 3_EDA	0.0398	0.0134	/

### 3.2. SURVEY ANALYSIS

The questionnaire reveals that the difference in environmental settings of each route result in a difference in walking experience (Table 7). First, the comfort level while walking directly affects pedestrians' route preference: the more comfortable they feel during walking, the more likely to choose the same route. The comfort level of each route is ranked as: Route 3 (100%) > Route 2 (0%) > Route 1 (-33.3%). The participants' route preference is Route 3 (22.2%) > Route 2 (-22.2%) > Route 1 (-33.3%). Then, when selecting the route, the degree of openness of the environment (from the pedestrians' evaluation) and the pedestrians' route preference shows a negative correlation — the more open the route, the less likely to choose the route.

Although Route 1 has the most open environment settings (openness: Route 1 (66.6%) > Route 2 (-11.1%) > Route 3 (-33.3%)), people are the least likely to take this route (route preference: Route 3 (22.2%) > Route 2 (-22.2%) > Route 1 (-33.3%)). Even though Route 3 has the least open environment, it provides the best experience for people. The rating of Route 3 ranks No. 1 when asking people which route they prefer to choose again. But for Route 1, people feel uncomfortable and unlikely to choose it again, and for Route 2, people hold a neutral attitude to its comfort level and state they don't want to choose this route again. Finally, compared to the tired feeling while walking, the degree of openness and comfort of route play a more crucial role in pedestrians' route selection. Although participants feel more tired while walking on the Route 2 compared to Route 1 (tired level: Route 2 (44.4%) > Route 1 (33.3%)), they are more likely to select Route 2 (route preference: Route 2 (-22.2%) > Route 1 (-33.3%)).

Table 7: Route Evaluation from Participant (Source: Author)

Route No.	Participant (Age/Gender)	Comfort Level (3 Comfortable ~ -3 Uncomfortable)	Environment Setting (3 Open ~ -3 Closed)	Tired Level (3 Relaxed ~ -3 Tired)	Choose Same Route (3 Yes ~ -3 No)
Route 1	24/F	-1	3	-1	-1
	27/M	-1	2	-1	-1
	26/F	-1	1	-1	-1
Route 2	27/M	-1	1	-1	1
	29/F	-1	-1	-2	-2
	25/M	2	-1	-1	-1
Route 3	32/F	3	-2	0	-1
	19/M	3	-2	3	2
	24/F	3	1	0	1

#### 4. Conclusion

For the transition space from outdoor to indoor, the three routes' different environmental settings provide participants with different experiences that can affect their future route selection. Data Analysis indicates that the pedestrian's Route 3

experience is totally different from Routes 1 and 2, but the experience of Route 1 is similar to Route 2. This means that the different environmental settings between Route 3 versus Routes 1 and 2, such as the elevator hall and study area, cause these distinct experiences. Route 3 is more psychologically satisfying and leads to more significantly physiological changes in participants when compared with Routes 1 and 2. From the lens of subjective psychological experience, although the environmental setting of Route 3 is the least open, it makes people feel comfortable and relaxed. Thus, there is a 22.2% probability of them selecting Route 3 again.

Meanwhile, the willingness of pedestrians to choose specific routes depends on the level of comfort and openness of the route. The comfort level is positively correlated with the path preference: the more comfortable the route, the more likely a pedestrian tends to choose that route. Conversely, the openness level is negatively correlated with the path preference: the less open the route, the more likely the pedestrian is to choose the route. Compared with the comfort and openness indicator, the degree of fatigue caused by walking on a specific route has fewer effects on people's route preferences and selection. According to the survey, Route 3 is the most popular route among the three routes because of the highest comfort level, the lowest tiredness level, and the lowest openness.

One of the limitations of this research is the small sample size. Unfortunately, due to Covid-19 (2020), the researcher's ability to collect a larger research sample was restricted. Future research might have more generalizable and convincing results if a larger sample size from the UC is used. This could be accomplished by adding more sites and assessing a greater number of environmental settings. This would allow further discussion about what environmental factors impact route selection when pedestrians move from outdoor to indoor.

### Acknowledgements

We would like to show our gratitude to Na Chen (Professor, Sun Yat-sen University) for her guidance and advice on the data analysis part.

### References

- Barros, A. P., Martínez, L. M., & Viegas, J. M. (2015). A new approach to understand modal and pedestrians route in Portugal. *Transportation Research Procedia*, 10, 860-869. <https://doi.org/10.1016/j.trpro.2015.09.039>.
- Calvert, T. (2014). *An exploration of the urban pedestrian experience, including how it is affected by the presence of motor traffic*. [PhD Thesis, University of the West of England, EThOS]
- Cepolina, E. M., Menichini, F., & Rojas, P. G. (2018). Level of service of pedestrian facilities: Modelling human comfort perception in the evaluation of pedestrian behaviour patterns. *Transportation research part F: traffic psychology and behaviour*, 58, 365-381. <https://doi.org/10.1016/j.trf.2018.06.028>.
- Creech, S. (n.d.). *ANOVA*. Retrieved June 5, 2020, from <https://www.statisticallysignificantconsulting.com/Anova.htm>
- Anonymous (n.d.). *E4 Wristband: Real-time physiological data streaming and visualization. Empatica*. Retrieved June 5h, 2020, from <https://www.empatica.com/en-int/research/e4/>.

- Guo, Z. (2009). Does the pedestrian environment affect the utility of walking? A case of path choice in downtown Boston. *Transportation Research Part D: Transport and Environment*, 14(5), 343-352. <https://doi.org/10.1016/j.trd.2009.03.007>.
- Guo, Z., & Loo, B. P. (2013). Pedestrian environment and route choice: evidence from New York City and Hong Kong. *Journal of transport geography*, 28, 124-136. <https://doi.org/10.1016/j.jtrangeo.2012.11.013>.
- Hollmann, C. (2015). *A cognitive human behaviour model for pedestrian behaviour simulation*. University of Greenwich. [PhD Thesis, University of Greenwich]
- Isaacs, R. (2001). The subjective duration of time in the experience of urban places. *Journal of Urban Design*, 6(2), 109-127. <https://doi.org/10.1080/13574800120057809>.
- Martínez, L. M., & Barros, A. P. (2014). Understanding the factors that influence pedestrian environment quality. *Paper presented at the Transportation Research Board 93rd Annual Meeting Compendium of Papers; Transportation Research Board: Washington, DC, USA*.
- Miller, E., & Buys, L. (2013). Exploring pedestrian accessibility and walkability: Insight from public place observations in inner-urban higher density Brisbane, Australia. *Urban Mobility: Textual and Spatial Urban Dynamics in Health, Culture, and Society*, 188-196.
- Anonymous (2016). *Understanding Analysis of Variance (ANOVA) and the F-test*. Minitab. Retrieved June 5, 2020, from <https://blog.minitab.com/blog/adventures-in-statistics-2/understanding-analysis-of-variance-anova-and-the-f-test>

# A FRAMEWORK FOR A GAMEFUL COLLECTIVE URBANISM BASED ON TOKENIZED LOCATION DATA AND LIQUID DEMOCRACY

*Early prototyping of a case study using e-bikes*

SALMA TABI<sup>1</sup>, YASUSHI SAKAI<sup>2</sup>, NGUYEN TUNG<sup>3</sup>, MASAHIRO TAIMA<sup>4</sup>, AQIL M. CHEDDADI<sup>5</sup> and YASUSHI IKEDA<sup>6</sup>

<sup>1,3,4,5,6</sup>*Keio University.* <sup>2</sup>*Massachusetts Institute of Technology.*

<sup>1</sup>*tabi@keio.jp, 0000-0002-2607-6287*

<sup>2</sup>*yasushis@media.mit.edu, 0000-0002-7570-602X*

<sup>3</sup>*tungnguyen173371@keio.jp, 0000-0002-9524-6622*

<sup>4</sup>*taima@keio.jp, 0000-0001-5420-201X*

<sup>5</sup>*aqil@sfc.keio.ac.jp, 0000-0002-4741-0950*

<sup>6</sup>*yasushi@sfc.keio.ac.jp, 0000-0002-2016-4083*

**Abstract.** The participation of citizens in designing their social and built environments is vital for the creation of sustainable cities and communities. However, in practice, collective decision-making remains challenging. Several researchers have proposed innovative models of governance to achieve a more democratic participation. This paper attempts to contribute to this topic from the viewpoint of urban planning. The objectives are twofold. First, to introduce a conceptual framework of a gameful collective process of urbanism based on location data. Second, to present an early stage of prototyping a case study using e-bikes. Research questions are elaborated as follows: How can collective processes of urban planning engage the collective intelligence and the local knowledge of the community? How to utilize technological tools to support new forms of participatory urban governance? The main contribution of this work lies in the combination of the concepts of temporal ownership of public space, tokenization of location data, and liquid democracy, to design a dynamic and gameful decision-making process that promotes collective intelligence.

**Keywords.** Collective Urbanism; Liquid Democracy; Temporal Ownership; Tokenization; Location Data; Data Dignity; Gameful Design; SDG 11.

## 1. Introduction

In this paper, the term collective urbanism describes a collaborative process of decision-making in urban planning and design, based on the tokenization of location data, to account for the community's influence on and control over how to improve local urban spaces. This work proposes a framework for building digital citizen

participation platforms and presents an early prototyping of the user interface of a case study that utilizes e-bikes. The rest of the paper is organized as follows. The first section, titled theoretical grounding, provides an overview of the concepts and theories on which this paper is based; followed by the results section, which presents the designed framework and the preliminary case study. Afterwards, the discussion section demonstrates the characteristics of this framework, its advantages, and limitations.

## **2. Methodology**

This paper introduces the conceptual part of a design-based research experiment grounded in a theory-driven approach. Based on literature review, a conceptual framework for digital citizen participation was designed. It is to be implemented and developed through an ongoing process of case studies. This research takes interest in the potential of location data in merging the physical and digital experiences of urban space and connecting online and offline communities.

## **3. Theoretical Grounding**

### **3.1. COLLECTIVE URBANISM**

Collective urbanism refers to the process of making decisions about urban space in a collective and decentralized way, whereby citizens are involved as generators of solutions (Sakai, 2017). Jon (2021, p. 324) states that urban planning research is about creating collectives, or entities that are sources of solidarity and togetherness, which are essential for creating more inclusive urban spaces. The author emphasises the need for urban planners to broaden the scope of legitimate forms of expertise, by valorising situated knowledge and perspectives often possessed by the public (Jon, 2021, p. 324).

#### *3.1.1. Citizen Participation in the Decision-Making Process*

Citizen participation can be defined as ‘the process by which members of a society (...) share power with public officials in making substantive decisions and in taking actions related to community’ (Roberts 2004, p. 320, as cited in Ertiö, 2015, p. 308). The participation of citizens in democracy achieves both the objectives of civic education and the creation of a feeling of belonging to the community (Michels & De Graaf, 2010, p. 480). Some of the challenges of "real" citizen participation are: the time consuming process and the increasing multiplicity of the stakeholders involved (Pilemalm, 2018, p. 5:5); in addition to the representation of the diverse groups of citizens (Michels & De Graaf, 2010, p. 479).

In the emerging contexts of civic engagement and efficient democratic processes, the integration of ICT solutions contributed to the development of e-government initiatives, to create efficient public services and to involve citizens through e-participation platforms (Pilemalm, 2018, p. 5:2). Estonia's e-government based on blockchain technology adopts internet voting and online platforms for public consultation and for citizens to submit collective initiatives (e-estonia.com). Digital technologies facilitate citizen participation in urban governance by overcoming the necessity of the physical presence at the same time and place, which is required by conventional citizen participation methods (Ertiö, 2015, p.3 03). Nonetheless, citizen

participation platforms still face the challenge of generating structured data that could be translated into practical solutions and actions, without constraining the ability of citizens to express their diverse opinions (Sakai, 2017, p. 27).

### *3.1.2. Technopolitics Discourse in Contrast with Smart Cities*

The 'Technopolitics discourse' is a sociological understanding of technology that contrasts with main smart cities' approaches, which are criticised for being managerial and seemingly participatory (Smith & Martín, 2021). Technopolitics did not develop in reaction to the smart city, but emerged as a bottom-up movement born out of digital activism (Smith & Martín, 2021, p. 314). Consul, an open-source software platform for citizen participation, is one of the tools developed within this discourse. It allows citizens to get involved in debates, proposals and participatory budgets (Secinaro et al., 2021). Smith and Martín (2021) explain that technopolitics is based on two main concepts: 'Democratic participation' using digital platforms to innovate in governance models; and 'Collective intelligence' where both citizens and experts are initiators of proposals. The collective urbanism framework presented in this paper aligns with the bottom-up collective approach of technopolitics. However, it differs by adopting liquid democracy as a model for democratic participation rather than direct democracy to which technopolitics is committed (Smith and Martín, 2021, p. 324).

## 3.2. TOKENIZED LOCATION DATA: TEMPORAL OWNERSHIP

### *3.2.1. Citizen Participation Applications Based on Location Data*

The technology of Global Positioning System (GPS) allows users to attach information to location points through map-based interfaces. Location data plays the role of a medium between virtual and physical layers of social and built environments. For instance, the Public Participatory Geographic Information System (PPGIS) is a user-friendly platform aiming at involving citizens in the collection of geographic data on a volunteer basis (Poplin et al., 2017). Yet, they remain primarily a tool for data collection through citizens (Nuojuua, 2010, as cited in Ertiö, 2015, p. 304). Another example is Commonplace, a platform that invite citizens to share geo-tagged comments about design proposals in their neighbourhood. Furthermore, Public Gratification Palace is a civic engagement and participation framework whereby citizens can interact with each other's comments only when they are physically present in the same location, promoting social interaction in urban space (Jiang et al., 2021).

### *3.2.2. Temporal Ownership: Tokenization of Location Data*

'The process of tokenization refers to issuing blockchain-based tokens that can be traded, stored, and transferred in the digital world. These tokens exist on the chain, act as a store of value and carry the rights of the assets they represent, while the real-world assets backed by these tokens continue to exist "off-chain."' (Becker, 2020, para. 2)

The concept of 'Temporal ownership of public space' is adopted from the work of Sakai (2017, p. 72). It can be defined as an immaterial ownership of public space based on the time people spend in it. In other words, if someone stayed in or passed through a public space, they "own" that space more than someone who has never visited it. This

concept is inspired by the temporal ownership of goods that exists in the sharing economy (Chi et al., 2017). In this paper, citizens' temporal ownership is accounted for through the tokenization of GPS data points shared by citizens (see the Results section). Location tokens act as a proof of temporal ownership of urban spaces.

### 3.3. LIQUID DEMOCRACY

Liquid democracy refers to a fluid voting model that combines direct and representative democracies. It has been implemented in politics, namely by eclectic political parties in Germany and Sweden (Kahng et al., 2021). In direct democracy, citizens vote directly on policies, whereas in representative democracy, citizens elect proxies to make policy decisions on their behalf. As societies have grown, direct democracy has become an impractical model. Therefore, representative democracy became commonly adopted by modern societies (Michels & De Graaf, 2010). Liquid democracy offers the advantage of allowing voters to delegate their votes temporarily to other well-informed voters, when their knowledge about an issue is limited, leading to better informed collective decisions (Kahng et al., 2021). Ramos (2015, p. 183) indicates that liquid democracy is part of a range of shifts in governance, technology, and social norms; and indicates a transitive voting process enabled by online systems.

### 3.4. GAMEFUL DESIGN AND MOTIVATION

The goal of informing the proposed framework by game design is to improve the quality of the experience of citizens, by making participation more worthwhile and/or enjoyable (Ryan & Deci, 2020, p. 3). This paper aligns with literature that considers gamification as a design approach of transforming systems to afford gameful experiences, in order to increase user engagement in various fields, including governance (Xiao et al., 2021). Deterding (2014, p. 312) clarifies that mainstream gamification is narrowly understood as the design of software systems and interfaces and needs to be rethought as a holistic design practice of socio-technical systems for motivational affordances. Woodcock & Johnson (2018) distinguish between imposed top-down gamification strategies, described as manipulative; and bottom-up gamification, which is based on the natural inclination towards enjoyable experiences.

The success of a collective process of urbanism depends on voluntary participation and action. Deterding (2014, p. 308) emphasises that what makes games enjoyable is the fact that gameplay is a voluntary action that delivers strong experiences of autonomy, without social or material pressure. Self-determination theory (SDT) is a framework that links intrinsic motivation and autonomous extrinsic motivation to three fundamental psychological needs: autonomy, competence and relatedness (Ryan & Deci, 2020). Autonomy refers to a sense of voluntariness (Ryan & Deci, 2017, p. 10) and ownership of one's actions, free from external control by rewards or punishments; competence describes a feeling of growth and capacity to succeed; and relatedness expresses a sense of belonging and connection (Ryan and Deci, 2020, p.1). According to this framework, collective urbanism could be a process of self-determination if it offers supportive contexts and conditions for autonomous motivational experiences.

### 3.5. DATA DIGNITY AND DATA OWNERSHIP



Data dignity is a concept that advocates for digital platforms' users to have complete control of how their data is used, and for them to be paid accordingly (Getoor, 2019, as cited in Cheng et al., 2021, p.1156). To implement data dignity, Lanier & Weyl (2018) propose to establish organizations called "mediators of individual data"(MIDs), which are groups of volunteers who will be in charge of evaluating the quality of data shared by people and negotiating data royalties with tech platforms. Considering that data dignity and ownership are fundamental for building a democratic participation, these concepts are integrated in the present framework through a tokenization system that records data transactions and generates 'e-coins' in return (see Figure 1).

4. Results

4.1. A FRAMEWORK FOR A COLLECTIVE URBANISM

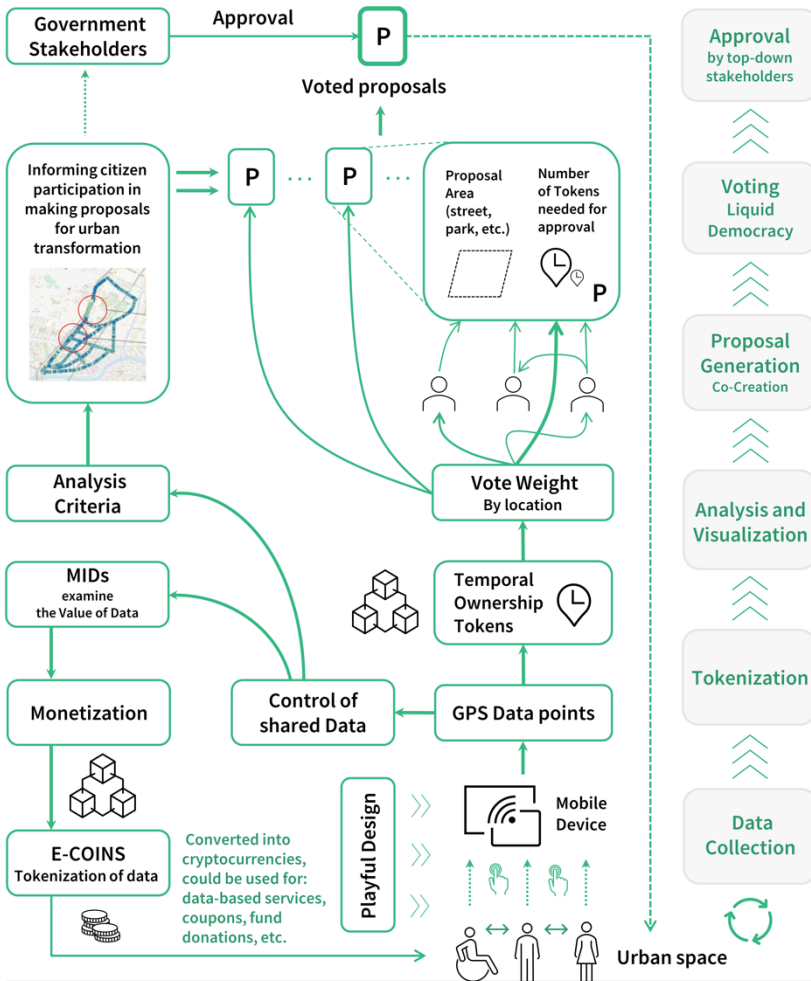


Figure 1: An overview diagram of the gameful collective urbanism framework

As illustrated in Figure 1, the participation process starts from citizens' interaction with a digital device, through which they decide to share their location-based data. By transforming GPS data points into tokens, citizen's temporal ownership of urban spaces is recorded. Citizens may accumulate variable amounts of tokens depending on how much time they spend in an urban space. When the proposal generation phase is concluded and the voting process begins, the collected tokens weigh citizens' votes. In this framework, time is the initial voting power that is equally given to everyone. The time spent in an urban space generates the right of the citizen to intervene in its transformation (Sakai, 2017, p. 72). When voting for urban proposals, only the tokens that citizens have collected within the proposal areas can be used for the voting. At this stage, proposal areas are defined according to urban zoning.

During the liquid voting process, citizens can choose to allocate their tokens to as many proposals as they want, expressing their various degrees of agreement or prioritization. In addition, they can decide to give a share of their temporal ownership of public space to another citizen. As a result, citizens can either choose to: delegate their vote partially or fully to someone else when they judge that the latter is knowledgeable about a specific issue they are not familiar with; or vote directly when they think they have enough knowledge and a clear opinion they want to express.

The present framework includes two tokenisation systems (two blockchain systems). The first is the tokenization of temporal ownership, whereby tokens are converted into a vote weight. These tokens are not monetized as they represent the various degrees of voting rights of citizens. The second tokenization system is the one that converts the value of the data shared by citizens into what "e-coins" (Figure1). These e-coins can be monetized in data marketplaces.

#### 4.2. EARLY PROTOTYPING OF A CASE STUDY USING E-BIKES

The first experiment of this framework is a case study that uses electronic bikes as the data collection device. At the time of writing this paper, a short test of the e-bikes was carried out, with the objective of conducting a preliminary data analysis and studying data visualization on the user interface (Figure 2, (1)). The analysis criteria are still under development. The test took place in Yokohama city in Japan, from 01/10/2020 to 06/11/2020, during which around 30 bikes were used. The bikers were mainly commuters from a train station, using e-bikes for transportation and shopping. Logs-in cannot be disclosed due to business terms related to the project.

Figure 2 displays prototypes of three main interfaces of the mobile application bike riders are expected to use to participate in designing urban spaces. The first interface (1) allows citizens to access an analysis of their riding behaviours and compare it to the analysis of the collective data shared by the community. The second interface (2) informs about the levels of temporal ownership of urban space. The third interface (3) shows the dynamic process of liquid voting on urban projects. It informs users about the amount of vote tokens other citizens possess as an indication of their knowledge about an urban space. Although this project is still under development, Interface design aims at encouraging a playful mood in the users, when interacting with the application.

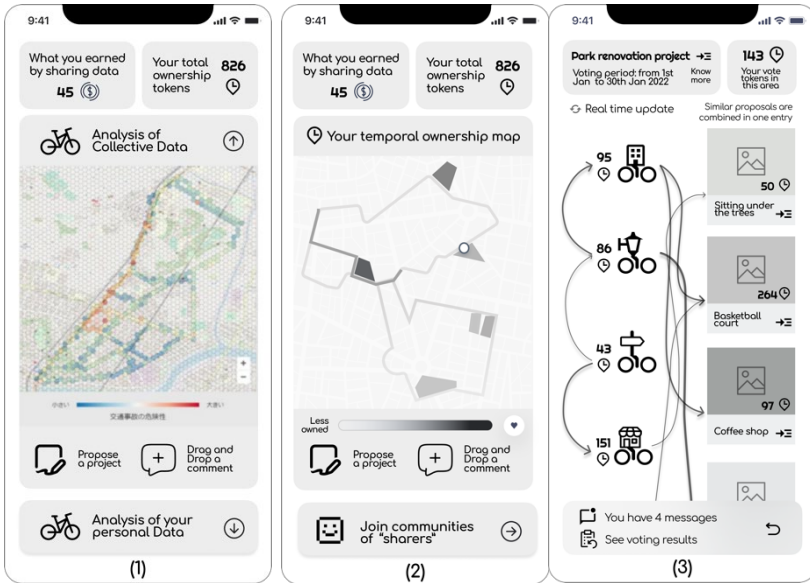


Figure 2: Early Prototyping of user interface: (1) Data visualization, (2) Temporal ownership, (3) Open voting process

**5. Discussion**

**5.1. FOSTERING COLLECTIVE INTELLIGENCE**

The proposed framework's main contribution is the combination of the concepts of temporal ownership of public space and liquid democracy, to create a voting model that fosters collective intelligence. On the one hand, temporal ownership based on the tokenization of location data, accounts for the usage of urban space, which is a valuable indication of local knowledge. Liquid voting on the other hand, allows citizens to benefit from each other's knowledge and expertise of the everyday realities of urban spaces, a knowledge that is crucial for improving the act of urban planning and its consequences on those inhabiting urban spaces (Jon, 2021). This paper focuses on the final voting stage to stress the importance of a generative citizen participation that leads to urban transformation. The proposed framework, however, incorporates citizen participation in both civic learning and collective reflection activities (Devisch et al., 2016), such as debates and workshops. These activities generate rich qualitative and quantitative data that needs to be structured into proposals, to translate into urban actions.

**5.2. BUILDING LOCAL COMMUNITIES**

This framework has the potential of harnessing an intentional interaction with urban space due to the dynamic nature of the voting system. Citizens can change their votes endlessly and keep increasing the number of their tokens in an urban space by simply visiting it more often within the voting period. Hence, this could result in creating

"collectives" (Jon, 2021) where the source of solidarity is the physical togetherness and the common experience of an urban space, thereby multiplying chances of social interaction and knowledge sharing. This framework bridges between virtual and local communities, by bringing together in an urban space, groups of citizens who are interested in having a say on how to improve it, leading to urban interventions that reflect daily realities and consider the scale of the neighbourhood. Respectively, this framework could promote sustainable bottom-up urbanization based on participatory and inclusive planning.

### 5.3. A GAMEFUL SOCIO-TECHNICAL SYSTEM

The objective of informing the present framework by game design is to harness an autonomous motivation to participate in collective decision-making and foster social interaction within local communities. Based on the game design framework of Mechanics, Dynamics, and Aesthetics (MDA) (Hunicke et al., 2004), the collective urbanism framework presented here has the potential of generating gameful dynamics and aesthetics. Firstly, during the open and dynamic voting process, citizens have access to updated voting results, they can continuously increase their temporal ownership tokens while carrying out daily life activities, and they are able to modify the allocation of their tokens as many times as they want during the voting period. Despite the fact that these dynamics are 'serious' aspects of the decision-making process, which objective is to activate the usage of urban space and multiply social encounters; they hold the potential of creating a playful interaction between the citizens and the system, by triggering a strategic behaviour and an autonomy of action, which aligns with SDT theory of motivational design (Ryan & Deci, 2017). These dynamics could also produce game aesthetics such as discovery and control. Secondly, based on the concept of temporal ownership, citizens can establish collectives that are centred around urban spaces. They can join hybrid communities (online and offline linked through location data) of 'sharers' or 'owners'. These dynamics could generate game aesthetics such as fellowship and expression. As a result, the proposed framework holds a great potential in creating a gameful experience of collective urbanism.

Thiel et al. (2016) indicate that older citizens could be opposed to gamification. However, further investigation is required, as there is limited research that focuses on the effect of gamification on older adults in various domains (Koivisto & Malik, 2021). With that being said, the gamefulness of the presented framework is not dependent on external game elements, but on the dynamic nature of the collective decision-making process. It could be playful or serious depending on the citizen's perspective when engaging the application. Furthermore, this framework can be adjusted to different implementation contexts, target groups and participation objectives, including adapting the degree of gamification, from playful interactions to full-fledged serious games.

### 5.4. LIMITATIONS AND FUTURE DEVELOPMENT

As described above, the right to vote about urban interventions is weighted by temporal ownership tokens. However, considering the following scenario of a person with disabilities, living right next to an open space, but unable to visit it. They would have no tokens to participate in deciding how to transform that open space despite the

possibility of being negatively impacted by the consequences of the decision (noise disturbances for example). Therefore, to improve the current voting model, the right to decide how to improve an urban space could be determined by a combination of temporal ownership tokens and ‘proximity to project’ tokens, to consider the potential repercussions of urban projects on individuals. Furthermore, the risk of potential manipulation of the vote weights for personal benefits is acknowledged. For instance, gathering a large amount of tokens by placing the data collection device in a location and leaving. This could be countered by adopting an authentication system that verifies the validity of the collected tokens. A more detailed description of these models is not covered in this paper, but it is planned for future research. At this stage, the potential of this framework is still hypothetical, and experimentation is needed to reveal more about its implications. Nonetheless, this paper presents a novel form of collective reflection on urban space, supported by digital tools, to create a dynamic model of urban decision-making that continuously adapts to the needs of local communities.

## 6. Conclusion

This paper contributes to existing research by proposing a technology-supported framework for a gameful collective urbanism, incorporating the concepts of temporal ownership of urban spaces and liquid democracy, to design a decision-making model that promotes local knowledge and collective intelligence. This framework is not intended to be comprehensive, but rather it can be extended in various ways. It addresses important discussions about the value of local knowledge in urban decision making and could contribute to the development of digital applications that support citizen participation in deciding how to design their living environments, which is fundamental for the creation of sustainable cities and communities.

## References

- Becker, S. (2020) Tokenization 101: How tokenization of physical assets enables the economy of everything. Riddle&Code - The blockchain interface company. Retrieved December 4, 2021, from <https://www.riddleandcode.com/tokenization-101>.
- Cheng, L., Varshney, K.R., Liu, H., (2021). Socially Responsible AI Algorithms: Issues, Purposes, and Challenges. *The Journal of Artificial Intelligence Research*, 71, 1137–1181. <https://doi.org/10.1613/JAIR.1.12814>
- Deterding, S. (2014). Eudaimonic Design, or: Six Invitations to Rethink Gamification. In Fuchs, M., Fizek, S., Ruffino, P., Schrape, N. (Eds.), *Rethinking Gamification*, (305-331). Meson press, Hybrid Publishing Lab.
- Devisch, O., Poplin, A., & Sofronie, S. (2016). The Gamification of Civic Participation: Two Experiments in Improving the Skills of Citizens to Reflect Collectively on Spatial Issues. *Journal of Urban Technology*, 23(2), 81–102. <https://doi.org/10.1080/10630732.2015.1102419>
- Ertiö, T.-P. (2015). Participatory Apps for Urban Planning-Space for Improvement. *Planning Practice & Research*, 30(3), 303–321. <https://doi.org/10.1080/02697459.2015.1052942>
- Fan, M., Heikkilä, J., Li, H., Shaw, M., & Zhang, H. (2017). *Internetworked World*. Springer.
- Hunicke, R., Leblanc, M. G., & Zubek, R. (2004). MDA: A Formal Approach to Game Design and Game Research. In *Proceedings of the AAAI Workshop on Challenges in Game AI* (Vol. 4, No. 1, pp. 1722).

- Jiang, J., Spencer, L., & Werner, L. C. (2021). Public Gratification Palace: A Framework for Increased Civic Engagement. *Media Architecture Biennale 20*, 131–140. <https://doi.org/10.1145/3469410.3469423>
- Jon, I. (2021). The City We Want: Against the Banality of Urban Planning Research. *Planning Theory & Practice*, 22(2), 321–328. <https://doi.org/10.1080/14649357.2021.1893588>
- Kahng, A., Mackenzie, S., & Procaccia, A. D. (2021). Liquid Democracy: An Algorithmic Perspective. *The Journal of Artificial Intelligence Research*, 70, 1223–1252. <http://dx.doi.org/10.1613/jair.1.12261>
- Koivisto, J., & Malik, A. (2021). Gamification for Older Adults: A Systematic Literature Review. *The Gerontologist*, 61(7), e360–e372. <https://doi.org/10.1093/geront/gnaa047>
- Lanier, J., & Weyl, E. G. (2018). A Blueprint for a Better Digital Society. *Harvard Business School Publishing Corporation*. from [http://eliassi.org/lanier\\_and\\_weyl\\_hbr2018.pdf](http://eliassi.org/lanier_and_weyl_hbr2018.pdf)
- Michels, A., & De Graaf, L. (2010). Examining Citizen Participation: Local Participatory Policy Making and Democracy. *Local Government Studies*, 36(4), 477–491. <https://doi.org/10.1080/03003930.2010.494101>
- Pilemalm, S. (2018). Participatory Design in Emerging Civic Engagement Initiatives in the New Public Sector: Applying PD Concepts in Resource-Scarce Organizations. *ACM Transactions on Computer-Human Interaction*, 25(1), 1–26. <https://doi.org/10.1145/3152420>
- Poplin, A., Shenk, L., Krejci, C., & Passe, U. (2017). Engaging Youth Through Spatial Socio-Technical Storytelling, Participatory GIS, Agent-based Modelling, Online Geogames and Action Projects. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-4/W3, 55–62. <https://doi.org/10.5194/isprs-annals-IV-4-W3-55-2017>
- Ramos J. (2015) Liquid Democracy and the Futures of Governance. In Winter J., Ono R. (Eds.) *The Future Internet* (pp. 173–191). Springer, Cham.
- Ryan, R. M., & Deci, E. L. (2017). *Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness*. Guilford Publications.
- Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, 61, 101860. <https://doi.org/10.1016/j.cedpsych.2020.101860>
- Sakai, Y. (2017). *Bikebump: Collective urban design* [Thesis, Massachusetts Institute of Technology]. from <https://dspace.mit.edu/handle/1721.1/114065>
- Secinaro, S., Brescia, V., Iannaci, D., & Jonathan, G. M. (2021). Does Citizen Involvement Feed on Digital Platforms? *International Journal of Public Administration*, 0(0), 1–18. <https://doi.org/10.1080/01900692.2021.1887216>
- Smith, A., & Martin, P. P. (2021). Going Beyond the Smart City? Implementing Technopolitical Platforms for Urban Democracy in Madrid and Barcelona. *Journal of Urban Technology*, 28(1–2), 311–330. <https://doi.org/10.1080/10630732.2020.1786337>
- Thiel, S.-K., Reisinger, M., & Röderer, K. (2016). “I’m too old for this!”: Influence of age on perception of gamified public participation. *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia*, (pp. 343–346). <https://doi.org/10.1145/3012709.3016073>
- Woodcock, J., & Johnson, M. R. (2018). Gamification: What it is, and how to fight it. *The Sociological Review*, 66(3), 542–558.
- Xiao, R., Wu, Z., & Hamari, J. (2021). Internet-of-Gamification: A Review of Literature on IoT-enabled Gamification for User Engagement. *International Journal of Human-Computer Interaction*, 1–25. <https://doi.org/10.1080/10447318.2021.1990517>

## SENSING THE CITY

*Leveraging geotagged social media posts and street view imagery to model urban streetscapes using deep neural networks*

JAYEDI AMAN<sup>1\*</sup>, TIMOTHY C MATISZIW<sup>2\*</sup>, JONG BUM KIM<sup>3</sup>  
and DAN LUO<sup>4</sup>

<sup>1,2,3</sup>*University of Missouri Columbia*

<sup>4</sup>*The University of Queensland*

<sup>1</sup>*jayediaman@mail.missouri.edu, 0000-0002-6128-8293*

<sup>2</sup>*matisziwt@missouri.edu, 0000-0003-2159-4904*

<sup>3</sup>*kimjongb@missouri.edu, 0000-0001-8762-9554*

<sup>4</sup>*d.luo@uq.edu.au, 0000-0003-1760-0451*

*\*Contributed equally*

**Abstract.** Understanding the relationships between individuals and the urban streetscape is an essential component of sustainable city planning. However, analysis of these relationships involves accounting for a complex mix of human behaviour, perception, as well as geospatial context. In this context, a comprehensive framework for predicting preferred streetscape characteristics utilizing deep learning and geospatial techniques is proposed. Geotagged social media posts and street view imagery are employed to account for individual sentiment and geospatial context. Natural Language Processing (NLP) and computer vision (CV) are then used to infer sentiment and model the visual environment within which individuals make posts to social media. An application of the developed framework is provided using Instagram posts and Google Street View imagery of the urban environment. A spatial analysis is conducted to assess the extent to which urban attributes correlate with the sentiment of social media postings. The results shed light on sustainable streetscape planning by focusing on the relationship between users and the built environment in a complex urban setting. Finally, limitations of the developed methodology as well as future directions are discussed.

**Keywords.** Urban Sustainability; Data Mining; Pedestrian Sentiments; Transportation Behaviour, Street Level Imagery; Transformers; SDG 11.

### 1. Introduction

Two-thirds of the world's population, 6.5 billion people will live in urbanized areas by 2050. As communities have become more global, dynamic, and transitory, sustainable

urban development can be difficult to achieve unless the way cities are planned and managed changes dramatically. Urban streetscapes are thought to play an essential role in influencing social interactions (Salesses et al., 2013). The design, condition, and spatial configuration of urban features such as buildings, traffic infrastructure, and parks can trigger a variety of collective and individual human experiences in this sense. Thus, urban planning and regulatory agencies are investing extensively in enhancing the condition of the streets in order to improve how they are experienced by individuals. To this end, it is increasingly recognized that the linkage between individuals' sentiment and the location(s) associated with sentiment is an important consideration. The spatial context explicitly enables and stimulates human activity. In return, these activities impact the form and perception of streetscapes (Batty et al., 2010). As a result, assessing people's perceptions of urban streetscapes could provide important insights on street semantics, leading to more informed community planning decisions.

Various approaches for gauging perception of streetscapes in urban studies have been proposed. Traditionally, data on humanistic experience of urban environments have been by way of field surveys, interviews, and questionnaires (Montello et al., 2003). Such modes of data collection often entail complex research protocols to ensure use/applicability over wide geographic areas or time periods. Smart city technologies such as drones, street view imagery and social media data have recently facilitated the ability to sense the urban streetscape in the field of urban sustainability research (Ilieva and McPhearson, 2018). In particular, Guerrero et al. (2016) explored how social media posts can be used to learn about how people interact with and perceive urban green spaces, and how this can be used to inform future sustainable infrastructure planning. More recently, Marti et al. (2017) used social media posts and urban cartographies to show what features may make open public spaces, such as urban plazas, more successful than others.

These newer approaches provide information on a variety of facets of urban processes and the dynamics thereof. In order to make full use of more robust streams of data though, there is a need to identify mechanisms for linking them with indicators of human perception of urban streetscapes. Furthermore, detailed accounts of the use of geolocated spatiotemporal social media data and street view imagery for urban perception analysis are relatively rare. To this end, a novel analysis framework is proposed to better integrate individual indicators of urban sentiment with geospatial context in predicting preferred streetscape and identifying linkages between streetscape features and individual sentiment. Open-source geospatial datasets such as images of urban features (Google Street View), roads, and public social media (Instagram) postings are analysed using deep learning algorithms to quantify individual sentiment relative to geospatial context. The developed framework is applied to a medium sized urban area to demonstrate its applicability.

## **2. Analytical Framework**

The proposed analytical framework is depicted in Figure 1. In Phase 1, information that can serve as indicators of individual sentiment is collected and processed. Social media postings are a rich source of data in this respect given that they can contain texture and tone reflective of an individual's sentiment. In many instances, social media postings also contain indicators of the geographic location (latitude, longitude) as well



as temporal references (e.g., date/time stamps). Once relevant posts have been collected, they can be processed by a natural language processing (NLP) algorithm, such as a 'text categorization' technique, to infer the nature of the poster's sentiment. Hot spot analyses can then be performed to assess the extent to which spatial dependency among posts along streetscapes may exist.

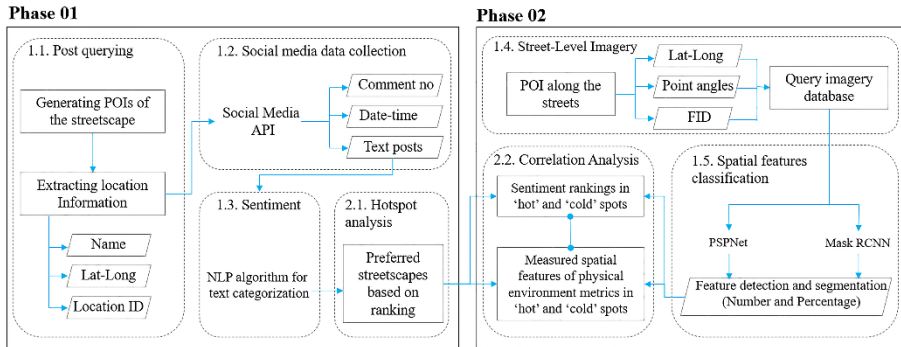


Figure 1. Overview of the methodology

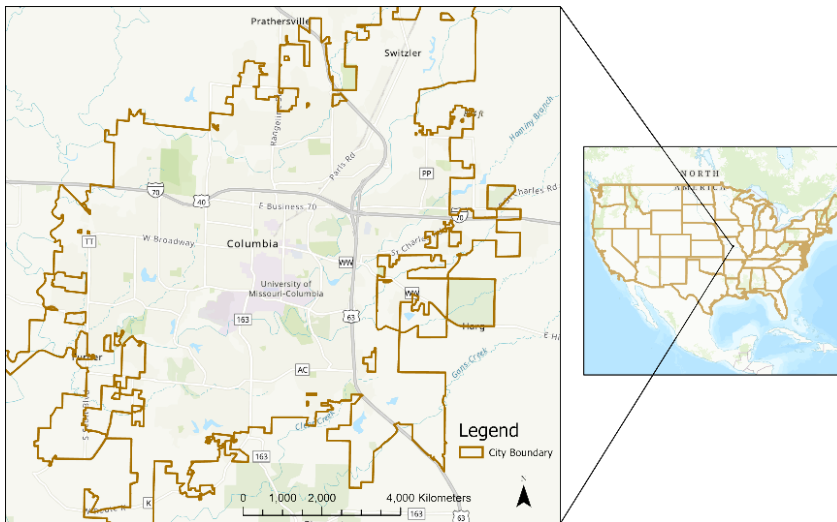


Figure 2. Study Region

In Phase 2, metrics of urban form such as enclosure, human dimension, and complexity in determining sentiment are derived for street segments. This can be accomplished via analysis of driver/pedestrian-perspective imagery (e.g., Google Street View (GSV)) for a set of points of interest (POI) along city streets. Spatial features, such as the sky, vegetation, roads, buildings, and sidewalks, can then be identified and quantified using computer vision methods such as PSPNet and Mask RCNN. Finally, level of enclosure, visibility, and complexity can be evaluated with respect to the sentiment metrics.

### 3. Application

To demonstrate the analysis framework (Figure 1), an application to the City of Columbia, MO, USA was examined. Columbia (Figure 2) is a medium sized city (population of 126,254) in the midwestern US which hosts several major universities/colleges (United States Census Bureau, 2020).

#### 3.1. SOCIAL MEDIA POSTING LOCATIONS

In this application, public posts from the Instagram social media platform are used to represent expressions of sentiment for the city's streetscapes. Several major roads traversing the city were selected for analysis and 16 POIs along these roads were generated to assist in the search for proximate Instagram posting locations. The coordinates of the POIs were then used to search for Instagram posting sites using the *instagrapi* (<https://github.com/adw0rd/instagrapi>) Python package as shown in Figure 3. Querying the locations of the 16 POIs resulted in the identification of 135 unique Instagram posting sites in the study region.

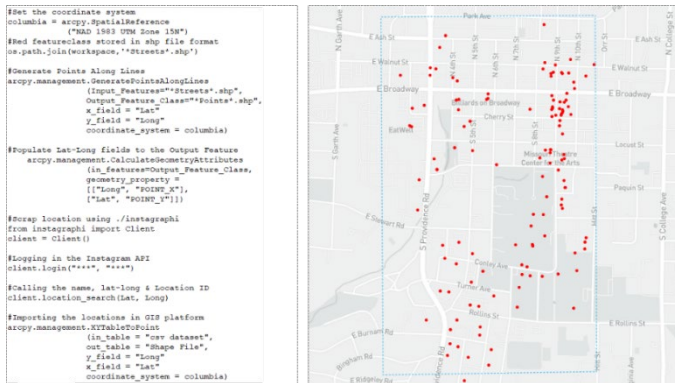


Figure 3. Code and mapping of Instagram posting locations

#### 3.2. GEOTAGGED SOCIAL MEDIA DATA COLLECTION

Relevant social media posts were obtained from InstaLoadGram, a third-party data provider. A total of 111 locations (out of the initial 135 sites) featured usable public posts, yielding a total of 63,861 posts spanning a six-year period (1/1/2015 - 20/11/2021). A few example postings are shown in Table 1.

Location Name	Lat	Long	Caption	Date	Likes/ Comments
Francis Quadrangle	38.946111	-92.328889	Just some undergrads, nothing to see here. Keep scrolling	Nov 14th, 2021, 1:54:14 am	81/02
Francis Quadrangle	38.946111	-92.328889	No better way to introduce our new mellophone band family than with the classic pose! 🤝👍	Aug 16th, 2021, 7:27:08 am	48/00
Francis Quadrangle	38.946111	-92.328889	So excited to welcome our new members to the MU Spirit Executive family! FIGHT TIGER & MIZ!	May 6th, 2021, 4:30:08 pm	35/00
Francis Quadrangle	38.946111	-92.328889	Soaking up the sun	Apr 26th, 2021, 3:42:02 am	19/00

Table 1. Example Instagram posts acquired

### 3.3. SENTIMENT ANALYSIS

Individual sentiment can be inferred using a well-trained NLP model to score text-based posts. For example, 'safe', 'lively', 'bored', 'affluent', 'gloomy', and 'beautiful' are six commonly used sentiment markers. In order to reliably classify post by sentiment, the geotagged postings were processed using a normalized rating system (Zhang et al., 2018). Several deep learning (DL) models were investigated in this respect. Ultimately, the pre-trained Transformers model was selected given its ability to recognize the context giving meaning to each word in the sentence, allowing for more parallelization and shorter training and prediction durations (Colón-Ruiz and Cristóbal, 2021). Specifically, the Hugging Face API, a Bidirectional Encoder Representations from Transformers (BERT) based pre-trained Transformers model, was utilized to predict multilingual sentiment from the Instagram posts given the capacity of BERT to create contextually appropriate word embeddings (Devlin et al. 2019; Colón-Ruiz and Cristóbal, 2021). The model, trained using the BERT approach on one million human evaluations, automatically translates each post into the normalized rating ranging from 1 to 5 stars (Very Negative = 1 star; Negative = 2 stars; Neutral = 3 stars; Positive = 4 stars; Very Positive = 5 stars), an example of which is depicted in Figure 4 (Hugging Face, 2021).

```

ModelName = 'nlpopen/bert-base-multilingual-uncased-sentiment'
TextClassifier = pipeline('sentiment-analysis', model = ModelName, Device = 0)

SentimentResults = TextClassifier(
    ["Just some undergrads, nothing to see here. Keep scrolling",
     "No better way to introduce our new meliphone band family than with the
     classic pose! 📸👏",
     "So excited to welcome our new members to the MU Spirit Executive family!
     FIGHT TIGER & MIZ!",
     "Soaking up the sun"])

o {'label': '3 stars', 'score': 0.4023219347000122}
o {'label': '5 stars', 'score': 0.32480156421661377}
o {'label': '5 stars', 'score': 0.8477996587753296}
o {'label': '4 stars', 'score': 0.4071483314037323}
    
```

Figure 4. Example of sentiment scoring input/output

### 3.4. STREETVIEW IMAGERY

GSV images were retrieved for spatial feature classification. GSVs provide street-level and profile views of the urban environment and thus represent what individuals navigating the city may encounter.



Figure 5. POIs along streets (left panel); GSV images per point (right panel)

Overall, 341 POIs along roads in the study region were generated to represent locations that would be traversed by individuals. Because the intention was to gather consecutive images of urban features, the POIs were collected at 30-meter intervals and then utilized to retrieve GSV images using Google API (Figure 4 - left panel). Four different viewing perspectives were considered for each POI  $i \in I$  given the sequence in which the POIs would be traversed along the roads: a) in the direction of the next POI  $i+1$ , b) 90.0 degrees from POI  $i+1$ , c) 180.0 degrees from POI  $i+1$ , and d) 270.0 degrees from POI  $i+1$  as illustrated in Figure 5 (right panel). Whereas the horizontal field of view was kept at 90.0 degrees, the "pitch" of the returned 800x400 pixel images was set to 0.0 degrees. In total, 1,364 images were retrieved for the sampled POIs.

### 3.5. SPATIAL FEATURE CLASSIFICATION

Enclosure, human scale, and complexity measurements are environment metrics thought to be strongly connected to the pedestrian experience but challenging to assess at urban scale (Miranda et al., 2021). Two prominent image detection and semantic segmentation models, PSPNet and Mask R-CNN, were used to classify features in each image. Enclosure and human scale measures were evaluated using PSPNet. The model was trained using the 150-category ADE20k dataset (Miranda et al., 2021). The proportion of an image classified as sky, walls, fences, trees, and buildings was used to represent the level of enclosure that persons were likely to experience. Human scale was estimated as the percentage of an image classified as sidewalk or road. The level of complexity present in the urban environment was characterized as the number of people, bicycles, motor bikes, cars, and streetlights present in each image as computed using the Mask R-CNN (Qiu et al., 2021).

## 4. Results

Computations for both the sentiment analyses and spatial feature classification were performed using the Google Colab Pro platform with a system configuration of Nvidia

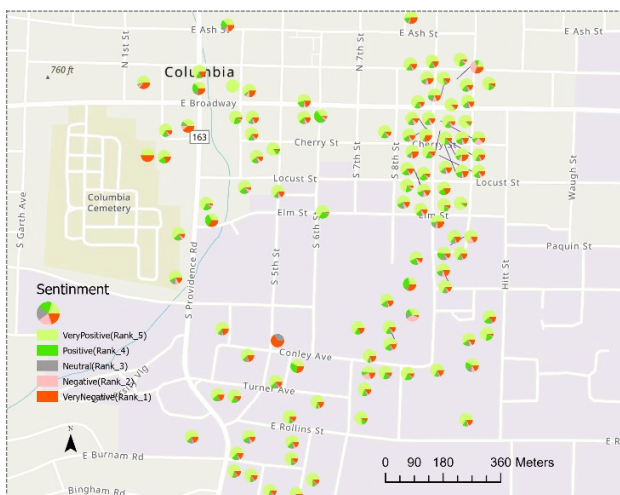


Figure 6. Summary output of sentiment analysis

Tesla P100-PCI-E GPU and 16GB RAM. Assessing the sentiment associated with each of the 63,861 postings took 35,756 seconds of computational time. Figure 6 summarizes the proportion of each sentiment classification at each posting site. At many posting sites, a majority of the posts are classified as very positive. There is quite a bit of variation though in the spatial distribution of negative, very negative sentiment, with locations having higher proportions of such sentiment located in the Northwest, Northeast, and Central portions of the study region.

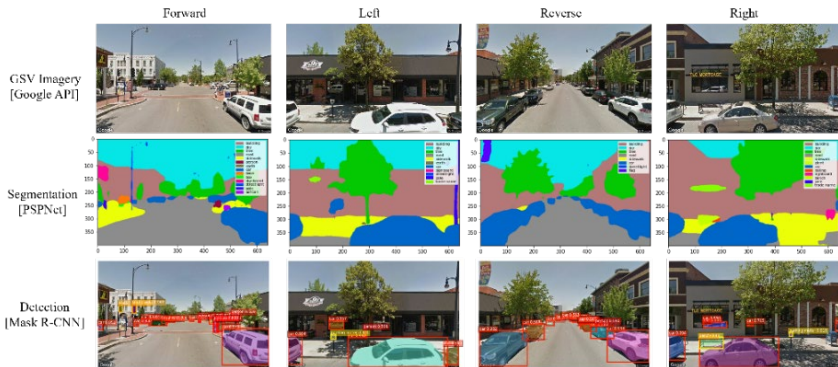


Figure 7. Spatial feature classification

The retrieval and classification of the 1,364 GSV images required 10,912 seconds of processing time. Figure 7 illustrates the output of the four images processed for one example POI. The first row depicts the images as retrieved. The second row depicts the percentage of pixels associated with different spatial features. The summary metrics associated with this are provided in the first and second sections of Table 2. The third row displays the number of complexity features identified in each image. The summary metrics for this are provided in the third section of Table 2.

ID: 135	First section Enclosure					Second section Human Scale		Third section Complexity				
	Sky	Wall	Fence	Tree	Building	Roads	Sidewalks	People	Bicycles	bikes	Cars	Streetlights
Forward	0.430	0.002	0.021	0.105	0.222	0.385	0.085	3	1	0	8	4
Left	0.222	0.000	0.002	0.158	0.238	0.010	0.012	1	0	2	4	1
Reverse	0.330	0.004	0.000	0.025	0.276	0.320	0.002	2	0	0	11	1
Right	0.150	0.010	0.005	0.071	0.484	0.103	0.183	1	0	0	2	0

Table 2. Spatial features calculation based on segmentation and detection

A heatmap was performed based on the proportion of negative sentiment rankings associated with each posting site (Figure 8). The heatmap (Figure 8-left panel) indicates several areas of denser proportion of negative sentiments (purple circle-marked). Improved can be found in areas to the Southeast (enclosed in a red circle). Similarly, an analysis of spatial autocorrelation, the Getis-Ord  $G_i^*$ , of the proportion of positive postings at each location (Figure 8-right panel) reveals the presence of statistically significant spatial autocorrelation at several sites in the study region.

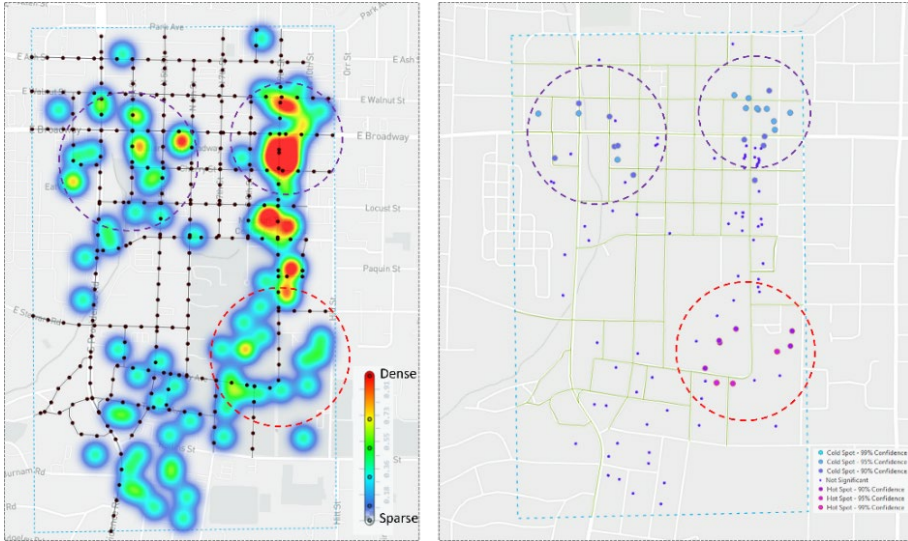


Figure 8. Heatmap based on negative posts (left panel); Hotspot analysis (right panel)

In particular, 19 sites in the South-eastern portion of the region were found to exhibit statistically significant positive spatial autocorrelation at the 0.90 confidence level or above (10 at the 0.95 confidence level or above, 4 at the 0.99 confidence level or above). Two areas to the North were found to exhibit statistically significant spatial autocorrelation of lower proportion of posts classified as positive at the 0.90 confidence level or above (11 at the 0.95 confidence level, 12 at the 0.99 confidence level or above).

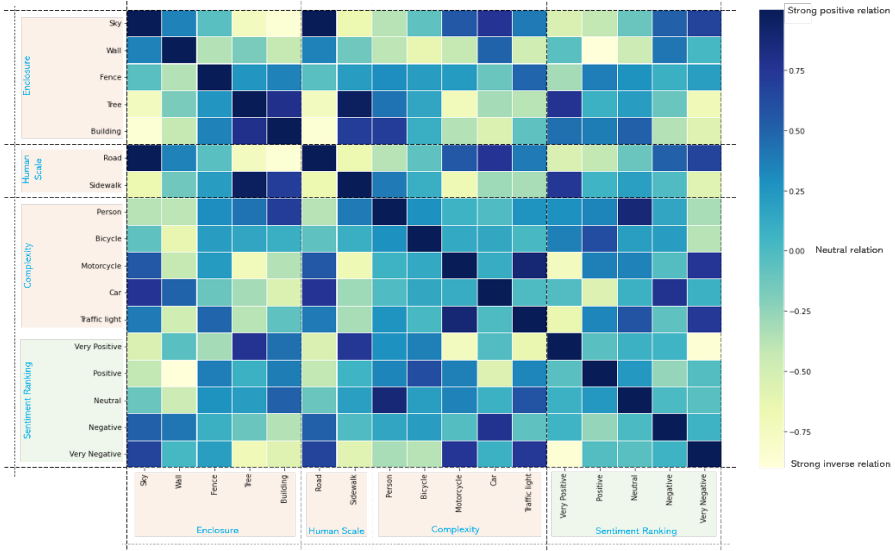


Figure 9. Correlation matrix of streetscape features and sentiment ratings

Finally, the characteristics of the streetscape were evaluated for posting locations associated with ratings of hot and cold spots to establish the extent to which streetscape features may be linked with individual sentiment. The results of the correlation analysis between streetscape features and sentiment ratings are summarized in Figure 9 and indicate that the streetscape features are strongly linked with individual sentiments.

In general, locations along streets of positive posting are strongly associated with the presence of vertical structures such as trees and buildings. Negative postings are more often associated with locations having fewer enclosure components and a greater amount of open sky. Both findings are consistent with those of Miranda et al (2021). Additionally, the findings show that streets with better pedestrian infrastructure, such as sidewalks, significantly contribute to positive sentiment, which is consistent with the human scale concept. Finally, areas with a higher concentration of human activity - people and cyclists - are directly associated with positive sentiments, whereas urban areas with heavy traffic - cars and motor vehicles - are inversely related to positive sentiments, as noted by Ye et al (2021).

## 5. Discussion and Future Works

Making urban human settlements more inclusive, safe, and resilient, as well as encouraging healthy lifestyles and well-being at all ages, are critical to long-term urban sustainability. The way people perceive and value places, infrastructures, and events in their daily lives has far-reaching implications for sustainability planning and implementation efforts. To this end, the aim of this research is to provide vital insights into street semantics to promote more informed community planning decisions (what SDG 11 aims to achieve by 2030). The paper details a framework for linking individual sentiment to characteristics of the streetscape. Open-source social media postings and are processed using deep learning methods to infer sentiment and characterize visibility at locations along a transportation system to make these associations possible. The framework is applied to a small case study to illustrate the procedures involved and demonstrate the applicability of the process.

There are various decision points, constraints, and extensions to this study that might be worthwhile directions for future research. Because of Meta's data policy at the time of the study, the investigators had limited access to social media. Furthermore, the sentiments expressed in the postings may not directly relate to the streetscape in all instances. Further, working with social media postings entails the complex task of interpreting the opinions of multiple individuals. In this study, a pre-trained model was utilized for this task. However, a more target-specific transformers model could be used instead, making use of a large amount of Instagram data, segmenting location-related data, and then training the transformers model for more accurate and relevant sentiment analysis. Given that social media postings are generally accompanied by a time/date stamp, it would also be intriguing to analyse sentiment in a more temporally disaggregated form if more frequent data on urban environment were available.

In the future, real-time social media datasets that incorporate both space and time into analysis, such as Emerging Hot Spot Analysis (EHSA) as a spatiotemporal application of the Getis-Ord  $G_i^*$  statistical analysis, would be more robust and efficient in analysing urban streetscape. Finally, developing a website-based user-friendly interface for the city residents can be quite beneficial in obtaining real-time responses.

## Acknowledgements

Research was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-21-2-0276. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein. Finally, the authors would like to express their gratitude to Raj Roy for sharing his expertise in the Transformers DL model for NLP.

## References

- Batty, M., Hudson-Smith, A., Milton, R., & Crooks, A. (2010). Map mashups, Web 2.0 and the GIS revolution. *Annals of GIS*, 16(1), 1–13. <https://doi.org/10.1080/19475681003700831>
- Colón-Ruiz, C., & Segura-Bedmar, I. (2020). Comparing deep learning architectures for sentiment analysis on drug reviews. *Journal of Biomedical Informatics*, 110, 103539. <https://doi.org/10.1016/j.jbi.2020.103539>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *ArXiv:1810.04805 [Cs]*. <http://arxiv.org/abs/1810.04805>
- Guerrero, P., Møller, M. S., Olafsson, A. S., & Snizek, B. (2016). Revealing Cultural Ecosystem Services through Instagram Images: The Potential of social media Volunteered Geographic Information for Urban Green Infrastructure Planning and Governance. *Urban Planning*, 1(2), 1–17. <https://doi.org/10.17645/up.v1i2.609>
- Hugging face API. (n.d.). Hugging face API. Retrieved November 22, 2021, from <https://huggingface.co/nlptown/bert-base-multilingual-uncased-sentiment>
- Ilieva, R. T., & McPhearson, T. (2018). Social-media data for urban sustainability. *Nature Sustainability*, 1(10), 553–565. <https://doi.org/10.1038/s41893-018-0153-6>
- Montello, D. R., Goodchild, M. F., Gottsegen, J., & Fohl, P. (2003). Where's downtown? Behavioral methods for determining referents of vague spatial queries. *Spatial Cognition and Computation*, 3(2–3), 185–204. <https://doi.org/10.1080/13875868.2003.9683761>
- Qiu, W., Li, W., Liu, X., & Huang, X. (2021). Subjectively Measured Streetscape Perceptions to Inform Urban Design Strategies for Shanghai. *ISPRS International Journal of Geo-Information*, 10(8), 493. <https://doi.org/10.3390/ijgi10080493>
- Salazar Miranda, A., Fan, Z., Duarte, F., & Ratti, C. (2021). Desirable streets: Using deviations in pedestrian trajectories to measure the value of the built environment. *Computers, Environment and Urban Systems*, 86, 101563. <https://doi.org/10.1016/j.compenvurbsys.2020.101563>
- Salesses, P., Schechtner, K., & Hidalgo, C. A. (2013). The Collaborative Image of The City: Mapping the Inequality of Urban Perception. *PLOS ONE*, 8(7), e68400.
- United States Census Bureau. (2020). *Census of Columbia [Data]*. <https://data.census.gov/cedsci/table?q=columbia,%20mo&tid=DECENNIALPL2020.P1>
- Ye, C., Zhang, F., Mu, L., Gao, Y., & Liu, Y. (2021). Urban function recognition by integrating social media and street-level imagery. *Environment and Planning B: Urban Analytics and City Science*, 48(6), 1430–1444. <https://doi.org/10.1177/2399808320935467>
- Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H. H., Lin, H., & Ratti, C. (2018). Measuring human perceptions of a large-scale urban region using machine learning. *Landscape and Urban Planning*, 180, 148–160. <https://doi.org/10.1016/j.landurbplan.2018.08.020>



# IMPACT OF COVID-19 ON ASSOCIATIONS BETWEEN LAND USE AND BIKE-SHARING USAGE

YUDI AN<sup>1</sup>

<sup>1</sup>*University of California, Los Angeles.*

<sup>1</sup>*anyudi@ucla.edu, 0000-0003-4324-4526*

**Abstract.** Bike-sharing as a human-centred, zero-emission, sustainable, alternative, and easily accessible transport mode has been implemented globally and consistently contributing to communities and the environment by alleviating consumption of natural resources, traffic congestion, and air pollution, which is considered a solution for future cities. The appearance of Covid-19 significantly impacts public transportation modes, including the bike-sharing system. The intention of this study was to investigate the spatiotemporal impact of the Covid-19 pandemic on associations between urban factors and bike-sharing usage in Los Angeles, United States, by analysing a sizeable actual trip dataset and employing geographically weighted regression (GWR) models. GWR was conducted for examining the varying spatial association between bike infrastructure, public transport, and urban land use factors, and bike-sharing trip volume. The results indicated that bike-sharing usage significantly decreased during the pandemic and essential service as restaurant was found consistently and positively associated with bike-sharing use. GWR provided clear spatial patterns of bike usage based on urban land use and big user databases. The outcomes of this study could inspire policymakers and shared mobility operators to support these safe, sustainable transport alters (such as rebalancing bike stations), help city resilience, and shape a sustainable future of mobility in the post-Covid-19 era.

**Keywords.** Bike-Sharing; Covid-19; Land Use; Geographically Weighted Regression; Big Data; SDG 11.

## 1. Introduction

As the bike-sharing system trend rapidly spreads worldwide, this flexible, affordable, easily accessible public mobility has become a solution for the first/last minutes of a daily trip, helping to facilitate the intermodal connection between users' homes and mass public transits within a short mediating distance. (Lahoorpoor et al., 2019; Wu et al., 2019). Over the past decade, many cities have implemented bike-sharing systems considering the benefits of urban sustainability and human health. Using big data-based analysis, Zhang and Mi (2018) identified that bike-sharing helped Shanghai save 8,358 tonnes of gasoline and reduce carbon dioxide and nitrogen oxide emissions by 25,240 tonnes and 64 tonnes respectively, in 2016. Bike-sharing system as an alternative to

auto-commuting provides environmental and economic benefits which could contribute to future sustainable city strategies. (Martens, 2006).

After the appearance of the Covid-19 pandemic, it has significantly impacted urban mobility modes and travel behaviours around the world (Zafri et al., 2021). The travel restriction and lockdown measurement drastically reduced global and domestic traffic volumes and demands of public transportation (Nikiforiadis et al., 2020; Jenelius and Cebecauer, 2020). Aloi et al. (2020) found that public transport users dropped 93% in Santander, Spain. Much empirical research (Zafri et al., 2021; Abdullah et al., 2020; Jenelius and Cebecauer, 2020; Kubal'ák et al., 2021) points out that the outflow of public transport contains private cars, walking, and cycling. Recent studies examining Covid-19 impacts on traveller's attitudes (Nikiforiadis et al., 2020; Mujahed, 2021) and bike-sharing usage are based on a limited amount of survey results and global quantitative analysis. Respondents of studies showed positive attitudes to bike-sharing transport mode considering hygiene and safety. Few previous studies have quantitatively explored the spatiotemporal impacts on bike-sharing usage by applying actual trip data before and during the pandemic. Wang and Noland's (2021) initial research investigated the bike-sharing usage pattern changes during the pandemic with Citi bike trip data and found reductions of trip volumes associated with mixed, commercial, and other public land use in September 2019 and 2020. Pase et al. (2020) affirmed that socioeconomic factors and bike lane networks contribute to bike-sharing traffic in an area. The lack of related studies is insufficient to understand the changes of quantitative relationships between land use factors and bike-sharing usage and the spatial visualization of associations.

This paper aims to address the research gap by analysing the Covid-19 impacts on the association between urban land use and shared bike ridership by comparing actual trip data during pre-Covid-19, highest morbidity period, and after vaccination implementation in Los Angeles County, United States. The study addresses several questions on how shared trip volumes changed compared to pre-Covid-19 time; how land-use factors were associated with ridership in three study periods; how the associations vary between different locations. The research employed global regression models and Geographically Weighted Regression models to investigate shared mobility patterns further to quantify trip information and provide spatial visualization. This presented method of modelling allows researchers to visualize raw data from open-source platforms and helps policymakers to improve bike-sharing system implementation regarding spatial variations to create more efficient and smarter cities.

## **2. Methodology**

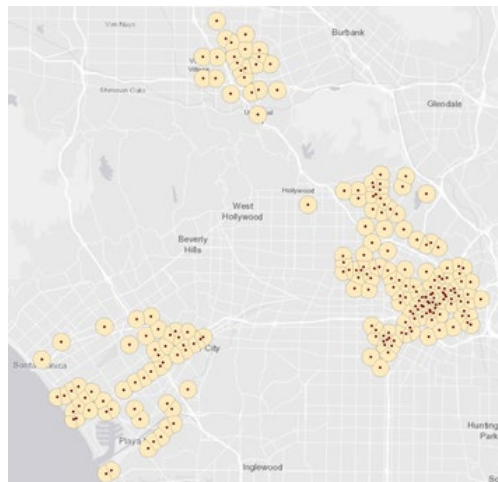
### **2.1. STUDY AREA**

The study area covers Downtown Los Angeles (DTLA), Santa Monica, and North Hollywood, including 214 bike-share stations. All three neighbourhoods have flat road slopes, high population density, young age population, and influential economic activities in Los Angeles County, with plenty of time and sufficient sunshine having a minor weather influence on bike riding relatively. Moreover, Los Angeles County experiences the most severe Covid-19 pandemic having most Covid-19 cases nationwide and is one of the first few counties issuing "Stay at home" orders. With

unique pandemic experience, multimodal transportation implementation, and data availability, Los Angeles County is a valuable study area with bike-sharing system development potentials.

## 2.2. DATA RESOURCE

The shared bike trip data was provided by Bicycle Transit Systems managing the Metro Bike Share program in Los Angeles and making trip data available to the public through their open portal. The raw dataset was downloaded in April 2021, containing information about the trip origin, destination, start time, end time, duration, and trip route category (one way or round trip). The data included three time periods which correspond to pre-Covid-19 time, the highest test rate, and decrease of morbidity with the beginning of vaccination, ranging from December 2019 to January 2020, December 2020 to January 2021, and February 2021 to March 2021. According to a report from the Los Angeles County Department of Public Health, the average daily new cases reached the peak from December 14th, 2020, to January 29th, 2021. Regarding shared bike as the first/last-mile transportation, 500m radius buffers around bike stations were applied and the Thiessen-polygon method was introduced into the study to avoid the buffer zone overlapping issue. The Thiessen-polygon defined edges to make all locations in partitioned areas closest to area centres that ensured buffer zone not overlapped and urban data in areas not double-counted (Rhynsburger, 1973). Figure 1 shows Thiessen-polygon buffers created in ArcGIS 10.7 with clear boundaries.



*Figure 1. Bike stations with 500m Thiessen-polygon buffers in study areas*

The original dataset contained a total of 103746 bike shared trip records, constituted by 50307 trips, 25170 trips, and 28269 trips, respectively. Due to the lack of actual trip route information, the calculation of trip distance was based on longitudes and latitudes of start points and endpoints. There was filtration of trips' unrealistic durations and distances that lasted less than 1 minute or longer than 1440 minutes and shorter than 100m or more prolonged than 50,000m. The round-trip removal was removed from

the raw dataset considering the accuracy of trip distance calculation and ensuring actual bike usage. The actual working data were 42255 trips, 17892 trips, and 19932 trips, respectively, after the data cleaning process.

Acquisition of land use, built environment, and urban infrastructure information was through OpenStreetMap (OSM) API, including primary roads, secondary roads, university area, commercial area, park area, industrial area, and retail area. Bus stops, metro stations, healthcare facilities, restaurants, and schools emerged from the points of interest (POIs) dataset from OSM API. Healthcare facilities consisted of clinics, doctors, and hospital POIs information. Data on sharing bike dock numbers per each bike station were from Bicycle Transit Systems open portal. The bikeway information came from the City of Los Angeles Hub API as open data. Furthermore, calculating all other land uses and built environment density indicators was based on previously mentioned numbers within ArcGIS buffer areas.

### 2.3. MODEL BUILD

The study first conducted a backward stepwise regression for each period to examine the relationship between trip frequencies with all variables. The method started with all variables included in the model, eliminated the least significant variable, and refitted in the statistical model until the most substantial explanatory variables got selected for all three models. The collinearity diagnostics were conducted simultaneously to ensure no intercorrelation was present on the multi-collinearity potentials between independent variables.

Second, the study employed Geographically Weighted Regression (GWR) models to investigate the spatial autocorrelation between variables which provided a local regression showing the spatial relationship (Fotheringham et al., 2002). Compared with global regression, the local statistic is not based on a single value but considered and examined the coefficients of variables in the vicinity of locations within the study area. Given by Fotheringham et al., the GWR model is shown as:

$$y_i = a_0(u_i, v_i) + \sum_k^m a_k(u_i, v_i) x_{ik} + \varepsilon_i \quad (1)$$

where  $y_i$  is the trip frequency in  $i$  Thiessen-polygon buffer;  $a_0(u_i, v_i)$  is the intercept parameter at  $i$  location;  $x_{ik}$  is an independent ( $k$ th) variable at  $i$  location;  $a_k(u_i, v_i)$  is the local coefficient of  $k$ th variable;  $\varepsilon_i$  is the random error at  $i$  location;  $m$  is the number of independent variables. Based on the equation (1), the estimated local coefficient of each location can be different within the study area which can preserve spatial relationship heterogeneity in this study. The large size of trip data can calibrate some bias and make standard errors lower during GWR estimation process (Fotheringham et al., 2002). The observed indicator closer than the location has a stronger influence than farther data. That can be presented as,

$$\hat{a}(u_i, v_i) = [X^T W(u_i, v_i) X]^{-1} X^T W(u_i, v_i) y$$

where,  $\hat{a}$  is the estimation of  $a$ ;  $W(u_i, v_i)$  is a  $n \times n$  matrix of geographical weighting of data at  $i$  location;  $y$  is the dependent variable.

### 3. Results

This study was to investigate how the pandemic impacts the association between bike-sharing usage and land use. Table 1 presented the summarized usage pattern of bike-sharing from Dec 2019 to Jan 2020 (Pre-Covid), Dec 2020 to Jan 2021 (Peak), and Feb 2021 to March 2021 (Early vaccination). It was noted that high bike sharing usage was recorded before the pandemic. The lowest trip volume occurred during the second “Stay at home” order period with significantly declined trip numbers. Accompanied with implementing vaccination and slight order release, bike-sharing usage began to increase. However, the trip volume was reduced during the pandemic, the trip duration and trip distance steadily increased. The trend might project that people were more likely to choose bike-sharing instead of shared riding (i.e., Uber and Lyft) concerning the unwillingness to stay with strangers in a private small automobile.

Table 1. Summary statistics for trip records in Los Angeles, United States.

	Pre-Covid (total 42255 trips)	Peak (total 17892 trips)	Early vaccination (total 19932 trips)
Trip duration (minute)			
Mean	17	28	32
Medium	10	16	16
Std. Dev.	50.7	71.9	83
Trip distance (meter)			
Mean	1389.6	1834	2013.3
Medium	1082.1	1309.2	1508.6
Std. Dev.	1040.8	1732.2	1836.7

#### 3.1. STEPWISE LINEAR REGRESSION MODEL AND VARIABLES

Three backward stepwise linear regression models by SPSS for different study periods were employed to examine the impact of Covid-19 on the association between sharing usage and land use, which could avoid multicollinearity problems and select the most influential explanatory indicators in the models. Table 2 presents the output of three models corresponding to different periods. The outcomes for peak time and early vaccination period had more similarity, mainly possessing the same predictors in final models, excluding primary road density. Additionally, bike path density, industrial area density, retail area density, school density, university density, and healthcare POIs density were eliminated from all three models. The backward stepwise linear regression results demonstrated that commercial area density and restaurant POIs density significantly correlated with bike-sharing usage. The centre of the commercial area usually limits automobile accessibility and parking spaces that can attract people to choose shared bikes. It was perceived that restaurant as essential service keeping high significance level which might be related to food pick-up services and people more likely to choose closer restaurants. Nevertheless, the park area density variable was excluded during the pandemic to reduce the rate of people going out only for leisure and recreation reasons to avoid exposure to Covid-19 in large crowds.

Table 2. Backward stepwise linear regression output

Explanatory variables	Pre-Covid		Peak		Early vaccination	
	Coef.		Coef.		Coef.	
Bike dock density	.320	***	.239	***	.223	**
Bike path density						
Primary road density	.148	**			-.204	**
Secondary road density			-.199	***	-.298	***
Intersection density			.147	***	.180	***
Bus stop density			.173	**	.184	*
Metro rail station density	-.097	*				
Commercial area density	.205	***	.189	**	.231	*
Industrial area density						
Retail area density						
Park area density						
University density	.141	***				
School POIs density						
Restaurant POIs density	.162	***	.262	***	.214	***
Healthcare POIs density						

Coef. = coefficient; "-" = retain variable in the model; \* $p \leq 0.1$ ; \*\* $p \leq 0.05$ ; \*\*\* $p \leq 0.01$ .

### 3.2. LAND USE AND BIKE-SHARING USAGE IN GWR MODELS

One of the primary aims of this study was to understand the pandemic impact on the association between urban land use and bike-sharing usage locally. GWR has been extensively used and proved a better fit for exploring predictors' spatial heterogeneity than global regression models (Fotheringham et al., 2002; Lu et al., 2014; Zhang et al., 2016). To further examine the variation of the local coefficients before the pandemic and during the pandemic, three GWR models with land use factors and other explanatory variables were applied to different study periods, respectively.

Figure 2 and figure 3 present the local coefficients of restaurant POIs density and commercial area density in three GWR models, considering the significant association between these two explanatory variables and bike sharing usage in the results of backward stepwise linear regression models. The local coefficients of restaurant POIs density appeared to positively associate bike-sharing use both before and through the pandemic. Before the pandemic, the spatial distribution of the strongest association was observed near the downtown area. The distribution became polycentric in the peaked period, centred on west Santa Monica and the whole North Hollywood area. During the early vaccination period, the centre of positive association moved from the west side of Santa Monica to the east. Regarding the local coefficients of commercial area density, the association was relatively weak compared with restaurant POIs density. Unexpectedly, the spatial distribution maps before and during the morbidity peak shared similarities that Santa Monica and North Hollywood were clustered to have strong associations and showed a decreased trend to the downtown area. It offers a

weak negative association near Santa Monica, and the trend generally fell from north to south after vaccination implementation.

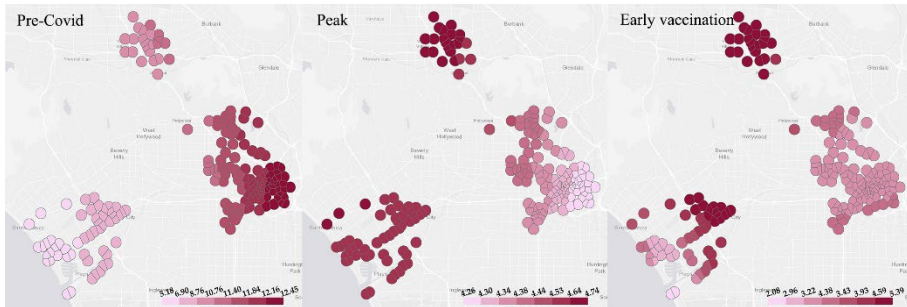


Figure 2. Restaurant POIs density spatial distribution of coefficients in three periods

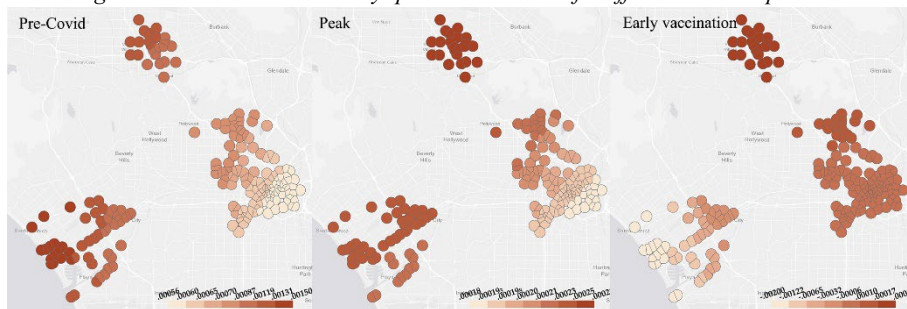


Figure 3. Commercial area density spatial distribution of coefficients in three periods

In general, the restaurant POIs density had a continuous significant positive contribution to bike-sharing usage throughout three study periods. It implies that restaurants attract more customers who prefer to ride shared bikes. Neighbourhoods near restaurants might be challenging to find parking spaces, and parking fees are usually higher than in other areas. Although the coefficients during the Covid-19 pandemic are relatively lower than before, food services as an essential need are still a strong reason for people coming outside and keeping demand for shared bikes. The discrepancy of spatial distribution might cost from the business closure and the decrease of commute population in the downtown area. The insignificant and negative coefficients of commercial density might have two reasons. First, most people were working from home, leading to fewer work-related trips, and unwilling to have nonessential outdoor activities to avoid contact with crowds. Second, due to the government orders and policies, commercials were kept closed through the most severe period to reduce gathering possibility. According to the results, North Hollywood and Santa Monica have a bigger potential to increase bike-sharing usage due to the considerable number of restaurants and commercial areas so expanding bike numbers in the areas might be considered as a rebalancing suggestion. Therefore, the study of the pandemic impact on the association of land use and bike-sharing in GWR models is worth discussing, which projects the spatial variation of different land-use factors. Thus, it gives clear visibility of how bike-sharing demand and spatial pattern change due to Covid-19.

### 3.3. OTHER VARIABLES

Bike dock density was the only significant factor other than land-use-related factors remaining in all three globe regression models. As figure 4 shows, most local shared bike dock density coefficients indicated a strong positive association with bike-sharing usage before or during the pandemic. In the pre-Covid and the peaked period, strong associations were observed around the downtown area, where a larger commuting population with smaller street block scales had. After vaccination coming out, the local coefficients tended to decrease from Santa Monica to northeast orientation, and the downtown area had consistently high local coefficient values. It is noteworthy that the North Hollywood region showed relatively weaker associations in all GWR models. It might be due to less shared bike docks installation and larger block scales. The number of shared bike docks and essential infrastructures encourage bike-sharing usage in all three areas. Hence, North Hollywood deserves significant bike dock infrastructure improvement after the pandemic to increase shared bike usage.

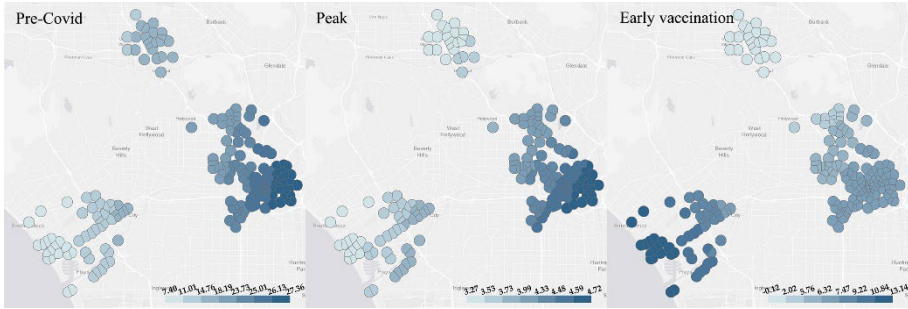


Figure 4. Bike dock density spatial distribution of coefficients in three periods

The results of backward regression models showed that secondary road and intersection density were both highly significant after the appearance of Covid-19. Figures 5 showed the spatial distribution of secondary road density local coefficients after the appearance of the Covid-19 pandemic. The secondary road density had a weak and negative association with bike-sharing usage in all study areas, reflecting that the secondary roads were not fit for bike-sharing users in Los Angeles.

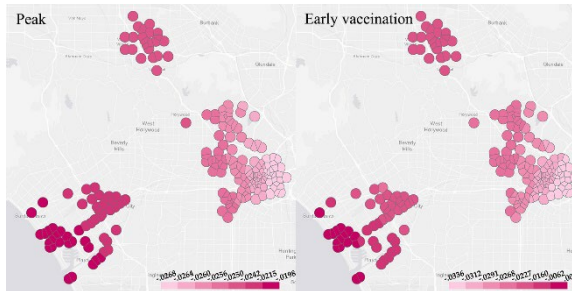


Figure 5. Secondary road density spatial distribution of coefficients in peak and early vaccination period



Figure 6 displays the local coefficients of intersection density. The presence of intersections negatively affected bike sharing usage. Roads with more intersections were not preferable for bike-sharing riders and were challenging to navigate while riding a bike.

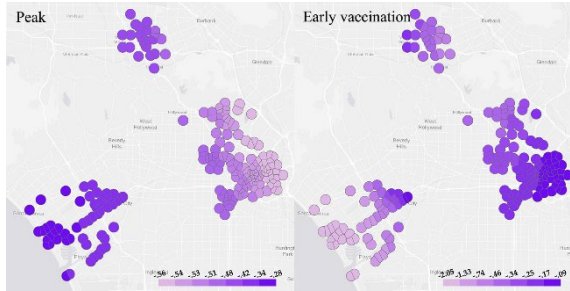


Figure 6. Intersection density spatial distribution of coefficients in peak and early vaccination period

#### 4. Discussion and Conclusions

This study aimed to investigate how the Covid-19 pandemic impacted the association between land use and bike sharing usage in Los Angeles. It contributes to the literature by analysing the influence of bike infrastructure, public transportation, and urban land use on trip volumes covering 3 different periods. Intuitively, the demand for bike-sharing was expected to drop after the Covid-19 pandemic appearance in general, but the trip volume was gradually increasing after releasing the lockdown order. The purpose of cycling might shift from work-related commute to more prolonged recreation or essential needs. Shared bikes might become a potential substitute for buses in medium-distance trips. The results of GWR revealed the spatial heterogeneity of the impact of land-use factors and other urban infrastructure variables on ridership in Los Angeles. GWR modelling as a tool can provide direct local urban analysis for spatial relationships of various urban database. This meaningful methodology presents a clear way to visualize spatial correlations of urban data and can be applied to other locations. GWR could overcome the limitations of global regression models which could be Hence, the results of the study could inspire policymakers and planners on how to implement bike-sharing system effectively based on spatial heterogeneity and make specific strategies for different locations. The pandemic negatively impacts bike-sharing demand and influences the spatial association between bike facilities, public transport accessibility, essential needs destination, open space, and bike-sharing system usage. According to the results, increasing bike dock quantity would strongly encourage commuters to ride shared bikes, releasing local traffic and contributing to sustainable city strategies. Riders are more like to avoid intersections and roads with large automobile volumes. While rebalancing existing bikes or installing new bike stations, policymakers and operators might consider ridership safety, essential needs for food services, groceries, and recreation destinations.

However, the pandemic still rages through countries; the current paper's main limitations concerning the timing of data collection are based on a limited period, in which individuals constantly change their travel behaviours and attitudes on sharing

transportation. Simulating and proposing implementation suggestions quantitatively based on this study results would be included in future study to provide further comparison. Besides, due to the specific bike station locations in Los Angeles, the usage pattern is centralized in specific areas. More broadly, additional data and surveys in other cities need further investigation.

## References

- Abdullah, M., Dias, C., Muley, D., & Shahin, M. (2020). Exploring the impacts of COVID-19 on travel behavior and mode preferences. *Transportation Research Interdisciplinary Perspectives*, 8, 100255.
- Aloi, A., Alonso, B., Benavente, J., Cordera, R., Echániz, E., González, F., Ladisa, C., et al. (2020). Effects of the COVID-19 Lockdown on Urban Mobility: Empirical Evidence from the City of Santander (Spain). *Sustainability*, 12(9), 3870. <http://doi.org/10.3390/su12093870>
- Fotheringham, A., Brunson, C., & Charlton, M. (2002). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Wiley.
- Jenelius, E., & Cebeauer, M. (2020). Impacts of COVID-19 on public transport ridership in Sweden: Analysis of ticket validations, sales and passenger counts. *Transportation Research Interdisciplinary Perspectives*, 8, 100242. <https://doi.org/10.1016/j.trip.2020>.
- Kubal'ák, S., Kalašová, A., & Hájnik, A. (2021). The Bike-Sharing System in Slovakia and the Impact of COVID-19 on This Shared Mobility Service in a Selected City. *Sustainability*, 13(12), 6544. <http://doi.org/10.3390/su13126544>
- Lahoorpoor, B., Farooqi, H., Sadeghi-Niaraki, A., & Choi, S. (2019). Spatial Cluster-Based Model for Static Rebalancing Bike Sharing Problem. *Sustainability*, 11(11), 3205. <http://doi.org/10.3390/su11113205>
- Martens, K. (2006). Promoting bike-and-ride: The Dutch experience. *Transportation Research Part A*, 41(2007), 326-338. <http://doi.org/10.1016/j.tra.2006.09.010>
- Mujahed, L. (2021). Urban Resilience: Relation between COVID-19 and Urban Environment in Amman City. *International Journal of Urban and Civil Engineering*, 15(3), 172-181.
- Nikiforiadis, A., Ayfantopoulou, G., & Stamelou, A. (2020). Assessing the Impact of COVID-19 on Bike-Sharing Usage: The Case of Thessaloniki, Greece. *Sustainability*, 12(19), 8215. <http://doi.org/10.3390/su12198215>
- Pase, F., Chiariotti, F., Zanella, A., & Zorzi, M. (2020). Bike Sharing and Urban Mobility in a Post-Pandemic World. *IEEE Access*, 8, 187291 - 187306. <http://doi.org/10.1109/ACCESS.2020.3030841>
- Rhynsburger, D. (1973). Analytic delineation of Thiessen polygons. *Geographical Analysis*, 5(2), 133-144.
- Wu, X., Lu, Y., Lin, Y., & Yang, Y. (2019). Measuring the Destination Accessibility of Cycling Transfer Trips in Metro Station Areas: A Big Data Approach. *Int. J. Environ. Res. Public Health*, 16(15), 2641. <http://doi.org/10.3390/ijerph16152641>
- Wang, H., & Noland, R. (2021, January 26). *Changes in the pattern of bikeshare usage due to the COVID-19 pandemic*. Findings. Retrieved May 16, 2021, from <https://findingspress.org/article/18728>
- Zafri, N.M., Khan, A., Jamal, S., & Alam, B.M. (2021). Impacts of the COVID-19 Pandemic on Active Travel Mode Choice in Bangladesh: A Study from the Perspective of Sustainability and New Normal Situation. *Sustainability*, 13(12), 6975. <http://doi.org/10.3390/su13126975>
- Zhang, Y., & Mi, Z. (2018). Environmental benefits of bike sharing: a bigdata-based analysis. *Applied Energy*, 220, 296-301. <https://doi.org/10.1016/j.apenergy.2018.03.101>

# MEASURING RESILIENT COMMUNITIES

*An analytical and predictive tool*

SILVIO CARTA<sup>1</sup>, TOMMASO TURCHI<sup>2</sup> and LUIGI PINTACUDA<sup>3</sup>  
<sup>1,2,3</sup> *University of Hertfordshire, UK.*

<sup>1</sup>*s.carta@herts.ac.uk, 0000-0002-7586-3121*

<sup>2</sup>*t.turchi@herts.ac.uk, 0000-0001-6826-9688*

<sup>3</sup>*l.pintacuda@herts.ac.uk, 0000-0002-9422-1810*

**Abstract.** This work presents the initial results of an analytical tool designed to quantitatively assess the level of resilience of urban areas. We use Deep Neural Networks to extract features of resilience from a trained model that classifies urban areas using a pre-assigned value range of resilience. The model returns the resilience value for any urban area, indicating the distance between the centre of the selected area and relevant typologies, including green areas, buildings, natural elements and infrastructures. Our tool also indicates the urban morphological characteristics that have a larger impact on the resilience score. In this way we can learn why a neighbourhood is successful (or not) and how to improve its level of resilience. The model employs Convolutional Neural Networks (CNNs) with Keras on Tensorflow for the computation. The outputs are loaded onto a Node.JS environment and bootstrapped with React.js to generate the online demo.

**Keywords.** Sustainable Cities and Communities; Resilient Communities; CNN; Urban Morphology; SDG 11; SDG 13.

## 1. Introduction

The configuration of the built environment has a significant impact on the ways in which people inhabit the urban space. Within the scope of this study, we consider resilience in terms of the ways in which urban communities respond to any event on the basis of the presence, location and configuration of the resources in their neighbourhoods and cities. In case of adverse circumstances (that can happen at a sudden as an earthquake or flooding, or more slowly like climate change), communities adapt, withstand and thrive in and around the physical elements of their urban space. This work is underpinned by the assumption that there is a significant correlation between urban morphology and the resilience of communities.

In previous work (Carta et al., 2021), we developed a quantitative method to evaluate the resilience of net-zero communities, based on position, density and proximity of physical resources. In this paper we present further findings where we apply our method to generate an online tool to automatically evaluate the level of

resilience of any neighbour and urban area based on their urban configuration.

The end-goal of our research is to develop a reliable method to calculate resilience values of urban communities based on urban morphology. In particular, we aim to use satellite images using models for object detection, so that our model can automatically assess and consistently provide a resilience value for any urban area in the world. The long-term plan is to train our model to classify resilience values directly from the images, by training it with a robust list of values of resilience for each urban area. In this study, we present the first step towards this plan where we use object-detection to identify relevant urban typologies in satellite imagery and evaluate resilience values directly on the web app.

## 2. Methods

We generated two datasets of satellite images. The first one is based on the DeepGlobe Land Cover Classification Dataset (Demir et al., 2018) and has been used for the training of the model on Tensorflow. We used a sample of 100 images (out of the 803 images in the original set) selecting the most representative with regards to the typologies observed.

The second dataset has been created for this project using QGIS to extract the OpenStreetMap features related to the physical environment (detectable in aerial images) as for our previous work (Carta et al., 2021). This second set has been used for the validation of the model.

### 2.1. WORKFLOW

For the resilience predictive model, we followed the workflow below:

Dataset: 1) collect satellite images from the DeepGlobe Land Cover Classification Dataset; 2) object labelling with VoTT (Visual Object Tagging Tool); 3) export the labelled images (in JSON format) (VoTT-JSON); 4) Import the labelled data (JSON) into roboflow; 5) Pre-process the images in roboflow (including resizing and augmentation); 6) in roboflow generate a new dataset (with pre-processed images) and create train/test split (70/20/10%); 7) export new dataset to YOLOv5/PyTorch format for training.

Training: 1) Import the pre-processed dataset from roboflow to Colab; 2) Clone Yolov5 on Colab; 3) Define model configuration and architecture; 4) Train custom Yolov5 Detector (approx. 4hr – using CUDA tensor types running on GPU); 5) Evaluate Custom YOLOv5 Detector Performance; 6) Run inference with training weights; 7) Export weights (with the best weight model) for tensorflow.js.

Web app visualiser: 1) Import tensorflow.js weights from Colab; 2) Apply the model to the on-screen satellite image selected by the user; 3) Compute the distance between each identified cluster to the GPS location on the centre of the screen; 4) show the final resilience score.

### 2.2. DATA PREPARATION AND LABELLING

The satellite images from the first dataset (all of which were at the same altitude and resolution) have been labelled one by one using Microsoft VoTT: Visual Object

Tagging Tool (Microsoft, 2019). Our previous work on the city of Copenhagen (Carta et al., 2021) showed that resilience values calculated on proximity and density of physical elements are mostly affected by green areas, natural elements and entertainment venues. Based on these previous findings, we focused on the 4 visually recognisable typologies below to identify relevant classes for our training: 1) Green areas, 2) Buildings (built areas vs unbuilt), 3) Large infrastructures (train stations, stadiums etc.) and 4) Natural elements (lakes, rivers, coasts etc.).

The labelling on VoTT allowed having a clearly tagged dataset where all images are mapped to the 4 classes. All images and labels have been pre-processed, including resizing all set to 416×416 pixels to ensure consistency in the training. We augmented the initial 100 images to 559 images including horizontal flip, rotation (-15° to +15°), saturation (-50% to +50%) and exposure (-25% to +25%). The dataset has been split into 70% for the training set (489 images), 20% for the validation set (46 images) and 10% for testing (24). The set ready for training has been exported in Yolov5 format, as the YOLO architecture allows for tensors and the use of GPU perfect for Tensorflow.

### 2.3. TRAINING DATA ON TENSORFLOW

The main idea underpinning this experiment is to use the YOLOv5 Detector (Redmon et al. 2016), which is an object detection model based on DenseNet/CNN backbone and train the weights for the model using inference (Gavali and Banu, 2019). Our model uses Keras on Tensorflow.

The You Only Look Once (YOLO) model was developed in 2015 and has attracted significant attention in researchers in computer vision since then. YOLO is defined as “an object detection algorithm that divides images into a grid system. Each cell in the grid is responsible for detecting objects within itself” (Ultralytics, 2021). One of the advantages of YOLO is that its system “computes all the features of the image and makes predictions for all objects at the same time. That is the idea of "You Only Look Once" (Thuan, 2021).

Released in 2020 by Glenn Jocher on GitHub (Jocher, 2020), YOLOV5 has been introduced as 'a family of object detection architectures and models [...] [that] represents [...] research into future vision AI methods' (Jocher, 2020). YOLOV5 object detection model is based on a DenseNet architecture (cf. EfficientDet architecture, which employs 'EfficientNet as the backbone network, BiFPN (bi-directional feature pyramid network), as the feature network, and shared class/box prediction network' (Tan et al., 2020:5).

For the training of our model, the following method has been followed: 1) Install dependencies, including TORCH.CUDA. This package adds support for CUDA tensor types, that implement the same function as CPU tensors, but they utilize GPUs for computation (PyTorch, 2021); 2) Import labelled dataset from roboflow server; 3) Define model configuration and architecture; 4) Train custom Yolov5 Detector. The model has been trained with 50 epochs and the network had 283 layers, 7,263,185 parameters, 7,263,185 gradients, 16.8 GFLOPs (Giga flops, or floating-point operations per second); 5) Evaluate Custom YOLOv5 Detector Performance, as shown in Figure 1.

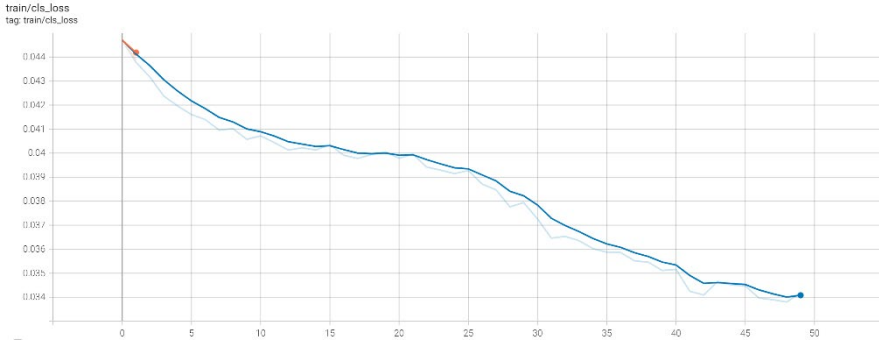


Figure 1. Visualising the performance of the object detector. The class-loss curve goes down to a value of 0.033 after around 40 epochs in this train.

6) Visualise training data with labels for control, as shown in Figure 2.

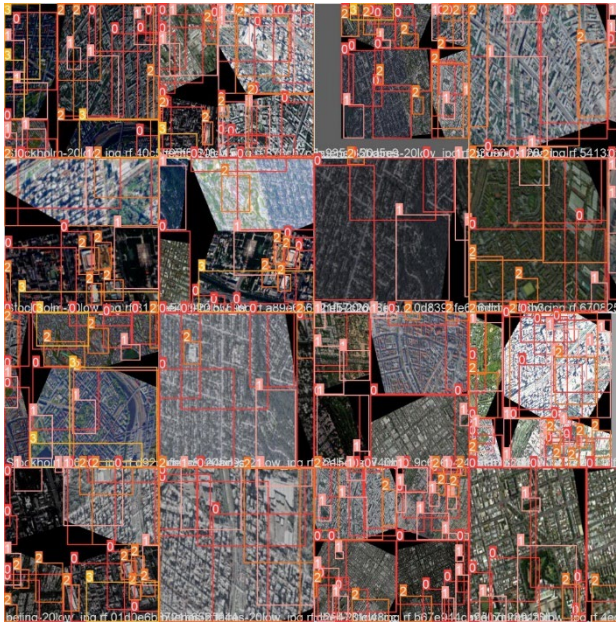


Figure 2. This figure shows the ground truth training data where the classifier has been initially tested with the training set. The classifier correctly assigned various classes (0 for green areas, 1 for buildings, 2 for large infrastructures and 3 for natural elements).

7) Run inference with training weights. The last part of this method is to use the best

weight in the training to run inference on a new set using the method `detect.py`. In this case, we use the testing set to initially evaluate the model within Colab. The best weight is used to evaluate the resilience value of any new map inputted to the model. Figure 3 illustrates how the model is able to detect the presence of buildings in a test image.



*Figure 3. detection of building groups with the testing set.*

#### 2.4. VISUALISE RESULTS ON THE WEB APP

For this experiment we developed a web app to calculate the  $R$  values directly on browsers, where we evaluate the distance in pixels between the centre of the map (as selected by the user) and the centre of the rectangular selection identified by the Yolov5 model.

The Web app is developed with the front-end Javascript framework React.js and uses Tensorflow.js to load and run the Yolov5 model online. It presents users with a Google Maps-like interface, where one can move around a satellite map of the World, pan and zoom in and out, and look for specific places by label. A button on the lower right corner triggers the execution of the classification over the white-overlaid area on the map, as shown in Figure 4. A static satellite image of the GPS location on the centre of the screen is fetched from Google Maps, encoded, and run through the Yolov5 model.

The results are pixel coordinates of the identified markers for each class (i.e., green areas, buildings, infrastructures, natural elements). We convert these coordinates back to GPS locations in order to draw the right squares onto the onscreen map (each with a different shade depending on the class it belongs to) and calculate the line-of-sight distance between each point and the centre of the screen. Finally, we calculate the  $R$  value using these distances and the gamma values for each class, displaying the result on screen. On this initial prototype, we are not using the uncertainty given by the model for each identified marker, but it could play a role in future instances of the model.



Figure 4. Web app interface where user can visualise  $R$  values based on area search.

### 3. Model Validation

In order to test the accuracy of our model, we used a sample of 15 cities, for which we compared the resilience values obtained by using two different methods. Firstly, we used data from existing literature and rankings, using data from the 2013 Grosvenor report (Barkham et al., 2013) and the Quality of Life global ranking from Numbeo (2021). Secondly, we evaluate the resilience values for the same cities using the method we developed in previous work (Carta et al., 2021), as detailed in Section 3.1. We then compared the three sets of values to estimate the precision of our model.

#### 3.1. RESILIENCE BY PHYSICAL ELEMENTS

In order to test our model, we calculated the resilience value of each neighbourhood using the method developed in Carta et al. 2021 which can be summarised in (1):

$$R = \sum_{i=1}^n d(\min)_n \gamma_n \quad (1)$$

where  $R$  is the overall resilience value for eth observed neighbourhood,  $d(\min)$  is the minimum distance between the centre of the neighbourhood and the urban typology considered (e.g. local school, park etc.) and  $\gamma$  is a coefficient that considers the quality of the distance of the typology based on Kronberg et al. (2019), Government Office for Science (2019) and Knupfer et al. (2018) for the private/public transportation ratio and National Travel Survey (2014) for education settings. The  $\gamma$  values are calculated following Table 1:

Good	Fair	Bad
< 15 min	15 min / 30 min	> 30 min
< 1,260 Km	1,260 to 2,520 Km	> 2,520 Km

Table 1. We considered walking distance in minutes and Kilometres: average of 1.4 metres per second or 5 km per hour.



As this method requires the georeferenced position of urban typologies (like schools, train stations, parks etc.), we created a specific map for each city using QGIS to import the OpenStreetMap features related to the physical environment at the scale of the city and the pin as from the Google Earth image. To calculate the values for each of the 4 typologies included in this study (green areas, buildings, large infrastructures and natural elements), we run a simple Grasshopper definition to compute the equation (1) using the maps created with OpenStreetMap and QGIS. The Grasshopper definition provided the values shown in the third column of Table 2.

### 3.2. VALUES COMPARISON

The resilience values obtained with these two methods have been compared with those calculated with the Yolov5 model as shown in Table 2 below.

	R (QoL)	R (GSVN)	R (GH)	R (Yolov5)
Munich	27	24	14.14	9.48
Seattle	26	11	7.56	3.08
Beijing	234	39	7.39	1.43
Tokyo	87	26	6.97	1.47
Melbourne	44	13	6.7	1.24
Rio	232	45	6.7	1.22
Pittsburgh	47	5	6.55	-
Cairo	227	48	5.89	5.48
Moscow	202	37	5.78	-
Delhi	236	42	4.92	0.79
London	149	18	4.82	-
Toronto	101	1	4.69	1.31
Singapore	113	32	4.53	-
Bno Aires	213	36	4.06	1.12
Stockholm	92	6	3.92	1.25

Table 2. Comparison of different values of resilience. QoL (1st column) from Numbeo (2021) indicating Quality of Life values, GSVN (2nd column) from Gosvernors (2014), R values obtained with our Grasshopper script (3rd column) and values with our Yolov5 model (4th column).

### 3.3. ANALYSIS OF RESULTS

The four sets of R values summarised in Table 2 differ significantly in relative magnitude between sets. However, they appear consistent throughout in relative values per city, as shown in Figure 5. For example, we see that Buenos Aires has lower R values than Delhi and higher than Melbourne in all 4 sets. This is consistent in all cities observed.

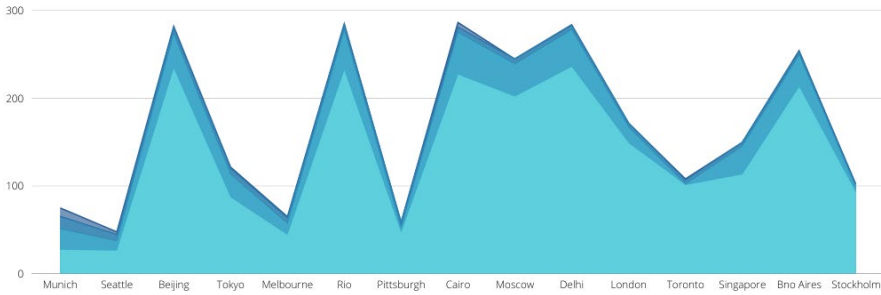


Figure 5. Plot of the  $R$  values with the 4 methods from Figure 2.

#### 4. Limitations and Next Steps

In the development of this model, we identified the following limitations. Firstly, we used a limited dataset (around 500+ images) for model training. An architecture like that of Yolov5 would require a larger number of images per class (in the order of thousands). However, Yolov5 is able to provide accurate results even with low numbers of images (Ultralytics, 2021). Moreover, the dataset used for the web app is based on DigitalGlobe Basemap +Vivid (Maxar, 2021) and differs from those of Google Maps and this might lower the accuracy of our results.

Secondly, we employed a limited number of categories to define the urban morphology of the areas observed. We limited the classes to 4 (green areas, buildings etc.). A more comprehensive experiment should include more categories, for example distinction between buildings (schools, residential, museums etc.). This approach will move from the analysis of simple satellite imagery to augmented images through the integration with OpenStreetMap data.

Another point to consider for refinement of this work is the way in which different values in (1) for each typology are summated and averaged. In this experiment, we did not attribute different weights to points. By acknowledging the individual contribution of each element, using for example a coefficient that accounts for distance from the centre of the area, or the size of the typology, we could yield more accurate results. Also, the model returns a value of uncertainty of each identified cluster, which might also be included in the  $R$  measure.

We also noticed that the model predicts differently depending on the zoom of the visualised area on screen (larger zoom results in more features being detected). This is partially related to the scale of the images in our training set and to the resolution of images loaded on the web app when calculating  $R$ . In future iterations of the app, we will consider constraining the zoom of the area, once the user has made the selection.

Finally, a limitation we observed in using Yolov5 was the difficulty in creating custom datasets without using roboflow to pre-process the dataset and labelling.

The next step for this project is to build a larger dataset with labels in order to refine the model and its accuracy.

## 5. Conclusions

The model we developed compute the resilience value  $R$  of any urban area in the world. For each selected area, the app returns an overall  $R$ , as well as a breakdown of the average of each cluster of urban typologies sorted by category, as shown in Figure 6. With this study, we introduce the early stages of a novel tool to analyse urban areas and their level of community resilience, as well as elements to suggest how to intervene to improve it.



Figure 6. Screen grab of the app showing the overall value  $R$  and 4 clusters of buildings detected by the model. The distance from the centre of the area to each cluster is used to calculate the individual contribution of that cluster to  $R$  (0.31, 0.28, 0.26).

This tool can be helpful to analyse areas of the world that are not currently included in world's resilience rankings, and to help designers to test design hypothesis for new schemes. For example, planners can analyse georeferenced master plan images to compare current levels of resilience to those resulting from the proposed schemes. We this initial results, we aim at sharing with colleagues our findings and gather feedback on how to improve the tool in the next stages of development.

## Acknowledgements

Our YoloV5 model, the Colab notebook, the Grasshopper definition and the dataset used for training can be found at: <https://anonymous.4open.science/r/caadria2022-4687>. The web app can be found here: <https://caadria2022.netlify.app/>

## References

- Barkham, R.J., Brown, K., Parpa, C., Breen, C., Carver, S. and Hooton, C., (2013). *Resilient cities: A Grosvenor research report. Grosvenor Global Outlook*. Available at: <https://www.alnap.org/system/files/content/resource/files/main/resilient-cities-a-grosvenor-research-report-2014.pdf>
- Carta S, Pintacuda L, Owen I. W. and Turchi T (2021) Resilient Communities: A Novel Workflow. *Frontiers of Built Environment*. 7:767779. <https://doi.org/10.3389/fbuil.2021.767779>

- Demir, I., Koperski, K., Lindenbaum, D., Pang, G., Huang, J., Basu, S., Hughes, F., Tuia, D. and Raskar, R., 2018. Deepglobe (2018). A challenge to parse the earth through satellite images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 172-181). From: <https://www.kaggle.com/balraj98/deepglobe-land-cover-classification-dataset>
- Gavali, P. and Banu, J.S. (2019). Deep convolutional neural network for image classification on CUDA platform. In *Deep Learning and Parallel Computing Environment for Bioengineering Systems* (pp. 99-122). Academic Press. <https://doi.org/10.1016/B978-0-12-816718-2.00013-0>
- Government Office for Science (2019). *A time of unprecedented change in the transport system*. GfOS. Retrieved from [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/780868/future\\_of\\_mobility\\_final.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/780868/future_of_mobility_final.pdf)
- Jocher, G. (2020). *YOLOV5*. Ultralytics/yolov5. GitHub. Available at: <https://github.com/ultralytics/yolov5>
- Knupfer, Stefan M. (2018). *Elements of success: Urban transportation systems of 24 global cities*. McKinsey & Co.
- Kronberg, Nico et al. (2019) *Transport Statistics Great Britain 2019*. Department for Transport. Available at: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/870647/tsgb-2019.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/870647/tsgb-2019.pdf)
- Maxar (2021). *DigitalGlobe Basemap +Vivid*. Maxar. Retrieved from <https://www.maxar.com/products/imagery-basemaps>.
- Microsoft (2019) *VoTT*. Commercial Software Engineering (CSE) group. Retrieved from <https://github.com/microsoft/VoTT>.
- National Travel Survey (2014). *Travel to school*. Department for Transport.
- Numbeo (2021). Available at: <https://www.numbeo.com/quality-of-life/rankings.jsp>
- Redmon, J., Divvala, S., Girshick, R. and Farhadi, A., 2016. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).
- Roboflow (2020). Available at: <https://roboflow.com/>
- Tan, M., Pang, R. and Le, Q.V., (2020). Efficientdet: Scalable and efficient object detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 10781-10790).
- Thuan, D., (2021). *Evolution of yolo algorithm and yolov5: the state-of-the-art object detection algorithm*. <https://www.theseus.fi/handle/10024/452552>
- Ultralytics (2021). *YOLOv5 Documentation*. From: <https://docs.ultralytics.com/>

# A WEB-BASED INTERACTIVE TOOL FOR URBAN FABRIC GENERATION

*A Case Study of Chinese Rural Context*

QIYAN ZHANG<sup>1</sup>, BIAO LI<sup>2</sup>, YICHEN MO<sup>3</sup>, YULONG CHEN<sup>4</sup> and PENG TANG<sup>5</sup>

<sup>1,2,3,4,5</sup>*School of Architecture, Southeast University, Nanjing, China.*

<sup>1</sup>*zhangqiyang@seu.edu.cn, 0000-0002-1702-2041*

<sup>2</sup>*jz.generator@gmail.com, 0000-0002-8713-4954*

<sup>3</sup>*moyichen@seu.edu.cn, 0000-0002-2844-7678*

<sup>4</sup>*skadeshat@gmail.com, 0000-0001-9015-131X*

<sup>5</sup>*tangpeng@seu.edu.cn, 0000-0003-1658-6774*

**Abstract.** The design of rural fabric is significant for making sustainable communities and requires innovative design models and prospective work paths. This paper presents an interactive tool based on the web to generate block fabric that responds to the Chinese rural context, consisting of streets, plots, and buildings. The tool is built upon the Browser/Server (B/S) architecture, allowing users to access the generation system via the web simply and to have interactive control over the generation process in a user-friendly way. The underlying tensor field and rule-based system are adopted in the backend to model the fabric subject to multiple factors, with rules extracted from the rural design prototype. The system aims to integrate the procedural model with practical design constraints in the rural context, such as patterns, natural boundaries, elevations, planning structure, and existing streets. The proposed framework supports extensions to different urban or suburban areas, inspiring the promising paths of remote cooperation and generative design for sustainable cities and communities.

**Keywords.** Generative Design; Web-based Tool; Urban Fabric; Rural Context; Procedural Modelling; Tensor Field; SDG 11.

## 1. Introduction

Making sustainable rural communities with better diversity, green and liveable environments, local distinctiveness, and cultural identity is essential for developing sustainable cities and communities. With the enormous impact of urbanization comes the attendant rural decline and loss of organically grown fabric (Liu and Li, 2017; Kiruthiga and Thirumaran, 2019), while the homogenizing pattern in urban sprawl is severed from the local context. There exists an urgent need to propose a more scientific and prospective methodology to inspire integrated and sustainable urban planning and

the rural community's ongoing progress, which would simultaneously contribute to promoting positive regional development and urban-rural linkages.

Urban and architectural design is concerned with an integrated complex adaptive system (CAS) that requires responding to internal and external constraints from the local environment (Li and Han, 2011). With the development of computer-aided design (CAAD) recently, the generative design method for cities have shown significant advantages: content abstraction into a set of procedures, synthesis of complex correlated rules, effective generation of various schemes, and user operation through parametric control without understanding the internal mechanisms (Smelik et al., 2014). Since the pioneering research of Parish and Müller based on L-systems (Parish and Müller, 2001), procedural modelling for urban layout has gathered much attention. Several algorithms have been applied to model cities in previous studies, such as rule-based systems, multi-agent systems, stochastic geometry, tensor fields, data-driven.

The grid model was experimentally implemented by Gretel et al., proposing a real-time procedural system that generates individual buildings from an integer grid (Greuter et al., 2003). Jing Sun et al. converted frequent road patterns into templates to generate a traffic network within the constraints of input maps of the site boundary, terrain elevation, and population density (Sun et al., 2002). Many studies have begun to concentrate on organically grown settlements (Duarte et al., 2007; Emilien et al., 2012). The study carried out by Li et al. introduced a model of evolutionary multi-agent system (Li et al., 2015). Given the morphological prototype of Ji Village, the system achieved the optimal generation of natural village fabric. Several researchers have developed interactive tools to utilize the approach better and indicated some methods to provide intuitive control for users, including sketch-based techniques (de Villiers and Naicker, 2006) and merging operations (Lipp et al., 2011).

The above procedural approaches show great potential in enhancing the urban fabric design process. However, the correspondence between the procedural model and the specific design context for sustainable rural fabric has not been well addressed. As a result, this paper introduces an interactive tool that users can access via a web page and generate site-specific rural fabric that responds to the local environment and cultural characteristics. Two aspects are crucial to developing the system:

- The system uses the tensor field and rule-based system to establish a complex adaptive system with multiple constraints to adapt the generated fabric to the rural context. Chen et al. firstly employed the tensor field to offer interactive control over the streets generation process (Chen et al., 2008). In this research, the tensor field is used to simulate the impact on physical forms from potential forces of rural context. The system constructs a procedural model for organically grown fabric using the rules from rural prototypes. This paper mainly describes methods for generating tensor fields, streets, and plots, not involving methods for buildings.
- The web tool can provide simplified access and novel media to a large number of users, along with flexible and intuitive operations, showing the potential to provide co-design access to participants in different roles. This tool was developed based on the open-source web-based framework proposed by Yichen Mo (Mo, 2021). Specific work involves data structure design, display style design, detailed operation steps and interactive control, and data transfer between backend and frontend.

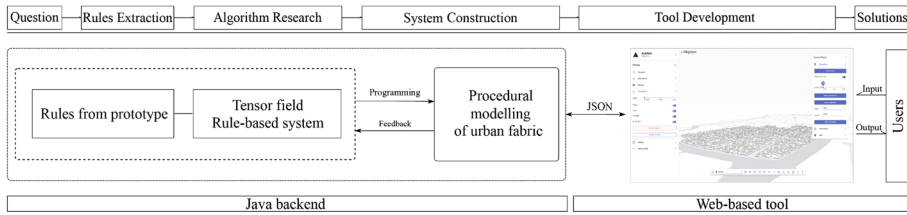


Figure 1. Research overview

## 2. Generation system

### 2.1. UNDERLYING TENSOR FIELD

There exist underlying forces from the local context that shaped and reshaped the physical forms of the urban fabric (Meng, 2015). This paper presents a generative system that uses tensor fields to model potential mechanisms and control the range and intensity of forces exerted by different constraints (such as boundaries, rivers, elevations, and axes). Thus, the generated fabric can respond flexibly to the environmental conditions and planning intentions, with the perspective to shape sustainable rural communities and realize a symbiotic relationship between the fabric and the natural environment.

Tensors are mathematical objects defining the multilinear mapping between sets of algebraic objects in a vector space. In order to form a smooth tensor field  $T$ , each point  $p = (x, y) \in R^2$  in the space is associated with a second-order tensor  $T(p)$  (Delmarcelle and Hesselink, 1994), and then derives the value of each tensor to make the whole into a continuously varying state. Streamlines tangent to the eigenvector field in space can be used to model roads in cities (Chen et al., 2008). This paper implements tensor field generation in the constraints of geometry and images via java and Gurobi Solver, and contributes to relating the method to specific design situations.

#### 2.1.1. Tensor Field Generation

Factors that affect urban texture can be identified in two types based on the data format input: geometries, such as site outlines, natural boundaries, road lines, axes, and central nodes; images, such as the grayscale map of elevation or population density. Though the specific data processing methods are different, both types follow a unified framework for constructing the underlying tensor field.

The tensor  $t$  defined in this system consists of four symmetric unit vectors, with basic properties including the position  $p(x, y)$ , the angle  $\theta$  of the vector in the first quadrant, and its projection on the coordinate axes:  $x_{vec}$ ,  $y_{vec}$ . Four major steps are adopted to evaluate the results. First, after inputting the constraints, an underlying two-dimensional mesh with Half-Edge Data Structure is generated, along with a hash map that stores each vertex and its related tensor. Secondly, the framework calculates the values of each tensor to derive a smooth field, following which the bilinear interpolation is applied to obtain the values at any other location. Finally, along the eigenvector field, traces the streamlines through moving the agent until it reaches outside the boundary or irrational region. The details for the second step of the above

two types are described below.

Geometric elements include polygons, polylines, and central nodes (for radial field). Let set  $V$  represent all the vertices of the underlying mesh. The first step involves identifying all the vertices of the mesh that overlap with these geometric elements (stored into set  $Q$ ) and setting the tensors of these vertices along the direction of the geometries (stored into set  $T$ ). Next, the system employs a quadratic programming model to solve the tensors of other vertices on the mesh. For vertices in set  $V$  with total  $n$ , the projection values of their related tensor could be described as  $P = \{var_1, var_2, \dots, var_n\}$ . Due to the unit length of the vector, the variables are subject to:  $-1 \leq var_i \leq 1$ ,  $var_i \in P$ . Meanwhile, the variables associated with the vertices in  $V \cap Q$  should be consistent with the values in set  $T$ . The optimization goal is to minimize  $\sum_{i=1}^n (var_{i1} - var_{i2})^2$  (the vertices corresponding to  $var_{i1}$  and  $var_{i2}$  are adjacent). Gurobi is used to solve the model and finally obtain the results. Figure 2 shows the generation process.

For Images, the system applies the gradient to calculate the value of each tensor, by which the developed tensor field reflects the features of elevation or population density (Figure 2). The expressions for the gradient are as follows:  $\nabla f(x, y) = [\partial f / \partial x, \partial f / \partial y]^T$ ,  $mag(\nabla f) = \sqrt{(\partial f / \partial x)^2 + (\partial f / \partial y)^2}$ ,  $\phi(x, y) = \arctan((\partial f / \partial y) / (\partial f / \partial x))$ . 2D kernel (Sobel or Canny operator) is used to scan the x and y axes to obtain the gradient.

The generative approach described above can be extended to a variety of design conditions and reflects the integrated effect of geometric and graphical constraints (Figure 3), shaping a space of forces that would control the forms of the urban fabric. The users can interactively add nodes or polylines of spatial axes for editing the tensor field.

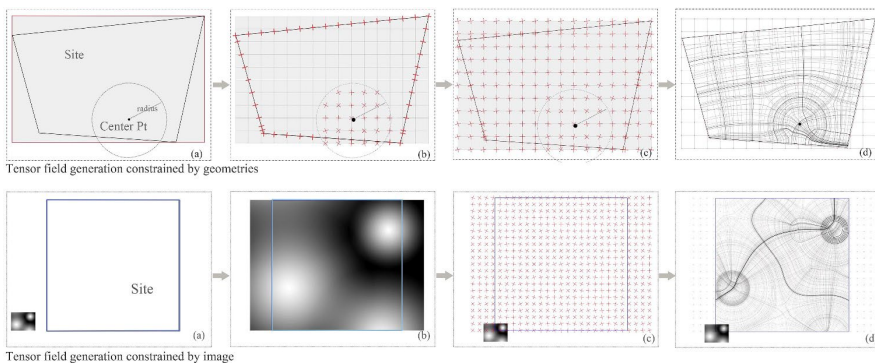


Figure 2. The process of tensor field generation

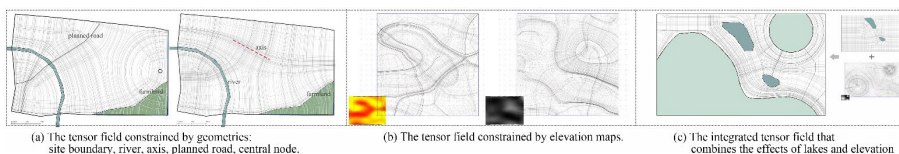


Figure 3. The tensor fields generated under diverse constraints



2.1.2. Application

With a view to further application of tensor fields in design, this study investigates three directions: user adjustment of the weight of constraints, control over the layout of roads, and organization of the road network.

The roads are simulated by the streamlines, forming a graph as a road network consisting of nodes and edges (representing the connection relationships). When adding different factors or changing their weights of influence, the generated tensor field and the consequent road deformation would change accordingly (Figure 5), thus allowing users to operate interactively from design requirements. Figure 4 shows a set of results in comparison: the tensor field when the city is influenced by the natural environment more than the axes, and the results when the city is planned more artificially. With the underlying tensor field, the approach can derive the graph of streets network adapted to the constraints (Figure 6).

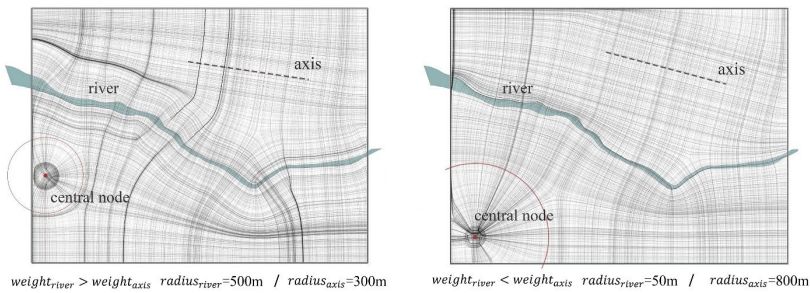


Figure 4. Comparison of the tensor fields constrained by factors of different weights and effect radii

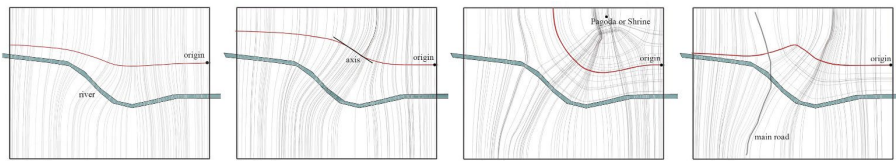


Figure 5. Changes of street deformation correspond to different constraints

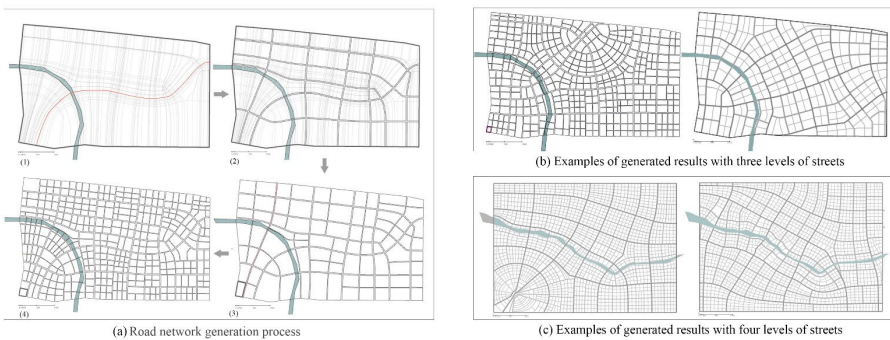


Figure 6. Generation of street network along the tensor field

## 2.2. FABRIC MODELING FROM PROTOTYPE

Rooted in the rural context, many researchers in China have examined the forms, patterns, and structures of the organic settlements (Shou and Zhong, 2016) and generalized the relevant parameters and factors (Ge and Tong, 2017). These rules from rural prototypes are the key to making settlement humanity in scale and form and achieving local distinctiveness representations. We extracted critical rules for streets and plots from others' previous research on three directions, hierarchy, geometry, and topology. The procedural approach synthesizes complex rules to build a hierarchical rural model of streets and plots, and rough buildings are generated to demonstrate the three-dimensional forms of the results.

### 2.2.1. Street Network Generation

The system involves four levels of roads (major roads, secondary roads, branch roads, and alleys) that gradually divide the space into districts, blocks, plots groups, and plots. The generation framework focuses on the geometrical and topological sides of the street network.

The system traces the single street along the tensor field and recursively grows the streets to form a network until the area of the divided blocks meets the requirements. To simulate the organic street forms in villages, nodes of streets are defined in three types: intersection, turning, and deflection nodes (Figure 7). Major streets may have no turning points, while for branches and alleys, the higher the proportion of turning and deflection nodes, the higher the irregularity of the street line. Parameters (street length, width, angle, T-junction rate and intersection rate of the network) with randomly given values within a rational range will lead to diverse generation results (Figure 8).

Several street patterns have been designed to provide choice for users, type of branch, radial, and organic network (Figure 9a). The system can organize several levels of streets in different patterns to obtain results that better match the planning requirements. Several practical factors can affect the values of the parameters; for example, different site locations and traffic volumes can correspond to different connectivity, density, and irregularity of the street network, generating fabric with various features (Figure 9b). Some other operations enrich the generated forms, such as the generation of street public spaces, the consideration of qualitative indicators for population and economy.

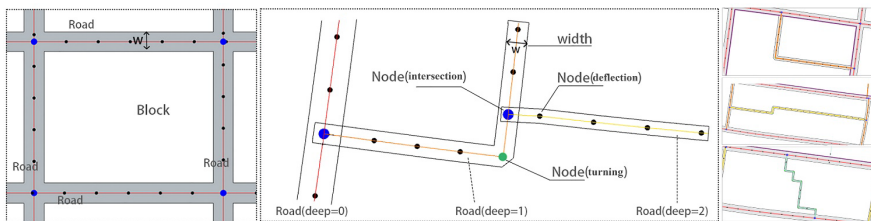


Figure 7. The definition of street lines

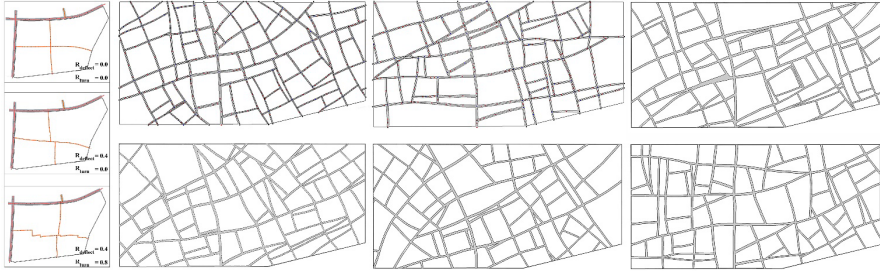


Figure 8. Diverse generation results controlled by parameters

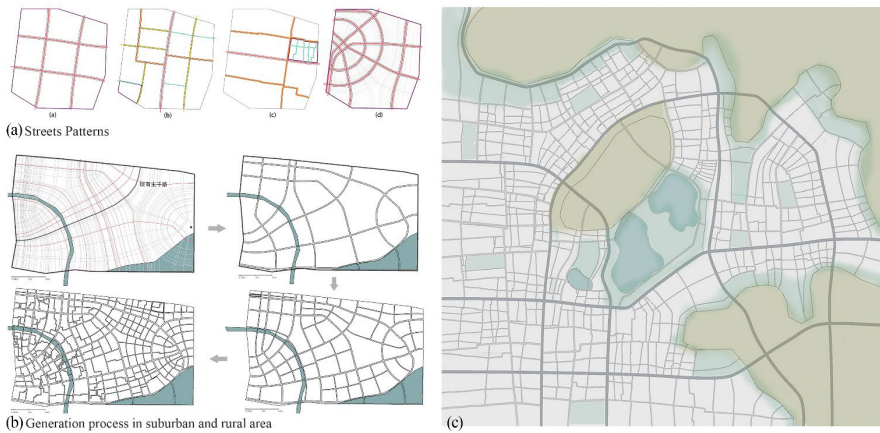


Figure 9. Street patterns defined and examples of generated results

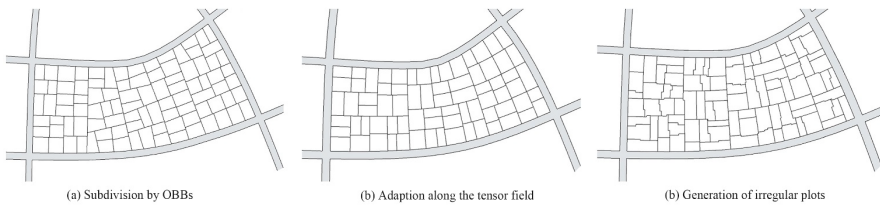


Figure 10. The process of generating plots

### 2.2.2. Plots Generation

The plot is the primary unit for the development and evolution of the block fabric, with the area, aspect ratio, orientation, and land use as its essential control parameters. This study implements adaptive spatial partitioning through subdivision based on oriented bounding boxes (OBB). In order to further adapt the orientation and form of the plots to the local site, the optimization process involves reconstructing the split lines along the tensor field during each iteration of the subdivision, selectively adding turning nodes to generate irregular plots (Figure 10). Through the evolution of the overall methods, the system derives results with generated buildings, as shown in Figure 11.



Figure 11. Generated rural fabric via the presented methods

### 3. Tool implementation

#### 3.1. WEB-BASED USER INTERFACE

This tool on the web uses a WebGL frontend to establish a user-friendly interface. The WebSocket protocol and the JSON data format are applied to connect the ends of the browser and server. Java backend calls the corresponding functions following the received command and passes the user-defined parameters to the functions. It then sends the generated geometries back to the frontend and draws them on the web page (Figure 12).

The tool is designed with a control panel and a settings column that enables mode selection, parameter adjustment, file import, and view conversion. The direct manipulation of geometric elements is supported, for instance, adding or deleting paths in the road network graph, dragging and dropping road nodes, merging adjacent parcels, and adding constraints to alter tensor fields. Also, the results produced on the interface are real-time updated, and relevant statistics can be displayed in parallel. The tool can support both personal use and remote collaboration for future practice, demonstrating the promising paths for sustainable urban planning.

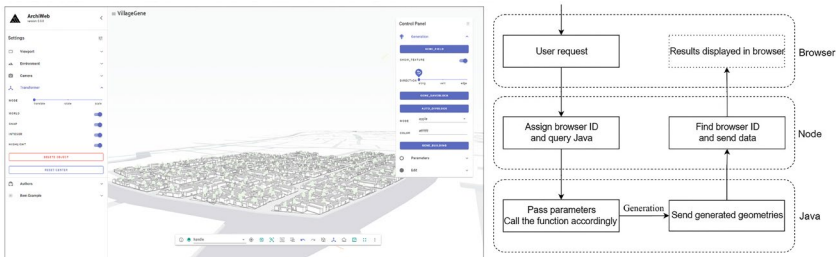


Figure 12. Interface and data pipeline

3.2. DESIGN EVOLUTION PROCESS

In order to finally get a model of the generated block fabric using this tool, the system requires five steps of design evolution when facing application scenarios, including expanding the existing villages and planning new villages nearby, and Figure 13 illustrates the data interaction in the process. The first step is to input the file of the design site, in which each type of polyline is placed in the correct layer as required. Then comes the generation of the underlying tensor field. The user can add axes or nodes based on their design intentions or enter other factor maps as constraints. After clicking on the button to send the command, the field streamlines generated by the backend are drawn on the web page in real-time. A button determines whether the tensor field is visible or not. Next, the streets network can be generated level by level or directly as a whole. The system allows users to locate the critical roads according to the landscape or planning structure before generating the whole network; the road graph can also be edited after generation. The fourth step is to generate the parcels and the general land use layout. The user can define the function and scale of the generated plots and control by editing the partial tensor fields. Eventually, buildings are generated based on the properties of different plots and rendered on the web page. Users can export the model files for further design.

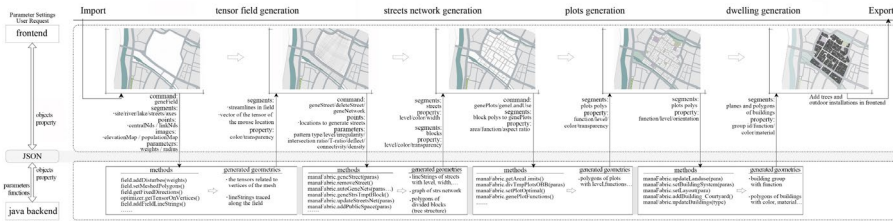


Figure 13. Generation steps and data interaction in the tool

4. Discussion

This paper has provided a deeper insight into the procedural modelling of block fabric in the Chinese rural context and further proposed a web-based generation tool framework that shows important implications for the design of sustainable rural communities. We achieved a more in-depth application of the tensor field to synthesize morphological factors. By complying with the sustainable perspectives, the system realizes the block fabric generation of sustainable rural communities adapted to the local context, natural environment, and local cultural characteristics.

The system's rationality would be more considered in this generation process for further work. More effort needs to be put into analysing the human activities and mechanisms behind the urban fabric. The system can involve more constraints related to social and economic factors for sustainable communities, making it more feasible and flexible. Optimizing the overall tool framework will also be part of further research to supply users with more intuitive and practical control over the design process. The application and development of generative design tools will always be the focus of research in the coming period, illustrating the potential to be an effective way to facilitate the active participation of local residents together with decision-makers,

academics, and professionals in the regional sustainable planning in the future.

### Acknowledgements

This research is funded by National Natural Science Foundation of China (NSFC), Research on Architectural Spatial Combination and Generative System Guided by Eigenvector Matrix Operation (No. 51978139).

### References

- Chen, G., Esch, G., Wonka, P., Müller, P., & Zhang, E. (2008). Interactive procedural street modeling. *ACM Transactions on Graphics*, 27(3), 1–10. <https://doi.org/10.1145/1360612.1360702>
- de Villiers, M., & Naicker, N. (2006). *A Sketching Interface for Procedural City Generation* [Technical report]. University of Cape Town. <https://pubs.cs.uct.ac.za/id/eprint/333/>
- Delmarcelle, T., & Hesselink, L. (1994). The topology of symmetric, second-order tensor fields. *Proceedings Visualization '94*, 140–147.
- Duarte, J. P., Rocha, J. M., & Soares, G. D. (2007). Unveiling the structure of the Marrakech Medina: A shape grammar and an interpreter for generating urban form. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 21(4), 317–349. <https://doi.org/10.1017/S0890060407000315>
- Emilien, A., Bernhardt, A., Peytavie, A., Cani, M.-P., & Galin, E. (2012). Procedural generation of villages on arbitrary terrains. *The Visual Computer*, 28(6–8), 809–818.
- Ge, D., & Tong, L. (2017). Parametric analysis and reconstruction method of village's road morphology. *Journal of ZheJiang University (Engineering Science)*, 51(2), 279–286.
- Greuter, S., Parker, J., Stewart, N., & Leach, G. (2003). Real-time procedural generation of 'pseudo infinite' cities. *Proceedings of the 1st International Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia - GRAPHITE '03*, 87. <https://doi.org/10.1145/604471.604490>
- Kiruthiga, K., & Thirumaran, K. (2019). Effects of urbanization on historical heritage buildings in Kumbakonam, Tamilnadu, India. *Frontiers of Architectural Research*, 8(1), 94–105. <https://doi.org/10.1016/j.foar.2018.09.002>
- Li, B., Guo, Z., & Ji, Y. (2015). Modeling and Realizing Generative Design A Case Study of the Assignment of Ji Village. *Architectural Journal*, 05, 94–98.
- Li, B., & Han, D. Q. (2011). Understanding and prospects of architectural design generation techniques. *Architectural Journal*, 13, 96–100.
- Lipp, M., Scherzer, D., Wonka, P., & Wimmer, M. (2011). Interactive Modeling of City Layouts using Layers of Procedural Content. *Computer Graphics Forum*, 30(2), 345–354.
- Liu, Y., & Li, Y. (2017). Revitalize the world's countryside. *Nature*, 548(7667), 275–277.
- Meng, J. (2015). *Research of urban intermediate structure*. Southeast University Press.
- Mo, Y. (2021). *ArchIndex*. <https://index.archialgo.com>
- Parish, Y. I. H., & Müller, P. (2001). Procedural modeling of cities. *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques - SIGGRAPH '01*, 301–308. <https://doi.org/10.1145/383259.383292>
- Shou, T., & Zhong, W. (2016). The “Matrix” of Ji Village A Study Based on the Construction Strategy of Self-Organization in Rural China. *Architectural Journal*, 08, 66–73.
- Smelik, R. M., Tutenel, T., Bidarra, R., & Benes, B. (2014). A Survey on Procedural Modelling for Virtual Worlds. *Computer Graphics Forum*, 33(6), 31–50.
- Sun, J., Yu, X., Baciu, G., & Green, M. (2002). Template-based generation of road networks for virtual city modelling. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology - VRST '02*, 33. <https://doi.org/10.1145/585740.585747>

# REBUGGING THE SMART CITY

*Design Explorations of Digital Urban Infrastructure*

NICK FÖRSTER<sup>1</sup>, GERHARD SCHUBERT<sup>2</sup> and FRANK  
PETZOLD<sup>3</sup>

<sup>1,2,3</sup> *Technical University of Munich.*

<sup>1</sup>*nick.foerster@tum.de, 0000-0002-4274-8127*

<sup>2</sup>*schubert@tum.de, 0000-0003-1385-5369*

<sup>3</sup>*petzold@tum.de, 0000-0001-8974-0926*

**Abstract.** Smart Cities are presented as a straightforward solution to diverse urban problems. On a closer look, however, the discourse on ‘Smart Cities’ seems wicked in various ways: vaguely defined, speculative, and fragmented into incommensurable positions. Focussing on this ‘wickedness,’ we explore the potential of design approaches to pervade the obscurities and discursive segregations around digital urban infrastructure. Insights from critical design theory lead us to an engagement with digital design not only as validation and enhancement of Smart City projects but as contingent and political exploration. Design becomes an investigation and remaking of what a ‘Smart City’ means in a concrete context. Hence, this approach allows an intersection of social and technical, affirmative and critical perspectives. We explore this approach through an experimental workshop. Hence, we discuss the unfolding of two design engagements: the reframing of ‘Smart Lighting’ as cosmopolitical controversy and the hacking of pedestrian navigation as urban exploration. This approach shows a double potential: On the one hand, it makes digital design practices aware of their ambiguous and political effects. On the other, we scrutinise the possibility of sociotechnical design perspectives as a research approach towards ‘Smart City’ projects and digital urban infrastructure.

**Keywords.** Smart City; Design Theory; Prototyping; Digital Infrastructure; Urban Studies; Critical Making; Speculative Design; SDG 9; SDG 11.

## 1. Introduction

‘By 2050, half of the world's population will live in cities, and cities cause most of the world's CO<sub>2</sub>’ (or similar) precedes most *Smart City* (henceforth: SC) research as a fateful preamble. Following this urgency, optimising such future cities and mitigating disastrous effects seems inevitable. Municipal and corporate-led projects promote the

potential of technical innovations regarding social, ecological, and economic issues. 'Sustainable Cities and Communities' are envisioned through technology and infrastructure. Conversely, urban scholars question this simplistic attitude towards complex urban problems (e.g., Marvin et al., 2016). However, both critical and affirmative approaches often remain in 'technological determinism,' either praising or condemning the SC generically (Fariás & Widmer, 2018, p. 44). Using the term 'Ordinary Smart Cities,' Fariás and Widmer suggest 'decentering' this discussion towards a more contextual and contingent understanding of how SC projects remake urban environments (ibid.). Hence, SC projects are not only confronted with urban complexity but become a part of it. Instead of a univocal and omnipotent solution, the SC itself resembles the 'Wicked Problems' described by Rittel and Webber (1973). Literally, it seems 'ill-defined,' involving heterogeneous aspects such as ISO-norms, governance practices, everyday activities, and digital networks. Thus, the SC entails diverse socio-political conflicts, while its realisation remains vague between actual implementation and wild speculation. This makes it challenging to follow what a SC means within a specific context and how it remakes urban environments.

Rittel and Webber describe 'design' as the reflective practice of exploring, framing and reframing such complex problems (1973). Building upon this perspective, we explore the potential of design approaches to investigate and remake the sociotechnical arrangements of the SC. Instead of planning digital infrastructure as direct 'solutions' regarding specific sustainable development goals, we consider (digital) design as a critical perspective to deconstruct such imperatives and analyse how these goals are negotiated and contextualised within SC projects. The metaphor of 'rebugging' becomes our narrative to engage with friction, contingencies, and possible alternatives through experimental design activities. This paper discusses how (digital) design practices allow new insights regarding two related questions: What is a SC? And what could it be?

On the one hand, we discuss how design practices contextualise the SC in concrete urban contexts. Could design intersect incommensurable perspectives, like affirmative and critical arguments, political and technical engagements? How do SC projects intertwine social practices, political controversies and infrastructure?

On the other hand, we investigate design as an open-ended or even speculative practice, through which we explore contingencies, tensions, and friction within digital infrastructure. How could designing remake monodirectional optimisation as a contingent and political field? Discussing the possible trajectories of SC projects seems crucial regarding the dynamic development of this domain.

We investigate these questions following an experimental workshop. Firstly, we survey perspectives from critical design theory to reconsider (digital) design as a critical approach towards the SC. Thus, we explore how insights from design theory allow a combination of technical engagements and critical perspectives. We describe the unfolding of this approach through two design engagements emerging in the workshop. Thus, we explore how these design activities frame, contextualise, and remake a SC. Finally, we discuss the potential of (digital) design practices to engage with the complex sociotechnical configurations of SC projects. Furthermore, we scrutinise how this sensitivity makes the design of digital urban infrastructure and SC projects more attentive to the contexts and conflicts they are involved in.



## 2. Designing, Making and Fabulating Smart Cities

Design thinking, participatory projects, and experimental prototyping can be considered a key ‘modus operandi’ of many SC initiatives (Tironi & Criado, 2015). Nonetheless, these activities do not lead a priori to a critical engagement. Tironi argues that prototyping often follows a simplistic ‘validating’ rationale instead of exploring contingencies and conflicts. Thus, SC projects are only considered solutions to urban complexity, whereas their inherent ‘wickedness’ is neglected. However, Tironi continues, these processes may still generate unexpected responses and frictions, which are worth analysing (Tironi, 2020). So how to consider design as the open exploration of these contingencies instead of a strategy to solve wickedness?

While McFarlane and Söderström criticise context-ignorant, ‘post-political,’ economy-centred, and technocratic tendencies of SC initiatives (2017, p. 313), Rosner counters similar problems within the domain of design (2018). Reframing design as ‘Critical Fabulation,’ she suggests focussing on contingencies, situated knowledge, collaboration and the continuous challenging of presumptions (Rosner, 2018, pp. 184–186). This perspective resonates with approaches at the intersection of Science-Technology-Studies and design theory. Variations of ‘Making,’ ‘Design,’ and ‘Prototyping’ combine critical perspectives with material and technical engagements (Varga, 2018). For instance, Ratto describes ‘critical making’ as a process-oriented collaborative investigation instead of the search for a single solution. He suggests the reciprocal combination of concrete technical experimentation and theoretical discussions. Hence, ‘critical making’ attempts to resolve the dilemma of ‘innovation’ vs ‘critique’ by ‘reintegrat[ing] technical and social work and thereby innovat[ing] both.’ (Ratto, 2011, p. 258)

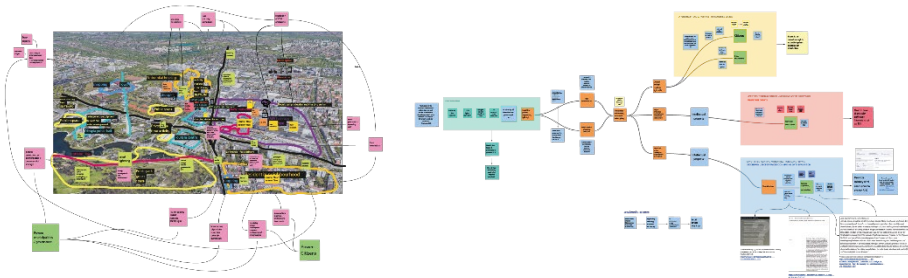
However, this synthesis of critique and innovation should not enforce false consensus over controversial problems. Varga demonstrates how design perspectives invoke a plurality of political engagements: materially inscribed ‘subpolitics’ in the sense of Foucault; controversies arising around artefacts following Latour; or as the composition of a common world through Stenger’s ‘cosmopolitics’ (Varga, 2018, pp. 33–35). Designing becomes a critical activity that deconstructs presumptions, discovers controversies, and slows down to recompose an urban cosmos. For scholars like Wilkie, designing represents a speculative method to explore possible political trajectories of digital technologies (2015). According to Di Salvo, the engagement with prototypes becomes not only a form of inquiry but a potential to confront and remake problematic conditions (2014).

Scholars like Tironi (2018) or Hollands (2015) discussed experimental prototyping to reframe the SC as context-specific, participatory, and knowledge-based. However, these studies mostly stay vague on the role of digital infrastructure within this remade ‘Smartness.’ While these approaches represent a crucial intervention into a technology-dominated discourse, we consider the concrete disposition of digital infrastructure as a relevant part of broader and ‘societal’ discussions around the SC. Hence, we aim to complement these perspectives through a concrete engagement with digital technologies. Building upon the introduced approaches, we discuss the potential of digital design to draw relations between technical and sociopolitical dimensions of the SC, juxtapose critical and affirmative perspectives, and combine analysis with speculative trajectories.

### 3. Investigating the Smart City by Remaking the Smart City

A workshop with a small group of Master level students from Architecture and Landscape Architecture at the Technical University of Munich became an opportunity for this experiment. The workshop was introduced by an intense research phase of three days, followed by three months of project development. This format allowed a playful investigation of digital urban infrastructure based on the described methods. It introduced mapping, reverse-engineering, and prototyping activities to generate a concrete understanding of digital urban infrastructure, intervene speculatively, and explore unexpected trajectories. While we carefully planned a structure of different methods, we were open to contingencies and dynamics. Instead of focussing on final results, this paper follows the workshop process through a series of design engagements.

Though, how to approach a wicked SC? Whereas Ratto (2011, p. 253) suggests combining technical making with critical research and discussions, the first open question was where to start these activities. The discourse on SCs invokes a multitude of topics and heterogeneous topologies, from class diagrams to urban spaces, everyday practices, and infrastructure networks. Furthermore, digital urban infrastructure often seems abstract, immaterial, and context-less. SC projects appear in different states of realisation, ranging from almost banal implementations to Sci-Fi speculation. Inspired by Rittel and Webber, we introduced designing as a practice of framing and reframing what ‘*Smart City*’ pragmatically means. Initially, a series of speculative mapping tasks were introduced as a design-based exploration of the SC and its involvement in urban environments.



*Figure 1: Spatial Mapping and Infrastructural Mapping*

The participants started mapping a selected urban area. They gathered spatial patterns, activities, and infrastructure as a collage of existing maps, diagrams, and sticky notes. They researched different SC projects and discussed how these would integrate into the observed spaces. The resulting cartography served as a first ‘mapping’ of what a SC means in the local context (see Figure 1, at the left). Hence, the map became a speculative tool to contextualise and relate diffuse definitions, speculative projects, and abstract technologies.

Consecutively, the participants focused on one digitally transformed practice or infrastructure in this map. By this, we turned towards the disposition of digital

infrastructure and its entanglements in urban environments. The participants scrutinised the observed system and created a plan of its components, processes, and media. Inspired by the mapping project ‘Anatomy of an AI System’ (Crawford & Joler, 2018), they discussed how the analysed infrastructure reconfigures spaces, generates conflicts and becomes political. Thus, digital infrastructure is interpreted as a ‘matter of concern,’ which interconnects various urban controversies (Latour, 2007, p. 815). Maps of Smart Parking, Energy Grids or Crowd Monitoring emerged and related heterogeneous perspectives like class diagrams and claims for a ‘right to the street’ (see Figure 1, at the right). While these maps showed that digital technologies are neither abstract nor separated from political questions, we wondered how to explore this dimension in more detail and move beyond a dichotomic projection of technology on space or politics on technology.

Hence, we encouraged the participants to choose one situation in their mappings and rearticulate it through a design intervention. After the initial block phase, these experiments were developed through longer and less structured design experiments, accompanied by drawing storyboards, making conceptual prototypes, and open discussions.

### 3.1. SMART LIGHTING COSMOPOLITICS

The first discussed project examined the digital transformation of urban lighting. The initial mapping showed street lanterns in heterogeneous spaces such as parks, parking lots, highways and residential streets. These lanterns were involved in diverse urban activities, from traffic to nightly table tennis. Enhancing streetlights represents a prominent use case for various SC initiatives. Technical research on ‘Smart Lighting’ revealed how digital infrastructure addresses the maintenance, control, and monitoring of street lanterns. Lighting infrastructure is enhanced through sensors, communication networks, and control mechanisms (M. Castro et al., 2013). The technical mapping showed how digital infrastructure reconfigures the circulation of energy, light, control signals, exchanged light bulbs, and sensor data. Additionally, this investigation revealed diverse problems addressed through ‘Smart Lighting:’ saving energy, enhancing the durability of light bulbs, improving maintenance efficiency, or enhancing the security of nightly streets.

Despite this diversity of problems, the examined ‘Smart Lighting’ projects followed the monodimensional rationale of enhancing positive effects (e.g., maintenance efficiency) while reducing negative impact (e.g., resource consumption). However, these clear design goals seemed far less unambiguous on closer examination. For instance, the implied correlation of light and safety turned out more fragile than expected: ‘Walking through a dark park, does light make you secure or perhaps exposed?’ Being lit up may not entail the same safety for everybody in every situation. This point was discussed with historical reference to the 18th century ‘Lantern Laws,’ which made urban lighting an infrastructure of racist oppression. (Browne, 2015). Another discussion thread addressed which needs and interests are considered within an automatically controlled lantern network. On the one side, discussing how SC projects reconfigure urban light became a reverse-engineering of hidden ‘subpolitics’ in the sense of Foucault (Varga, 2018, p. 33). On the other side, the question of who and what to take into account led the project to cosmopolitics in the sense of Stengers

(2005). Thus, the project asked how urban lighting co-creates specific urban atmospheres, which enable particular activities and frustrate others – lingering, walking quickly, jogging, feeling safe?

Leonie, who conducted this project, chose the problem of light regulation as a design brief. She rearticulated these logics of urban lighting to explore their presumptions, effects, and politics. How could this prototype investigate the involved rationales, actors, practices, and interests? The first concept (see Figure 2) introduced a light-jukebox as a metaphorical controller, which allows a tuning of the atmosphere by various actors. While this attempt seemed promising in general, it still involved a commodified and human-centred perspective, which seemed not to do justice to the environmental implications of light.



Figure 2: Storyboard of a Remaking of Smart Lighting  
(Image by Leonie Lux)

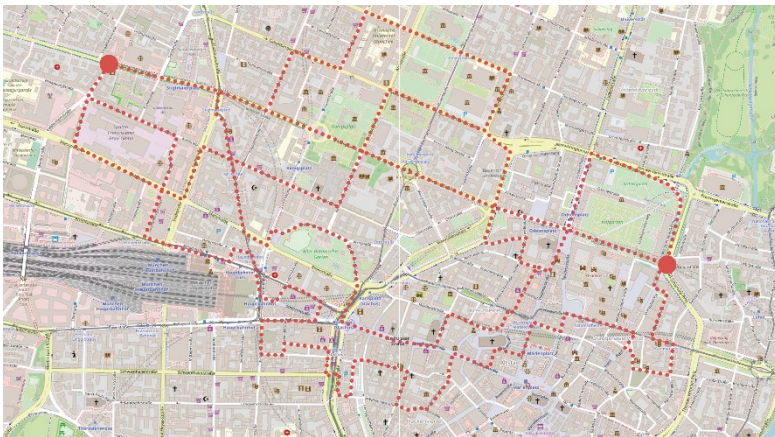
The next concept targeted opening the jukebox towards unexpected actors and incommensurable input. Designing the interactions of this experimental 'Smart Lighting' system became an investigation of *what* and *who* to take into account and *how* to frame the problem of light concerning human usages, safety, resource usage, and the affected fauna. We discussed how to inscribe these different perspectives into the digital control of urban light. The debate revolved around wildlife sensing, youngsters kicking out lanterns, and Dumbledore's deluminator. Finally, the proposal for a process-oriented prototyping project emerged. A few smart home light bulbs would be installed in a park near the university. Their brightness, light colour and animation are modulated through a simple micro-controller. The prototype's interface and control logic would be adapted successively in reaction to the local environment. Thus, the lighting system becomes successively attuned to unexpected conditions. The prototype development would trace the cosmopolitical negotiation over an atmosphere.

Probably because of the non-solutionist focus of this workshop, this promising concept remains an inspiring mock-up until now.

### 3.2. DIJKSTRA'S DRIFT

The second project scrutinised how the SC and digital infrastructure are involved in the walk of pedestrians. A mapping emerged from the initial tasks, laying out different digital technologies along a fictional stroll through the observed area. Thus, this map explored how sensors, databases, GIS, interfaces, and algorithms monitor and influence pedestrian movement. Technologies such as individual smartwatches and air-pollution models co-constitute the practice of *how* and *where* to walk. This collage touched diverse questions of health, safety, traffic, privacy. Ekaterina, who was responsible for this project, focused on the practice of pedestrian wayfinding for further exploration. How is the choice, where to walk, remade through air-pollution maps, traffic data, models, and algorithms?

Even though ‘wayfinding’ currently occurs individually and in the private sector, we did not want to delimit the discussion of what a SC means beforehand. Thus, Ekaterina investigated design interventions into the practice of navigation. A first prototype framed navigation as a negotiation among different factors instead of just the ‘shortest route.’ The user scenario depicted a parent and a child who playfully create a way to school which is safe and fun. A multi-criteria navigation app would support this wayfinding. A discussion emerged around this prototype: On the one hand, the ‘learning’ aspect and the exploration of different criteria seemed interesting. On the other, ‘improving the way to school’ appeared to fall into the trap of optimisation and securitisation. The prototype established a deterministic relationship between the user and digital infrastructure, child and parent. Hence, we wondered how to open this standard SC narrative up through prototyping.



*Figure 3: Resulting Journeys of the Dijkstra's Drift  
(Project by Ekaterina Tepliakova)*

Instead of a deterministic relationship, Kitchin and Dodge describe the involvement of digital technology in urban practices as a ‘collaborative,’ ‘contingent,’ and ‘context-specific’ ‘transduction’ (2011, p. 80). Following this thought, navigation becomes a sociotechnical negotiation between pedestrians and technology. This perspective resonates with de Certeau’s seminal description of walking as the tactical *reaction to*

and *re-appropriation of top-down planned spaces and infrastructure* (1984).

A second prototype started with a playful investigation of Google Maps. First, the application generated connections between two places. These routes were distorted and dragged according to different criteria by mouse clicks and disobedient walking. For a more detailed insight into this interaction, Ekaterina turned towards technical wayfinding models, namely Dijkstra's Algorithm. This algorithm addresses the shortest-way problem through a weighted network graph. A cost function includes different criteria in this weighting. (Velden, 2014) In the case of navigation, these are usually distance, congestion, or simplicity.

We discussed how to reinterpret these means-to-reach-an-end as devices for urban exploration through several prototyping steps. The project took inspiration from situationist psychogeography. This playful mapping practice rearranges the structuring and understanding of urban spaces by strolling/drift through the city (Flanagan, 2013, pp. 194–197). How could we remake navigation as drift through both physical and digital spaces?

Ekaterina experimented with different cost functions, adding weight to Dijkstra's network graph. Thus, she played with different logics to connect locations on a map: the shadiest, hottest, greenest, or most accident-ridden route (see Figure 3). By this, the data sets were experimentally related to navigation. According to these rulesets, the algorithm generated semi-absurd journeys to stroll through urban spaces. The walk along these navigation graphs was documented as video. These movies showed experiences of urban spaces, which were co-constituted by experimental navigation rules. Prototyping became an exploration of different rationales of moving through the city. On a second layer, the prototype allows exploring how these logics intersect with concrete urban spaces. Hence, it enabled the discovery of the underlying datasets in urban space. How is the experience of the 'highest ranked' way? How does it feel to wander through 'hot' climate data? The prototype recontextualises datasets and algorithms in the pedestrian's experience. Thus, the documented journey moved through both urban space and the navigation model. It remakes navigation as an open-ended exploration of the physical *and* the smart city.

#### 4. Discussion

This paper described the implementation of SC projects in urban environments as a wicked problem and experiments with design approaches to understand and remake how SCs reconfigure urban spaces and practices. Hence, we reviewed perspectives from critical design theory. Furthermore, we explored experimental and playful design interventions in a workshop and followed two projects through several design activities. By this, we discovered the sociotechnical involvement of digital infrastructure in concrete urban environments as well as possibilities to remake these constellations. In conclusion, we discuss what insights this perspective allows and how it is relevant for an extended digital design practice.

We introduced design activities to contextualise and reframe what a SC represents in a specific context – and what problems it entails. The initial mappings and the following prototyping tasks successively connected different topologies and aspects, like technical diagrams, urban spaces, and political issues. Furthermore, these

approaches concretised abstract and speculative SC projects through sociotechnical engagements. The first prototype revealed the complex social and political entanglements of lighting infrastructure. The 'Dijkstra's Drift' intimately related the abstract network space of digital models and the pedestrian's everyday experience.

Furthermore, both projects allowed de-centring the teleological focus of digital design and exploring contingencies in the SC. Engaging with 'Smart Lighting' deconstructed the usual rationale of 'optimisation' and revealed the political dimensions of this infrastructure. The second prototype reconsidered the target-oriented function of 'wayfinding' as the exploration of hybrid urban spaces. Thus, design engagements allowed critical insights by challenging presumptions and exploring alternative trajectories. However, this criticality was not a given but emerged through design activities, from continuously intersecting technical perspectives with theoretical discussions and reacting to nagging questions. Both projects even shifted back and forth between solutionist and explorative approaches. Optimistically, we interpret this as a productive co-existence of technical and critical perspectives.

On the one side, the presented methods allowed insights into the involvement of SC projects in urban environments. Of course, these methods are not a substitute for ethnographic or theoretical investigations of the SC. Nonetheless, the described design approaches could complement such studies, offering an exploration of the sociotechnical disposition of digital infrastructure. On the other side, the discussed methods represent a valuable contribution to digital design practices in the SC context. They challenge inherent presumptions and investigate how the developed projects entail contingent effects. At least, the proposed playful approaches offer relaxation to an often tense and deterministic discourse.

Nonetheless, the presented perspectives bear various potentials for further exploration. Even if the workshop aimed to contextualise digital technologies, both presented prototypes were only loosely related to concrete urban spaces. For future work, it is essential to explore how these practices reconfigure specific urban environments and how they are affected reciprocally. While this paper focussed on experimenting with digital technologies, a stronger connection to urban contexts remains a vast potential. Also, due to the current pandemic, the collaborative and participatory aspects of the introduced design approaches fell short. However, including situated perspectives and experiences is crucial to understanding how SC projects concretely affect different groups and stakeholders. Collaboration becomes an essential aspect for the further exploration of the presented methods.

Thus, this experiment only represents our first design steps towards wicked smartness. The documented activities made us problematise and rearticulate the SC in various ways. The workshop made it move from 'By 2050...' to urban spaces, negotiations around urban light, and situationist navigation tools. Where else could design take the Smart City?

### **Acknowledgements**

We thank all participants of the described workshop and hope that confusion and success were well balanced. Especially, we are grateful to Leonie Lux and Ekaterina Tepliakova, who allowed us to share their process and prototypes.

## References

- Browne, S. (2015). *Dark matters. On the surveillance of blackness*. Duke University Press.
- Castro, M., Jara A. J., Skarmeta, A. F. G. (2013). Smart Lighting Solutions for Smart Cities. In *27th International Conference on Advanced Information Networking and Applications Workshops* (pp. 1374–1379).
- Crawford, K. & Joler, V. (2018). *Anatomy of an AI System. The Amazon Echo as An Anatomical Map of Human Labor, Data and Planetary Resources*. Now Institute and Share Lab. Retrieved November 23, 2021, from <https://anatomyof.ai/>.
- De Certeau, M. (1984). *The practice of everyday life*. University of California Press.
- DiSalvo, C. (2014). Critical Making as Materialising the Politics of Design. *The Information Society*, 30(2), 96–105. <https://doi.org/10.1080/01972243.2014.875770>.
- Farias, I. & Widmer, S. (2018). Ordinary Smart Cities. How Calculated Users, Professional Citizens, Technology Companies and City Administrations Engage in a More-than-digital Politics. *Tecnoscienza: Italian Journal of Science & Technology Studies*, 8, 43–60.
- Flanagan, M. (2013). *Critical play. Radical game design*. MIT Press.
- Hollands, R. G. (2015). Critical interventions into the corporate smart city. *Cambridge J Regions Econ Soc*, 8(1), 61–77. <https://doi.org/10.1093/cjres/rsu011>.
- Kitchin, R. & Dodge, M. (2011). *Code/space. Software and everyday life*. MIT Press.
- Latour, B. (2007). Turning Around Politics. *Soc Stud Sci*, 37(5), 811–820. <https://doi.org/10.1177/0306312707081222>.
- Marvin, S., Luque-Ayala, A., & McFarlane, C. (2016). *Smart urbanism. Utopian vision or false dawn?* Routledge, Taylor & Francis Group.
- McFarlane, C. & Söderström, O. (2017). On alternative smart cities. *City*, 21(3-4), 312–328. <https://doi.org/10.1080/13604813.2017.1327166>.
- Ratto, M. (2011). Critical Making: Conceptual and Material Studies in Technology and Social Life. *The Information Society*, 27(4), 252–260. <https://doi.org/10.1080/01972243.2011.583819>.
- Rittel, H. W. J., Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sci*, 4(2), 155–169. <https://doi.org/10.1007/BF01405730>.
- Rosner, D. K. (2018). *Critical Fabulations. Reworking the Methods and Margins of Design*. MIT Press.
- Stengers, I. (2005). The Cosmopolitical Proposal. In B. Latour und P. Weibel (Eds.), *Making things public. Atmospheres of democracy* (pp. 994–1003). MIT Press, ZKM Center for Art and Media.
- Tironi, M. & Criado, T. S. (2015). Of Sensors and Sensitivities. Towards a Cosmopolitics of "Smart Cities"? *TECNOSCIENZA: Italian Journal of Science & Technology Studies*, 6(1), 89–108. <http://www.tecnoscienza.net/index.php/tsj/article/view/217>.
- Tironi, M. (2018). Speculative prototyping, frictions and counter-participation. A civic intervention with homeless individuals. *Design Studies*, 59, 117–138. <https://doi.org/10.1016/j.destud.2018.05.003>.
- Tironi, M. (2020). Prototyping public friction: Exploring the political effects of design testing in urban space. *The British journal of sociology*, 71(3), 503–519. <https://doi.org/10.1111/1468-4446.12718>.
- Varga, H. M. (2018). On Design and Making with STS. *Diseña* (12), 30–51. <https://doi.org/10.7764/disena.12.30-51>.
- Velden, L. (2014). *Der Dijkstra-Algorithmus*. Technical University of Munich. Retrieved November 24, 2021, from [https://algorithms.discrete.ma.tum.de/graph-algorithms/spp-dijkstra/index\\_de.html](https://algorithms.discrete.ma.tum.de/graph-algorithms/spp-dijkstra/index_de.html).
- Wilkie, A., Michael, M. & Plummer-Fernandez, M. (2015). Speculative Method and Twitter: Bots, Energy and Three Conceptual Characters. *The Sociological Review*. 63(1), 79–101. <https://doi.org/10.1111/1467-954X.12168>.



# MACHINE LEARNING-BASED WALKABILITY MODELING IN URBAN LIFE CIRCLE

PIXIN GONG<sup>1</sup>, XIAORAN HUANG<sup>2</sup>, CHENYU HUANG<sup>3</sup> and MARCUS WHITE<sup>4</sup>

<sup>1,2,3</sup>*School of Architecture and Art, North China University of Technology*, <sup>4</sup>*Swinburne University of Technology*.

<sup>1</sup>*gongpixinn@ncut.edu.cn, 0000-0003-4795-666X*

<sup>2</sup>*xiaoran.huang@ncut.edu.cn, 0000-0002-8702-2805*

<sup>3</sup>*huangchenyu303@163.com, 0000-0002-6360-638X*

<sup>4</sup>*marcuswhite@swin.edu.au, 0000-0002-2238-9251*

**Abstract.** With China's fast urbanization, the study of the walkability of residents' life circles has become critical to improve people's quality of life. Traditional walkability calculations are based on Lawrence Frank's theory. However, the weighted calculation method cannot be adapted to ever-changing and complicated scenarios as the scope and topic of research transforming. This study investigated walkability at the community level by combining machine learning techniques with multi-source data. Feature indicators affecting walkability were estimated from multi-source data. Machine learning was used to refine the weighting calculation under the previous indicator framework. We compared the performance of 20 regression models from 6 different machine learning algorithms for estimating the walkability of 14578 communities in downtown Shanghai. It is concluded that the Bagged Tree Model ( $R^2=0.86$ ,  $RMSE=0.36862$ ) achieves the best performance, which is used to revise the initial walkability index values. The workflow proposed in this paper allows for rapid application of expert empirical consensus to comprehensive urban design and detailed urban governance in the future.

**Keywords.** Life Circle; Walkability Indicator; Multi-source Data; Machine Learning; Refined Urban Design; SDG 3; SDG 10; SDG 11.

## 1. Introduction

Walking is the primary mode of short-distance travel for urban residents. High urban walkability can effectively increase residents' willingness to travel and boost urban vitality. However, during rapid urbanization, overemphasis on automotive traffic in metropolitans has raised issues such as traffic congestion and unpleasant walking environments (Vichiensan and Nakamura, 2021). Nowadays, the attention of urban construction has shifted to the living areas of ordinary inhabitants, and a healthy and sustainable urban life is essential for urban development. The life circle is a concept that has emerged in recent years to represent the walkable spatial range of activities of

community residents (Sun et al., 2012). The study of the walkability of residents' life circles has become a central topic for improving people's quality of life and increasing their health outcomes (Frank et al., 2010).

The advancement of big data and artificial intelligence technology in recent years has enabled large-scale, fine-grained studies. Compared to conventional studies with limited samples, new urban science research approaches based on the combination of massive urban data and machine learning algorithms are more efficient, fine-grained, and rapidly generalizable. With a large enough sample size, the application of machine learning algorithms to human-scale investigations of pedestrian environment quality could produce results with a high degree of trustworthiness and generalization performance.

In the previous studies, there is an increasing interest in determining the effect of built environment features on habitual physical activity. Frank et al., (2006), proposed a "walkability index" system as a composite measure of the built environment by summing the z scores for net residential density, street connectivity, land use mix, and retail floor area ratio for each census urban block, with street connectivity receiving twice the weight as the other three variables. Ewing and Handy, (2009), developed the concept of "Measuring the Unmeasurable," which quantifies five urban design qualities through the physical characteristics of streets and their edges: imageability, enclosure, human scale, transparency, and complexity. Daniel and Burns, (2018), incorporated the often-overlooked steepness variable as a key statistical element of the walking experience. Besides, thermal comfort also has been progressively promoted as a critical measure for walkability (Labdaoui et al., 2021). With the growth of modern urban science, new forms of data are being used to research urban walkability, including urban land use data (Wei et al., 2016), point of interest (POI) data (Zandieh et al., 2017), and location-based services (LBS) data (Yamagata et al., 2020). In terms of more recent research approaches, machine learning methods have been used to predict walkability (Yin and Wang, 2020). However, the majority of research has been undertaken at the metropolitan planning level, with an emphasis on the linear space of the street and little attention paid to the walkability of the life circle. Additionally, the prior studies have frequently employed the regular circular buffer zone or grid as the research unit, although the range of human mobility is not same in all directions. In this paper, the spatial boundary of the live circle is an irregular shape generated along with the road network, reflecting the possible range of activities of the residents.

## 2. Methods

In this study, based on the walkability framework proposed by Frank, this paper estimated the feature indicators and the walkability of the life circle through multi-source data. Moreover, the mapping of impact indicators to walkability was established by machine learning. Multi-source data about urban areas were collected from commercial and government websites. Prior to the data analysis, all database were purged to reduce the computation latency. Each indicator's values were normalized using the Z-score method, which maps all data ranges to the same interval for comparison analysis. Each indicator's results were calculated and presented using a vector geographic information system (GIS). The dataset for machine learning consists of walkability and five feature indicators. Supervised regression and six classes of

machine learning algorithms were chosen to model the dataset to avoid the inductive bias. The results were compared using a ten-fold cross-validation method, and the prediction model with the highest overall performance was selected for model optimization and application.

## 2.1. RESEARCH SITE

Shanghai is a classic Chinese metropolitan city which with numerous challenges to be addressed. One of the most notable issues is the low quality of walkability in the residents' life circle under scenario of rapid urban development. Fortunately, due to the advancement of the urban management technology, data is more transparent and freely accessible than ever before, offering possibilities for large-scale and refined measurement. In this study, seven districts of downtown Shanghai were chosen as the research areas, namely Huangpu District, Xuhui District, Changning District, Jing'an District, Putuo District, Hongkou District, and Yangpu District.

The indicators calculated in this paper are based on community life circles (Sun et al., 2012). The data on pedestrian traffic and other time circles were gathered using the open map platform Mapbox. A Python crawler application was created to extract 14578 communities (containing location coordinate, household count, average property price, and other information) from the commercial website (fang.com). The location coordinates of the communities were used to link with the Mapbox's Isochrone API to obtain geographic data of the life circle of all communities. The range of the life circle is set as a defined length of time from the central point of the community. The time threshold for this study was set at 15 minutes in accordance with the Shanghai Master Plan's current urban development criteria. The research site is shown in figure 1.



Figure 1. Research precincts in Shanghai

## 2.2. CALCULATION OF FEATURE INDICATORS

The following four feature indicators (land use mix, residential density, street connectivity, and retail density) are derived from Frank's system of walkability research. With the dawn of big data era, newly developed urban data types are more successful in picturing the indicators' dimensions. Thus, the data types used in this paper have been changed. Additionally, the unit of study was adjusted to a more realistic walking life circle based on trip time costs.

### *2.2.1. Land Use Mix*

The “land use mix” indicator is calculated using the ArcGIS platform for processing the land-use data of Shanghai. Nine land use categories were employed in the calculation (urban residential land, industrial, mining and storage land, public infrastructure land, public building land, transport land, commercial service land, landscape land, special land and unused land). 50m\*50m grids were used to cut the land use vector data, and the ratio of each land use category within each grid was estimated. The index of land use mix was calculated using the entropy approach (see table 1 for the calculation formula). In ArcGIS, the polygon data were translated into point data. The number of grid points in each life circle was tallied, and thereafter the entropy values were averaged and superimposed. The final result was used as the land use mix index value for each life circle.

### *2.2.2. Residential Density*

In comparison to demographic data, data gathered from commercial websites at the community level is more accurate, current, and convenient to use. The overall number of households in each community is derived from the housing stock's commercial websites (fang.com). According to the Shanghai Statistical Yearbook 2020, the average Persons per Household is 2.64, and the two figures are multiplied to determine the total population of each community. Due to the life circle does not comprise a single community, the population of all communities inside the live circle is aggregated and divided by the area to generate the “residential density” indicator.

### *2.2.3. Street Connectivity*

The “street connectivity” indicator is associated with ease of travel and plot accessibility. The Shanghai road network data was obtained via the Baidu API. After cleaning and processing the data, the road network was topologically processed in ArcGIS to extract road intersection nodes and spatially connect them within each community's life circle in order to calculate the number density of road intersections as a measure of street connectivity.

### *2.2.4. Retail Density*

Walkability is affected by the number of retail enterprises inside a residential area, and good accessibility boosts inhabitants' willingness to travel. Frank believed that the Retail Floor Area Ratio might increase the sensitivity to retail use, which was supposed to boost pedestrian traffic. However, the commercial floor space ratio does not always correspond to actual retail use, and Baidu's POI data can provide a full view of the city's current commercial activity. The results of this study were derived from Baidu Vector Map, and the retail density for each life circle was estimated.

## 2.3. CALCULATION OF WALKABILITY INDICATOR

### *2.3.1. Crowd Density*

High community walkability results in increased pedestrian flow, (Ewing et al., 2006),

and this study characterizes the walkability of life circles with pedestrian flow density. Traditional statistical approaches based on field observations are no longer adequate for large-scale fine-grained measurement. With the present prevalence of mobile internet, it has become fairly typical for people in a lingering, slow-walking position in their daily lives to use their mobile phones for reading, socializing, navigation, shopping, and enjoying other mobile internet services. When consumers use mobile devices such as smartphones to access APPs, they are typically standing or walking slowly. These mobile internet services generate the user's LBS location data, which includes the user's GPS-adjusted location and associated timestamp. This form of data is well-suited for sensing the intensity of activity in a community, as it records the position and duration of people's halting behaviour over a vast region and over an extended period. In this scenario, the data is a raster of crowd density derived through LBS data processing. The vector data was segmented into 100m\*100m grids in ArcGIS to determine the average population density within the life circle. Finally, the density of the crowd was determined by dividing the area of the life circle.

Table 1. The formulas of indicator

Indicator	Formula	Description
Land Use Mix	$H(X) = -\sum_{i=0}^n P_i \ln P_i$	$H(X)$ denotes the entropy of sample $X$ ; $P_i$ is the ratio of different site types.
Residential Density	$RD = \left( \sum_{i=0}^n K * X_i \right) / area$	$K$ is person-to-household ratio of 2.64, which is the number of households in the neighborhoods included in the life circle
Street Connectivity	$ID = I_i / area$	$I_i$ is the number of road intersections within the life circle and the area of the life circle
Retail Density	$RD = P_i / area$	$P_i$ is the number of POIs in the life cycle, and the area of the life circle
Crowd Density	$CD = \left( \sum_{i=0}^n C_i \right) / area$	$C_i$ is the number of crowd at the grid points contained within the life, and cycle e area of the life circle

#### 2.4. MACHINE LEARNING MODELLING

Prior to machine learning, the data in this study is dimensionless processed to eliminate unit limitations, allowing indications of varying units of magnitudes to be compared and weighted. In this study, each indicator is normalized by the method of Z-score.

$$y_i = \frac{x_i - \bar{x}}{s}, \text{ where } \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$$

The findings of Frank Lawrence's walkability study are based on sociological research and statistical analysis. The results are restricted by the data's trustworthiness and sample size, while the weighting of the four variables is less applicable. Moreover, the findings of studies undertaken in developed countries do not apply entirely to Chinese cities. This work offers a machine learning approach for supervised regression analysis in order to rectify the weight coefficients of each indicator under the Shanghai context.

To circumvent the issue of inductive bias, six machine learning algorithms were

chosen for the purpose of building prediction models for comparative examination. Each category was subdivided into distinct subcategories, yielding twenty prediction models. The 10-fold cross-validation was conducted to evaluate the models. The data set was divided into ten mutually exclusive subsets, with a concatenated set of nine subsets serving as the training set and the remaining subsets serving as the test set. With the acquisition of ten training and test sets, ten training and testing sessions are completed in total. Finally, the test findings' mean values are returned for further research (Figure 2).

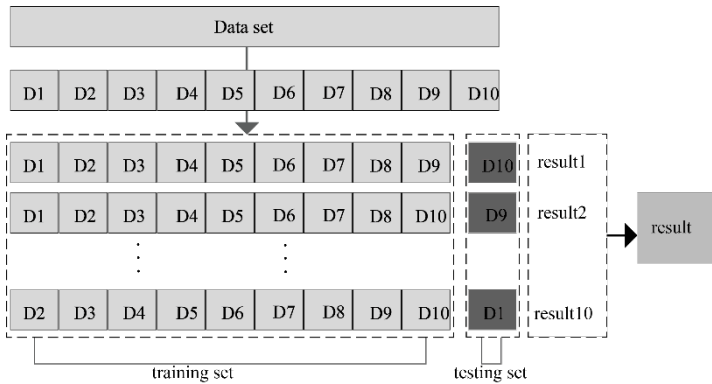


Figure 2. The process of 10-fold cross validation

### 3. Results and Discussion

#### 3.1. RESULTS OF FEATURE INDICATORS

Figure 3 shows the results of feature indicators. The result of the "Land use mix" computation is depicted in Figure 3(a). At the community level, the values of "Land use mix" are uniformly distributed without any discernible spatial agglomeration; At the sub-district level, DaNinglu, GongHeXinlu, HuaiHaiZhonglu, and BeiWaiTan have higher values, with linear spatial distributions and a general tendency along two water systems; At the district level, XuHui District and JingAn District have greater values, whereas ChangNing and YangPu districts have lower values.

The "Residential density" calculation outcome is shown in Figure 3(b). The distribution of residential density values reflects a characteristic of "low in the centre and high on the periphery," which is related to Shanghai's urban development: downtown has a more traditional residential plot with a lower floor area ratio, whereas the periphery of the city has a higher construction intensity due to land prices and policies, reflecting the obvious separation of work and residence in Shanghai; At the sub-district level, HuaYanglu, JiangSulu, XinHuaLu, JiangWan, and TaoPu have greater values, and the majority of them are located on the outskirts of the city centre; ChangNing District has the highest value, while Hongkou District has the lowest.

The result of the "Connectivity density" calculation is depicted in Figure 3(c). At the community level, the higher-valued living circles are primarily located in the northern section of the city center, where the road network is more dense and

connectivity is better. Dinghai Road, Ouyang Road, Quyang Road, North Sichuan Road, North Bund, and Tianping Road all have greater values at the sub-district level, although the sub-districts in Huangpu District generally have lower values, likely because the old city has more old communities; Hongkou and Xuhui districts have the highest values, while Huangpu District and Jing'an District have the lowest.

Figure 3(d) illustrates the result of the "Retail density" calculation. At the community level, life circles with a higher retail density are more likely to be found in the northern and southernmost portions of the downtown, implying that these communities are more accessible; At the sub-district level, Huanan Road, Ouyang Road, Sichuan North Road, Beiwai Tan Road, and Tianping Road all have greater values; Xuhui District and Hongkou District have higher total values, while Changning District and Jing'an District have lower values.

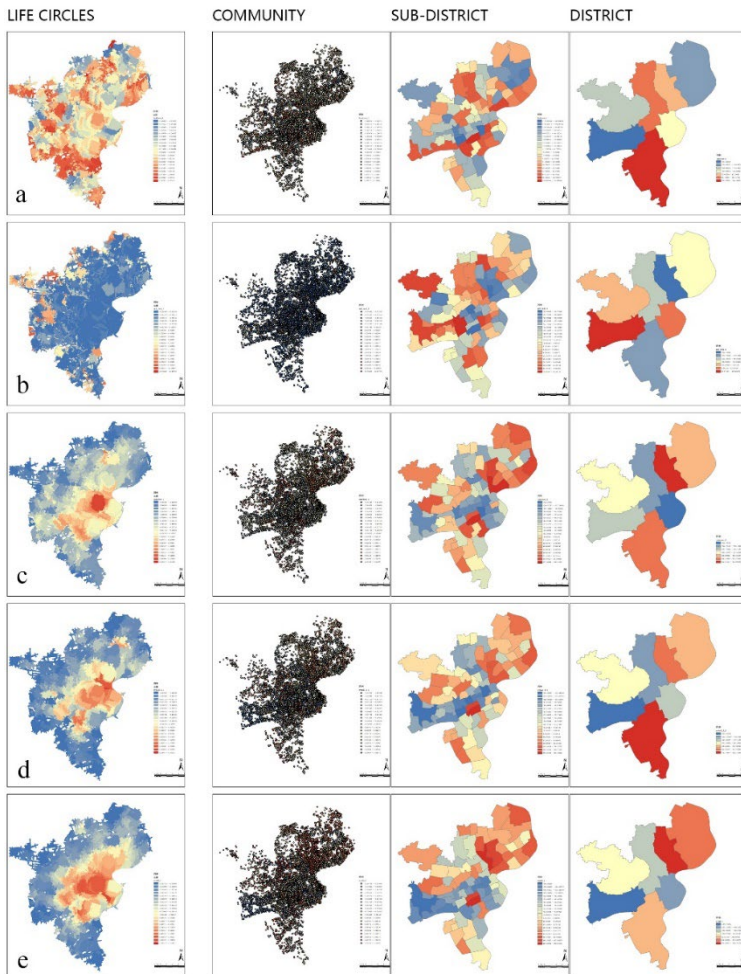


Figure 3. the simulation result of 5 indicators

The outcome of the "Crowd density" calculation is shown in Figure 3(e). At the community level, areas with a higher pedestrian density are located in the northern portion of the city centre, where additional elements, such as universities and commercial districts, draw more pedestrians due to their abundance and variety of activities. Guangzhong Road, Huanan Road, Ouyang Road, Sichuan North Road, and Tianping Road all have greater crowd density ratings at the sub-district level, and these sub-districts also have a higher retail density. Yangpu and Hongkou districts have greater overall values, whereas Changning District and Huangpu District have lower overall values. It is evident that older urban districts are less appealing to pedestrians, indicating the importance of optimizing the quality of life in older urban regions.

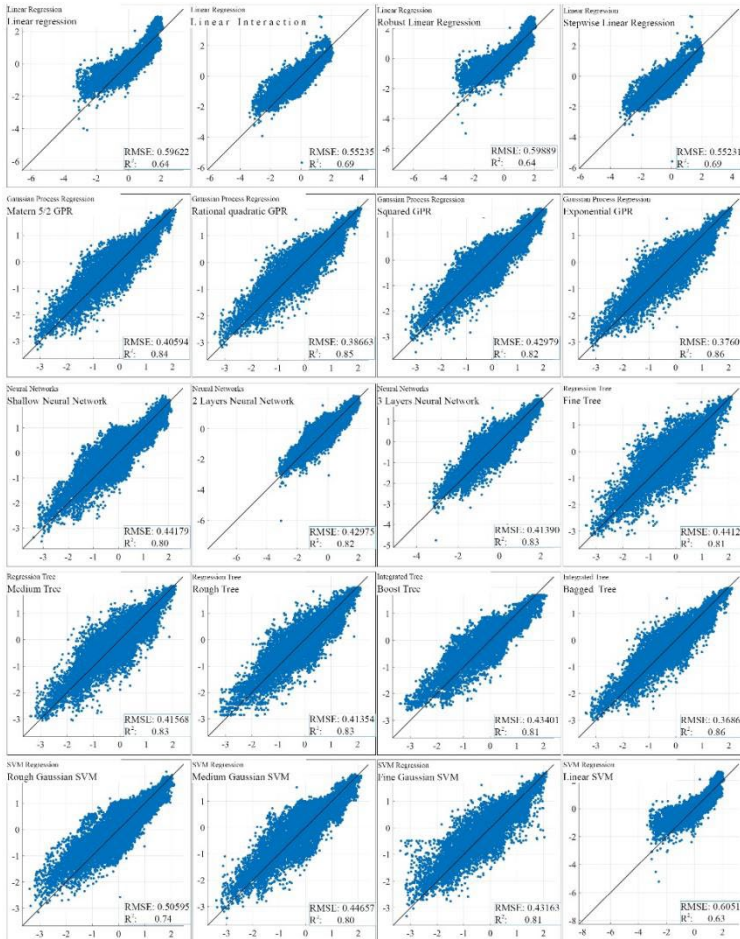


Figure 4. Regression performance of 20 models

### 3.2. COMPARISON OF DIFFERENT MACHINE LEARNING MODELS

Six different types of machine learning algorithms are used in this article to create



twenty machine learning prediction models. The tenfold cross validation approach is used to determine the model's generalization error and find the optimal selection. The outcomes are depicted in the figure 4.

As illustrated in the figures, both the bagged tree and exponential Gaussian process region prediction models have the highest Goodness of Fit, with an  $R^2$  of 0.86. However, because the Root Mean Square Error of the bagged tree model (RMSE=0.36862) is less than that of the exponential Gaussian process region model (RMSE=0.37609), the former is clearly superior. The prediction model implements Lawrence Frank's walking measurement index system's parameter correction. The developed model may be appropriate for future investigation. Simultaneously, the machine learning model may be imported into grasshopper via the MATLAB API interface and integrated into the urban design workflow.

Table 2. Model classification and evaluation results

CATEGORIE	SUBCATEGORIE	RMSE	R <sup>2</sup>
Linear Regression	Linear interaction	0.55235	0.69
	Linear	0.59622	0.64
	Robust linear	0.59889	0.64
	Stepwise linear regression	0.55231	0.69
Regression Tree	Fine tree	0.44125	0.81
	Medium tree	0.41568	0.83
	Rough tree	0.41354	0.83
SVM Regression	Linear SVM	0.60513	0.63
	Fine Gaussian SVM	0.44125	0.81
	Medium Gaussian SVM	0.41568	0.83
	Rough Gaussian SVM	0.41354	0.83
Integrated Tree	Bagged tree	0.36862	0.86
	boost tree	0.43401	0.81
Gaussian Process Regression	Squared GPR	0.42979	0.82
	Matern 5/2 GPR	0.40594	0.84
	Exponential GPR	0.37609	0.86
	Rational quadratic GPR	0.38663	0.85
Neural Networks	Shallow neural network	0.44179	0.8
	2 layers neural network	0.42975	0.82
	3 layers neural network	0.4139	0.83

#### 4. Conclusions

As urbanization progresses towards the infill phase, study on the walkability of urban space becomes critical for promoting urban people' walking activities and enhancing urban vitality. This study, based on Frank Lawrence's walkability theory, examines walkability at the community-scale by combining machine learning techniques with multi-source data. In order to compare the walkability of a total of 14,578 communities in downtown Shanghai, 20 regression models were created in this research using 6

different types of machine learning algorithms, with the Bagged Tree Model ( $R^2=0.86$ ,  $RMSE=0.36862$ ) achieving the greatest result. The machine learning model, which is highly suited for studies of urban walkability in China, can be integrated into urban design processes in the future by importing it into Grasshopper via MATLAB's API.

Of course, this paper contains flaws as well. With the rise of interdisciplinary research, additional indicator dimensions could have been included in the study of walkability. However, because this paper is based on Frank's theoretical findings, it focuses exclusively on the impact of the four indicator dimensions on walkability, leaving out other indicator dimensions. Meanwhile, due to the limited granularity of the data, there are some issues with the development of various indicators, which will be gradually resolved in subsequent research.

## References

- Daniel, P., & Burns, L. (2018). How steep is that street? Mapping 'real' pedestrian catchments by adding elevation to street networks. *Radical Statistics*, 121, 26-48.
- Ewing, R., & Handy, S. (2009). Measuring the unmeasurable: Urban design qualities related to walkability. *Journal of Urban Design*, 14(1), 65-84.
- Ewing, R., Handy, S., Brownson, R.C., Clemente, O., Winston, E., 2006. Identifying and measuring urban design qualities related to walkability. *J. Phys. Act. Health*, 3, S223
- Frank, L., Devlin, A., Johnstone, S., & Loon, J. V. (2010). Neighbourhood design, travel, and health in Metro Vancouver: Using a walkability index.
- Frank, L. D., Sallis, J. F., Conway, T. L., Chapman, J. E., Saelens, B. E., & Bachman, W. (2006). Many pathways from land use to health: associations between neighborhood walkability and active transportation, body mass index, and air quality. *Journal of the American planning Association*, 72(1), 75-87.
- Labdaoui, K. , Mazouz, S. , Moeinaddini, M. , Cools, M. , & Teller, J. . (2021). The street walkability and thermal comfort index (swtci): a new assessment tool combining street design measurements and thermal comfort. *Science of The Total Environment*, 148663.
- Sun, D., Shan, S., & WuTinghai. (2012). Life circle theory based county public service distribution:jiangsu pizhou case. *Planners*.
- Vichiensan, V., & Nakamura, K. (2021). Walkability perception in asian cities: a comparative study in bangkok and nagoya. *Sustainability*, 13.
- Wei, Y. D., Xiao, W., Wen, M., & Wei, R. (2016). Walkability, land use and physical activity. *Sustainability*, 8(1), 65.
- Yamagata, Y., Murakami, D., & Yoshida, T. (2020). Evaluating walkability using mobile GPS data. In *Spatial Analysis Using Big Data* (pp. 239-257). Academic Press.
- Yin, L., & Wang, Z. (2016). Measuring visual enclosure for street walkability: Using machine learning algorithms and Google Street View imagery. *Applied geography*, 76, 147-153.
- Zandieh, R., Flacke, J., Martinez, J., Jones, P., & Van Maarseveen, M. (2017). Do inequalities in neighborhood walkability drive disparities in older adults' outdoor walking. *International journal of environmental research and public health*, 14(7), 740.

# ADDRESSING FLOOD RESILIENCE IN JAKARTA'S KAMPUNGS THROUGH THE USE OF SEQUENTIAL EVOLUTIONARY SIMULATIONS

KIM RICAFORT<sup>1</sup>, ETHAN KOCH<sup>2</sup> and MOHAMMED MAKKI<sup>3</sup>

<sup>1,2,3</sup>*University of Technology Sydney.*

<sup>1</sup>*kimjerleemaeclimacosa.ricafort@student.uts.edu.au, 0000-0003-2898-2037*

<sup>2</sup>*ethan.r.koch@student.uts.edu.au, 0000-0002-1288-8858*

<sup>3</sup>*mohammed.makki@uts.edu.au, 0000-0002-8338-6134*

**Abstract.** The urban superblock of Kampung Melayu, located in Jakarta, Indonesia, is a typology amalgamated by the environmental and infrastructural challenges caused by Jakarta's urban sprawl. Rapid and unregulated urban growth, fluctuating tropical conditions, rising sea levels and unprecedented environmental stresses have led to a city that is sinking, leaving unregulated low-income settlements, such as Kampung Melayu, most vulnerable. To address these issues, the presented research employs the use of a multi-objective evolutionary algorithm for an in-depth analysis of the various relationships within the urban fabric. The simulations present an alternative urban approach to the design of a flood resilient Kampung; addressing environmental and demographic stresses while maintaining the irregularity that has become ingrained in the history of the urban form.

**Keywords.** Jakarta; Kampung Melayu; Sequential Simulations; Evolutionary Algorithm; Computational Design; Urban Growth; Flood Resilience; SDG 3; SDG 6; SDG 10; SDG 11; SDG 13.

## 1. Introduction

Annual flooding is a prominent urban disaster in many countries in the Asia Pacific region, particularly during the monsoon season. This leads to the contribution of urban implications in aspects including water, sanitation, waste management and destruction of property (Yenny et al., 2017). The effects of annual flooding are amplified in heavily populated and low-income urban centres, in which government funding and aid are not readily available (Alzamil, 2020). One prominent locale is Jakarta, Indonesia; a city that has been attributed to be 'sinking', (Lyons, 2015) causing the government to relocate the country's capital city to another locale, yet for the millions living within the city, the problem still remains.

The causes of flooding in Jakarta ranges from over-extraction of groundwater, deforestation in upstream areas, lack of river dredging, and the conversion of green areas to residential, commercial, and industrial developments. Climate change has also

exacerbated the floods in Jakarta by increasing the frequency of extreme weather conditions, endangering the dominating communities of *Kampungs* situated along the riverbanks (Dovey et al., 2019).

The paper examines one of the most populated urban fabrics in Jakarta that is affected by annual flooding, the *Kampung*. Through the use of a multi-objective evolutionary algorithm (MOEA) with a thorough selection process applied to the generated results, the presented experiment investigates an alternative urban approach to the design of a flood resilient *Kampung*, addressing environmental and demographic stresses while maintaining the irregularity that has become ingrained in the history of its urban form. This allows for a thorough examination of issues under the Sustainable Development Goals (SDGs) published by the United Nations (UN), addressing the provision of clean water and sanitation (SDG 6); production of sustainable cities and communities (SDG 11); the promotion of good health and well-being of the residents (SDG 3), reduced inequalities of living standards within Jakarta (SDG 10), and finally, addressing climate action through the integration of environmental and climatic design goals within the all aspects of the conducted experiment (SDG 13).

## **2. Context and Research**

### **2.1. KAMPUNG MELAYU, JAKARTA**

Located in the heart of central Jakarta, within the neighbourhoods of Kebon Pala and Tanah Rendah; *Kampung Melayu* is an unregulated, disorganised informal settlement home to a densely populated, compact community (Sihombing, 2004).

*Kampung Melayu* was a bustling trading centre during the Dutch colonial period in Indonesia, which lasted from the 16th through the 19th centuries. Established during the 17th century, it housed Malay communities from Malaysia and it thrived due to its location along the Ciliwung River route, which turned out to be the most active trading avenue for people and goods. Residents of *Kampung Melayu* were mainly traders or vendors who owned small businesses (Yenny et al., 2017). The majority of the current population is not of Malaysian heritage, but rather come from other regions of Java Island that have lived in the area for centuries.

### **2.2. FLOODING**

*Kampung Melayu*, which is only 15 meters from the fast-flowing Ciliwung River, is prone to considerable frequent flooding, especially during the rain season. Numerous floods have affected the residents of Jakarta, in 2002, 2007 and 2013, peaking at 3.5 metres in height (Budiyono et al., 2015).

The Ciliwung River is the largest among 13 rivers running through Jakarta and is heavily polluted with heavy metal concentrates, mainly lead and zinc. Studies of viruses, infectious diseases and bacterial indicators were found to be within the contaminated floodwaters, surging especially during the wet seasons between January and February (Mishra et al., 2018).

Despite the circumstances, the residents still choose to continue living next to the river, where it is used for a variety of activities such as washing, defecation and swimming. It was believed that residing in a place with significant health hazards,

limited infrastructure, inconsistent water and electricity supplies, and frequent floods was generally seen as an acceptable and typical part of everyday life (Purba et al., 2018).

In efforts of relocating the residents due to its unfavourable living conditions, the residents were apprehensive about moving or being transferred by the government to other sections of Jakarta with better living standards. Instead, public buildings, religious buildings, schools, and open spaces were used as evacuation zones for temporary shelters during times of flood (Yenny et al., 2017).

### 2.3. OVER-DENSIFICATION

Jakarta continues to grow and is beset by a slew of issues that have gotten worse over time. Growing environmental concerns, not only in Jakarta but also in the surrounding areas, are one of the most pressing issues it faces. According to reports, the carrying capacity of Java Island, where it is located, is already overburdened, due to land and water concerns (Rustiadi et al., 2021).

With an ever-growing population of 10,56 million residents in Jakarta in 2020 (BPS, 2021), the increased demand for housing is one of the consequences in irregular settlements or urban villages such as Kampung Melayu. In 2015, it had the greatest population density, with an estimated 640 people per hectare (Lestari & Sumabrata, 2018).

Houses were found to be overcrowded and have had poor levels of sanitation (Purba et al., 2018). Notwithstanding the circumstances, residents of Kampung Melayu refuse to be relocated as they worry about loss of income and cost implications when it comes to rental. It is understood that the fulfilment of living with their long-term neighbours is prioritised when given the opportunity of relocation despite urban or environmental implications. The preference for immediate solutions located close to their homes is observed in times of natural disasters and urban challenges (Yenny et al., 2017).

## 3. Method

### 3.1. EVOLUTION STRATEGY

The presented research examines the challenges listed above and aims to convert Jakarta's urban typology of the Kampung into an urban superblock which mitigates the current and future flooding and rising sea level challenges, whilst improving urban living conditions for the residents of the Kampung. The multi-objective evolutionary algorithm (MOEA) implemented in the experiment is NSGA-2, developed by Deb et al. (2000). NSGA-2 is the driving algorithm behind the software Wallacei, a free plug-in written for Grasshopper 3D (Makki et al., 2018).

The experiment is presented in four key stages. Stage one examines the site and deconstructs its core urban characteristics; stage two builds the evolutionary matrix for the MOEA, the matrix examines the relationship between the chromosomes (the parameters), the phenotype (the urban form), and the fitness objectives (design goals); stage three analyses the results of the algorithm and applies a thorough multi-step selection process and stage four identifies the selected solution and examines it within

an urban context in relation to the original design goals of the experiment.

Finally, the experiments address the challenge of selection usually associated with the utilisation of MOEAs in design, in which the optimisation of conflicting objectives generates multiple 'optimal solutions', whereas the user usually requires a single solution as a design output. The presented addresses this issue by integrating a thorough selection process that is both data driven, as well as user driven, to identify the best solution from the optimal solution set.

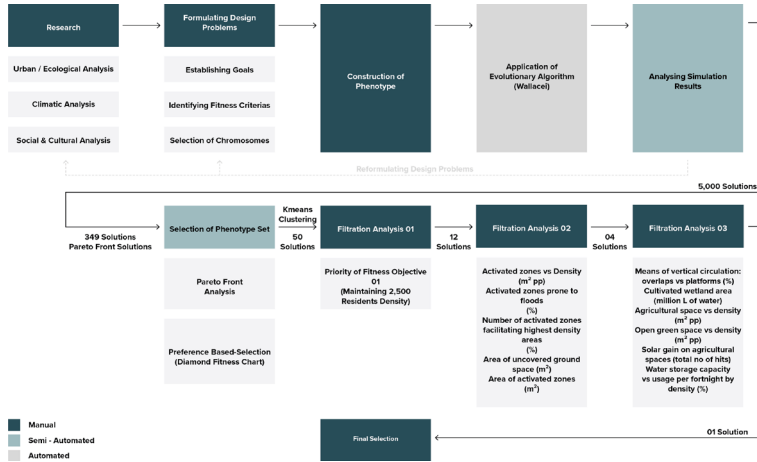


Figure 1. Pseudo Code

## 4. Experiment Setup

### 4.1. PARAMETRIC DEFINITION

The definition utilises the existing site boundary, dividing the site into 30m x 30m platforms. Through a series of trial and error, it was found that this size best accommodates variation in building footprints. The platforms were then moved in the X, Y and Z axis to re-introduce natural light to the ground plane while raising the platforms at varying heights between three to six metres to ensure flood resilience. Points along the existing site boundary were taken as reference and rotated, allowing the platforms to emulate the random orientation of the kampung and surrounding urban context. Platforms that encroached the site’s surrounding context were culled, and those that encroached on the boundary along the river were retained.

A condition unique to the phenotype’s construction was the typology of overlaps. Based on the analysis of the overlapping areas, larger overlaps were converted into voids to allow for more light to access the ground plane while smaller overlaps were converted into water storage units. After which, building footprints were created using the setback of the platform’s current footprint, allowing the primary network to also be defined at three metres in width. The building footprints were further divided into smaller components to replicate the building scale in the surrounding context, creating a secondary network with a 2m width.

The building footprints were then differentiated as viable and non-viable footprints. Non-viable footprints were blocks that were either too small or had odd corners deemed unfit for residential use; these footprints were converted into agricultural spaces on the platforms. Viable footprints were extruded to create buildings ranging from one to four storeys. Thereafter, the buildings that were only single storey high had their roofs converted into an open green space for the resident's recreational use. This summarises the approach taken to address the platform level.

To address the ground plane, the uncovered ground spaces were identified and the areas that had the largest areas of connected green space would be converted into activation zones for residents to gather. The centre points of these activation zones were used as a reference to create varying circumferences, in which the buildings that would fall within these circumference zones would be extruded to the ground plane as means of vertical connectivity. The introduction of cultivating wetlands and green tissues on site have been explored as it would help in flood inundation around nearby areas, and will aid in purifying the river water (Figure 2).

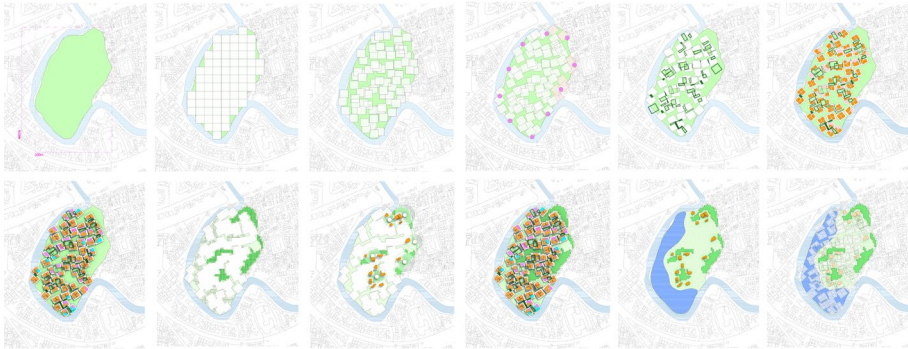


Figure 2. Key Steps of Construction of Phenotype

#### 4.2. FITNESS OBJECTIVES

The evolutionary matrix presented in Figure 3 highlights the relationship between the fitness objectives, chromosomes (explained in the previous section) and the phenotype. The matrix is vital in identifying the impact of each chromosome on each fitness objective, which is critical when debugging the design problem.

Fitness Criteria	Chromosomes				
	Chromosome 01 Pattern of platform	Chromosome 02 Shape of building footprint	Chromosome 03 Height of residential building	Chromosome 04 Area of residential building	Chromosome 05 Number of building connected to ground plane
Fitness Criteria 01 Maximum density of target platform (2000 people/10000m <sup>2</sup> )	X	X	X	X	X
Fitness Criteria 02 Maximize water green agricultural spaces on platform	X	X	X	X	
Fitness Criteria 03 Maximize number of vertical circulation	X	X		X	
Fitness Criteria 04 Maximize agricultural area on ground plane	X	X		X	X

Figure 3. Evolutionary Matrix

### 4.3. ALGORITHMIC SETUP

The algorithm evolved 5000 solutions (generation size 50 and generation count 100) and the simulation runtime was 11 hours and 19 minutes. The simulation was run on a consumer-grade PC, Intel(R) Xeon(R) W-10855M CPU, 2.80GHz with 32.0 GB of RAM.

## 5. Experiment Results & Selection Process

### 5.1. SIMULATION RESULTS

Examining the results produced by the algorithm (Figure 4) indicates that the simulation was successful in improving the mean fitness values for objectives 1, 2 and 3, while objective 4 (maximise ground level activation zones) struggled to optimise throughout the algorithmic run. Analysis of the standard deviation charts indicates that although the mean fitness was improving, the variation of solutions was fluctuating throughout and not converging towards a local or global optima.

Through the analysis of the results, it is clear that the complexity of the design problem necessitates a longer algorithmic run. Improving mean fitness while maintaining variation indicates that a higher generation count (which will also result in a longer simulation runtime) is critical to allow the algorithm to converge towards an optimal solution set. As such, the assessment criteria used to analyse the output of the algorithm, for the purposes of selection, will therefore be critical for a comprehensive understanding of the results, and to identify the impact of the various parameters on the morphology of the urban form.



Figure 4. Simulation Results

### 5.2. SELECTION PROCESS

Analysis of the Pareto Front, along with a series of manual filtration analyses are employed to identify the best solution evolved by the algorithm.

To better understand the amount of variation generated by the algorithm, the pareto front solutions (i.e., the solutions in the population that are not dominated by any other solution), were clustered using K-means clustering with a K-value of 50. In total, the algorithm outputted 349 Pareto Front solutions; the cluster centres of these solutions were selected for further analysis.



The first filtration analysis aims to prioritise maintaining the site's original density, despite the algorithm optimising to maximise it. Among these 50 cluster centres, ten solutions closest to the site's original population density of 2,500 residents were selected, along with two additional solutions that did not meet the original density but had unique fitness values. These 2 additional solutions aim to be experimental, kept as observation points to examine how they perform when compared to the other solutions optimised for density.

The second filtration analysis conducts manual calculations of density and the ground plane. A ranking matrix was used to score the solutions based on a series of primary and secondary requirements when compared with one another. These requirements were the analysis of the density distribution, activation zones and uncovered ground space. From this analysis, the top four solutions were selected based on their highest scores. Figure 5 presents this analysis using a colour scheme to identify the solutions with a poor, average and good ranking.



Figure 5. Ranking Matrix for the Purposes of Selection

The final filtration analysis focuses on the urban attributes of the platforms while calculating the impacts of solar gain and areas of cultivated wetland areas. Using the same ranking matrix as the previous step, it was observed that generation 66, individual 43 had the highest score of compared to the other solutions. This solution however was the 'experimental' solution that was selected in step one, a solution that did not meet the density requirements identified. This was a clear indication that the density requirements significantly affected the solution's performance value for all other criteria. Although solution Gen.66\_Ind.43 is only experimental and thus will not be selected, it demonstrates the value of assessing the solutions generated by the algorithm within the context of the design goals and objectives being pursued. As such, the best ranking solution that met the density threshold was generation 95, individual 24 (Figure 6).



Figure 6. Final Solution - Generation 95, Individual 24

### 5.3. ANALYSIS OF FINAL SOLUTION

Having selected generation 95, individual 24 as the final solution, two focal points within the site were selected as part of the manual design process. This aims to bring a higher level of design thinking and exploration of street typologies and urban form from the perspective of the end-user of the space.

Retaining the importance of communal culture and the spirit of a Kampung, an analysis of colours, materiality and urban characteristics were revisited, in order to implement a sense of place within the newly designed superblock. Integration of new ideas of gathering spaces and communal bridges aims to enhance the Kampung's overall connectivity.

A study of the components within each focal point was conducted to understand the urban form make-up of each area and the importance of ensuring a balance of public facilities and private useable spaces, despite high variation and orientation between the platforms (Figure 7). An analysis of the section also provides an insight of the functionality of the wetlands and the vertical connectivity of the spaces, ensuring ample spatial interaction between the platform datum and the ground plane.

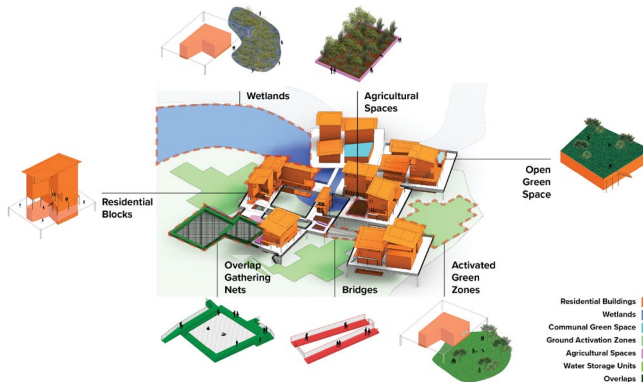


Figure 7. Components of Focal Point

This process of designing the evolved solution outside the scope of the algorithm aids in the critical reflection of the selected phenotype, addressing the above demographic and environmental challenges while comparing it to the pre-existing site condition.

## 6. Conclusion and Discussion

Kampung Melayu in Jakarta was used as a case study in the experiments presented to address demographic and environmental challenges affecting the urban fabric of high density, low-income superblocks impacted by extreme climatic conditions; in this case, flooding. The presented model, although applied within Jakarta, can be adapted to other regions as well. The approach of generating a strong relationship between parameters (genes), design goals (fitness functions) and morphological characteristics (phenotype) can be adapted to incorporate environmental, social, and cultural characteristics of sites located in alternative locations. The presented experiment is not intended to replace the existing Kampung Melayu; rather, it proposes an alternative solution to creating a superblock that incorporates a higher degree of variation throughout the urban tissue, allowing for more resilience to floods through localised parameters as opposed to top-down decisions. It is imperative that there is an attempt to identify the morphological urban characteristics that represent the cultural and social traits of the superblock being investigated. These characteristics must form a core component of the formulation of the design problem, as well as the criteria used to analyse and select solutions from the pareto front. In the selected solution, open space plays a key role in the superblock, however, considering urban patterns in Kampung settlements, built form supersedes open space, and so although the selected solution maintained the existing population density of the Kampung, additional research that addresses population growth within the settlement, and its impact on urban form, is critical.

Finally, a limitation of the computational setup resulted in the evolutionary algorithm unable to optimise and converge for the selected fitness objectives, although one solution is to run a longer simulation, the time needed to do so may not be feasible. As such, an alternative solution is to revise the evolutionary matrix (presented in Figure 3) and reformulate the design problem to revisit the impact the various chromosomes and genes have on the fitness objectives being optimised.

## References

- Alzamil, W. S. (2020). Housing Challenges for Urban Poor: Kampung in Jakarta, Indonesia. *Current Politics and Economics of South, Southeastern, and Central Asia*, 29(1), 81–112.
- Budiyo, Y., Aerts, J., Brinkman, J., Marfai, M. A., & Ward, P. J. (2015). Flood risk assessment for delta mega-cities: A case study of Jakarta. *Natural Hazards (Dordrecht)*, 75(1), 389–413.
- Deb, K., Agrawal, S., Pratap, A., & Meyarivan, T. (2000). A Fast Elitist Non-Dominated Sorting Genetic Algorithm for Multi-Objective Optimization: NSGA-II. *International Conference on Parallel Problem Solving From Nature* (pp.849–858). <https://doi.org/10.1007/s11069-014-1327-9>
- Dovey, K., Cook, B., & Achmadi, A. (2019). Contested riverscapes in Jakarta: Flooding, forced eviction and urban image. *Space and Polity*, 23(3), 265–282. <https://doi.org/10.1080/13562576.2019.1667764>
- Lestari, W. M., & Sumabrata, J. (2018). The influencing factors on place attachment in neighborhood of Kampung Melayu. *IOP Conference Series: Earth and Environmental Science*, 126, 012190. <https://doi.org/10.1088/1755-1315/126/1/012190>
- Lyons, S. (2015). The Jakarta floods of early 2014: Rising risks in one of the World's fastest sinking cities. *State of Environmental Migration*, Liège Université.
- Makki, M., Showkatbakhsh, M., & Song, Y. (2018). Wallacei: An evolutionary and Analytic Engine for Grasshopper 3D. *Wallacei*. <https://www.wallacei.com/>
- Mishra, B. k., Rafiei Emam, A., Masago, Y., Kumar, P., Regmi, R. k., & Fukushi, K. (2018). Assessment of future flood inundations under climate and land use change scenarios in the Ciliwung River Basin, Jakarta. *Journal of Flood Risk Management*, 11(S2), S1105–S1115. <https://doi.org/10.1111/jfr3.12311>
- Purba, F., Hunfeld, J., Fitriana, T., Iskandarsyah, A., Sadarjoen, S., Busschbach, J. J. V., & Passchier, J. (2018). Living in uncertainty due to floods and pollution: The health status and quality of life of people living on an unhealthy riverbank. *BMC Public Health*, 18, 782. <https://doi.org/10.1186/s12889-018-5706-0>
- Rustiadi, E., Pravitasari, A. E., Setiawan, Y., Mulya, S. P., Pribadi, D. O., & Tsutsumida, N. (2021). Impact of continuous Jakarta megacity urban expansion on the formation of the Jakarta-Bandung conurbation over the rice farm regions. *Cities*, 111, 103000. <https://doi.org/10.1016/j.cities.2020.103000>
- Sihombing, A. (2004). The tranformation of Kampungkota: Symbiosys between Kampung and Kota: A case study from Jakarta. In *International Housing Conference in Hong Kong, Housing in tlte 21st Century: Challenges and Commitments* (pp. 2-4).
- Yenny, R., Parnell, M., & Soebiyani, V. (2017). Understanding community-led resilience: The Jakarta floods experience. *Australian Journal of Emergency Management*, 32, 58–66.

# **Digital Heritage**



# EXPLORING THE TOPOLOGICAL SYSTEM OF DOUGONG

HAN-TING LIN<sup>1</sup> and JUNE-HAO HOU<sup>2</sup>

<sup>1,2</sup>*Graduate Institute of Architecture, National Yang Ming Chiao Tung University.*

<sup>1</sup>*hantinglin@arch.nycu.edu.tw, 0000-0002-4745-3520*

<sup>2</sup>*jhou@arch.nycu.edu.tw, 0000-0002-8362-7719*

**Abstract.** The large-span wooden construction project uses a sophisticated tenon joinery system to overcome the limitation on the size of the material. However, making a clear layout and knowledge transfer is an important issue under the complex structure. This research takes “Dougong” as an example to sort out the possible knowledge graph of Dougong. Through the geometric feature classification and the relationship between the joints, we found that the structural relationship of traditional Dougong is like the branch system of the L-system. But it has the characteristic of horizontal connections that make Dougong restrain one another more firmly. Besides a graphical representation of the complex joinery system, it can quickly visualize and adjust the type changes and therefore provide another network related to the building model. Besides computational geometry to traditional wood structure analysis and automation, we also explored two new types of Dougong from a perspective of the traditional wooden structure. So, in this research, we developed automatic digital tools for Dougong and propose new applications of Space Syntax, attempting to break through the existing limitations of Dougong.

**Keywords.** Dougong joint; Knowledge Graph Visualization; Parametric Design; Space Syntax; SDG 4; SDG 9; SDG 11.

## 1. Introduction

The role and importance of wooden structures have transformed and shifted. According to Yingzao Fashi (Li, 1920) in the Song dynasty, traditional buildings used special tenon-joining methods and systems to achieve large-span wooden constructions with smaller wooden materials. Since the industrial revolution, cement and steel structures used in place of wood for large projects. During the mid-to-late 20th century, wooden structures inspired some architectural designs even with cement, such as Arata Isozaki’s design of “Clusters in the air” in 1962. Until recently, in response to environmental sustainability, the importance of natural materials was once again emphasized. Shigeru Ban and Kengo Kuma went beyond traditional wooden structural systems in terms of digital tectonics and scale, such as the Tamedia Office Building. And in prospective research, such as the beaver digital wood structure (Cheng and Hou, 2016), which solve physical wooden system’s information into computer language,

and from digital organization to manufacturing. As stated above, the fabrication of unique joinery relationships fascinated different generations. They used the assembly methods or system to respond to the sustainability issues of the environment, culture, and future. And lead to the topological system we will investigate.

According to White's Concept Sourcebook (1975), architectural design often uses simple diagrams as transferring knowledge which behind them are the beginning of the construction of the system. So, it is necessary to comprehend the complexity of the joinery system through these simple diagrams. In the traditional wooden structure, Dougong is classic tenon joinery for beam-column support systems. It also highlights the craftsmanship in the history of over 2,000 years. And brought out this research takes the Dougong of the Tang Dynasty as an example. During the process, we use the Rhino/Grasshopper to analyze its topological relationships and explore corresponding knowledge graphs such as generating Dougong's classification, topology map, force transmission graph, and new form suggest for breaking through the perception of traditional Dougong.

## 2. Research Background

In the past, the stakeout drawing by artisans had important implications, including clear lofting on site, knowledge transfer between artisans, and passing on to the next generation. And now, we recognize the complex wooden joinery's typology and structure through 3D models or animations. However, whether the sophisticated structure such as Dougong no longer is viewed in the way of classic architecture, but a way of computer understandable topology system. The system not only reduces many calculations or artificial analyses, but also helps users to make the final decision. Currently, there is relatively little research on Dougong's design methods or systems. The following discusses the topology of components and the selection of digital tools through the literature.

### 2.1. FROM TYPOLOGY TO TOPOLOGY OF COMPONENTS

The structure of the Dougong is much more complicated than the general tenon joints, so there are many studies to discuss. In order to understand the geometric form and system topology of the Dougong, the following reviews the literature in three aspects: First is the geometric form of the Dougong. Second is the topology of the relationship between the components. Third, the generation rules of Dougong.

- Dougong's geometric form: Di uses Boolean operations such as the size and inner cut of the Dougong on the historical record of digital modeling (Di et al., 2014). In addition, Ranger explores the geometric proportions in the parametric construction of the Dougong and the Chinese palace (Ranger et al., 2020). The above two methods respectively describe the digital meaning of the Dougong and discuss the authenticity of the digital model.
- Components' topology: Jørgensen said that the IFC structure was just the hierarchical relationship of the data but didn't contain the relationship between the overlapped components. Then, he proposed a schematic diagram of the building components to express the importance of relational topology in the data (Jørgensen,



2011). Furthermore, Li used the transmission of a seismic force of Dougong to draw the concept diagram of the topological properties between the components (Li et al., 2018). Although the above are only conceptual diagrams, they reveal the need for a knowledge graph to describe the relationship between building components.

- Dougong's computer generative design: Wu (2003) applied the concept of Shape Grammar (Stiny and Gips, 1971) to find out the logical rules of Dougong and create a set of grammar to generate it. In addition, some studies tried to apply to the parametric design by the growth pattern of the L-system (Prusinkiewicz et al., 1988). Cai et al. (2020) used the Paifangs' composition system, which transforms a graph and built-in components with L-system grammar derived from architectural rules. Therefore, deconstructing the system of components and applying computer logic to parametric design is a way to inherit the original system and carry out the generative design.

Through the above literature review, we will discuss three aspects of Dougong for integrating into a system for exploring Dougong. The first simplifies the geometric understanding of Dougong. Second, develop the knowledge map of Dougong. Third, expand the possibility of Dougong.

## 2.2. DIGITAL TOOLKITS

We use Rhino/Grasshopper as the parametric platform, which will facilitate the exploration of Dougong geometry. In addition, we use Syntactic to analyze the relationship of components. Finally, use Wasp to explore the possibilities of Dougong components. Besides the existing tools, we have also developed an automated process in Chapters 3 and 4.

### 2.2.1. *Syntactic: A digital tool for spatial graph theory*

According to Space Syntax (Hillier, 2007). Pirouz Nourian developed the Graph Theoretical Methods and a tool: Syntactic in 2016. He reorganized the Space Syntax into algorithms and released them as open source. Syntactic uses Breadth-First Search: BFS to find the relationship value and depth value corresponding to each space. Then, it calculates four different parameters: Integration, Control, Choice, Entropy, and uses visualization to illustrate the mutual space relationships. This tool analyzes Dougong to explore whether the components can correspond to the relevance of the space.

### 2.2.2. *WASP: A digital Tool for Discrete Modeling*

Wasp is an expansion kit proposed by Andrea Rossi research (Rossi and Tessmann, 2018). Its purpose is to aggregate geometric components and apply them to discrete modeling and structural entities that allow assembly and disassembly. However, this tool will be suitable for use because the assembling of the Dougong is just like 3D puzzles. It will break through the imagination of traditional Dougong and provide possible future forms and even topological relationships.

### 3. Geometric Analysis of Dougong

In Rhinoceros3D, objects are constructed and organized mainly by geometric relationships, but lacking the order of composition. To solve this problem, we developed a set of automation processes in this chapter, including sorting, classification, and visualization of Dougong. The purpose is to deal with the arrangement of the Dougong components sequentially rather than structural order and simplify the complicated naming. These processes simplify the comprehension of the Dougong and facilitate subsequent research operations. Then, the following parts discuss the foundation of the Dougong. One is the order of the Dougong assembly. The second is feature classification.

#### 3.1. INTERLOCK SEQUENCE

We advanced the sorting function for the components to solve lacking information about the actual structural order of Dougong, which can automatically define IDs. The following process depicts the sorting of the interlock sequence (see Figure 1). First, we adjust the order of the Dougong structure based on the Z value of the center of mass of each component. We found that Ang (the extension member of the Dougong) would break the stacking order and cause the problem of sorting correctly. Therefore, it is necessary to find out the position of Ang. Next, we use the geometric points of each component to perform linear regression for the centerline. Then, identifying Ang by the Z values of the centerlines' vectors. Finally, we sort and arrange the components to get the correct interlock sequence.

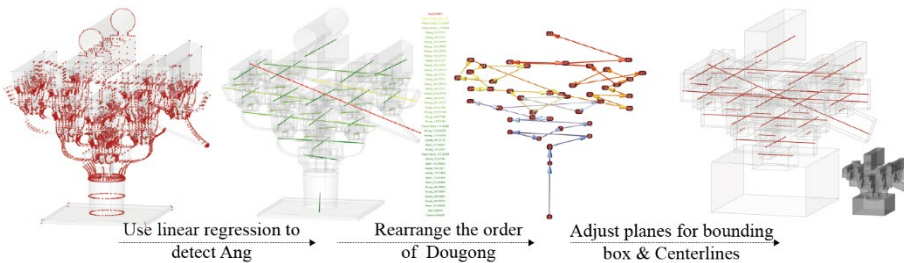


Figure 1. Automatic sorting diagram of Dougong.

#### 3.2. FEATURE CLASSIFICATION

This research also developed the automatic classification function, which can screen the relative characteristics and the numbering of the Dougong components, which will simplify the understanding of the complex structure and multiple naming of the Dougong. We use the Dougong component's unique geometric characteristics, such as proportion, vector, and section area, to reorganize it. Then test on different Dougong (see Figure 2).

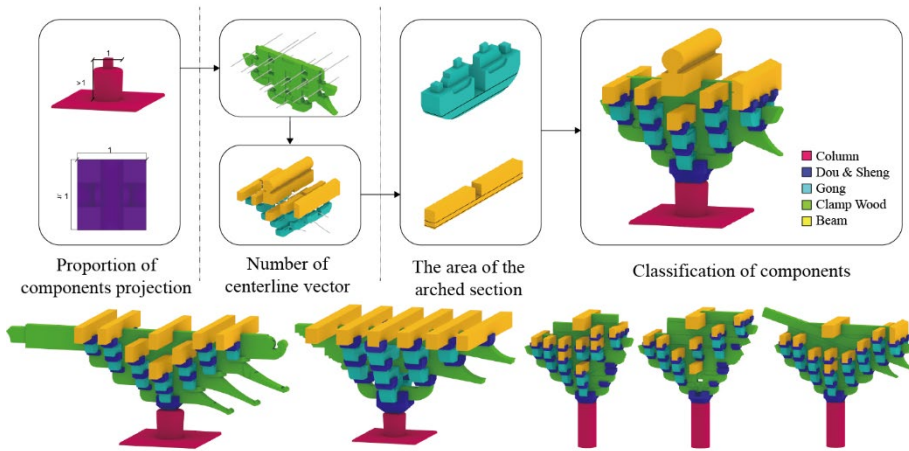


Figure 2. Feature classification of different types of Dougong.

#### 4. Knowledge Graph of Dougong

Continuing from the previous chapter, we developed two tools for automated analysis, including relational topology diagrams and force transmission diagrams, and use Syntactic for inspection and analysis. We explored the knowledge graph of Dougong through three steps: First, put forward three sets of knowledge graphs of the Dougong relationship. Second, use Syntactic for component analysis and visualization. Third, develop a knowledge map of the transmission of the Dougong force. Hence the design of novel wooden structures no longer needs to go through complicated drawings and physical simulation to its trivial details (see Figure 3).

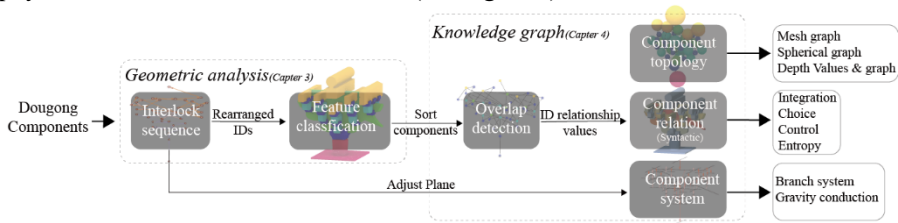


Figure 3. System structure.

#### 4.1. COMPONENT TOPOLOGY: GET RELATIONSHIP VALUE AND DEPTH VALUE THROUGH VOLUME OVERLAPS

To find out the relationship between the components, the automatic overlap detection of them will find out the others. They overlap each other and generate different relationship diagrams based on their ID relationship values. There are three types of graphics based on the relationship between components.

- Mesh topology graph: We connect the centroids to edges to form the graph  $G = (V, E)$  (Nourian, 2016). In addition to visualizing the edges, the process also records the overlapping relationships between the components. Comparing with topological

studies based on L-system, The system is a cross-linked branch system due to the horizontal connections between chains (see Figure 3).

- Spherical topology graph: We use kangaroo collision and extrusion to get the system topology based on the volume of the component. And then convert it into a sphere to understand the relationship between the component's connection (see Figure 4).

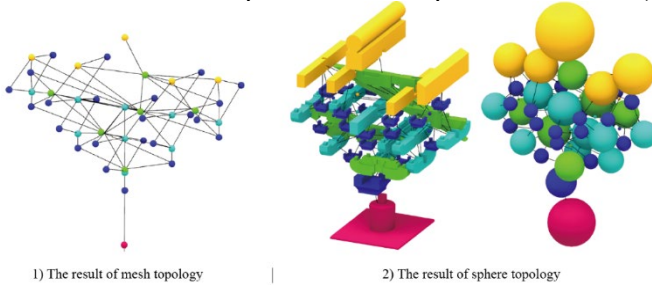


Figure 4. Components' knowledge graph.

- Depth topology graph: According to the ID relationship values above, the BFS algorithm calculates the shortest number of nodes for the depth value corresponding to each component. For example, the column reached through three or five nodes to Arch. So, the gong is at the third depth of the column. And conversely, the column is also at the third depth. According to the relationship, each component would derive a depth map (see Figure 5).

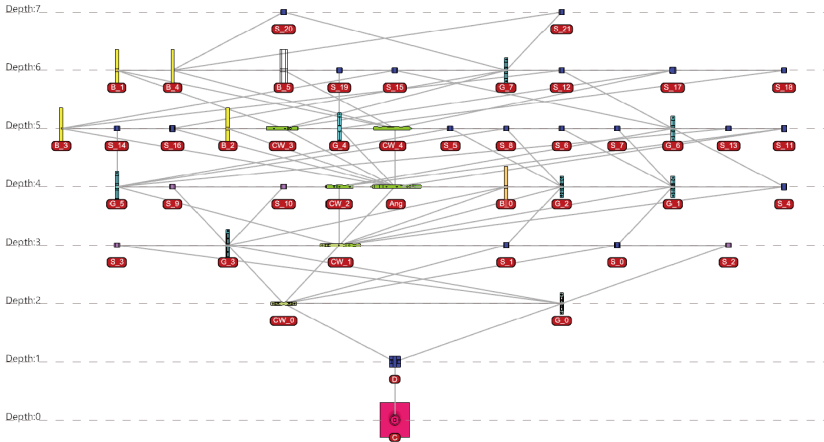


Figure 5. Depth topology graph.

#### 4.2. COMPONENT RELATIONS: THE NUMERICAL MEANINGS IN SYNTACTIC

To respond to whether the four relational values in Chapter 2 can apply to traditional wood structural systems. We used Syntactic for inspection and analysis based on the ID relation value and depth value of Dougong. And then reinterpret the component meaning of this value and visualize it (see Figure 6).

- Integration: The degree of proximity is a measure of the central component. The higher the value, the closer it is to the core position.
- Choice: Indicates the importance of a node, which is the total number of times that a component is located between other components. The higher it is, the more important it is.
- Control: It is the sum of the reciprocal of the depth values between the components. The higher the number, the more controlled and the harder it is to move.
- Entropy: The higher the value, the more difficult it is to connect its components to others.

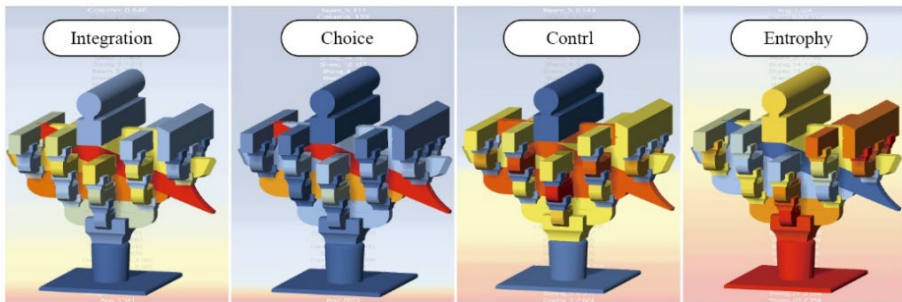


Figure 6. Visualization of Dougong with four numerical values.

Through the integration and the choice, we found Ang is the center of the overall composition and found its slope is significantly higher than other components. In addition, from the visualization of the control value, the middle gong is the maximum value of all. However, the topological diagram in section 4.2 cannot express the particularity of Ang. This phenomenon also leads to the need for a system topology with slope characteristics.

#### 4.3. COMPONENT SYSTEM: BRANCHING AND FORCE TRANSMISSION

In the previous paragraph's conclusion, the three topology graphs can't read out Ang's information. So, we turn the link of Dougong components' centroid points into a centerlines connection; we found it is essentially close to the branch system. Because of the characteristics of the horizontal connection, it's more like a cross-linked branch system and maybe more in line with the topological relationship between the components. Through the relationship between the branch system and gravity extrusion, Dougong components clamp each other to resist the seismic force. The following explores its system and force transmission in three parts (see Figure 7).

- Branch system diagram: Organize into three specific branch states to understand Dougong's system. (1) Vertical branch: bearing capacity. (2) Horizontal branch: the centerlines of the components. (3) Special branches: form a network of fixed vertical and horizontal branches and extend them.
- Force conduction vector diagram: According to the contact points of different branches, the force conduction simulation got by projection.

- Gravity conduction diagram of connecting components: Compare the force state of each component and its associated Dougong components.

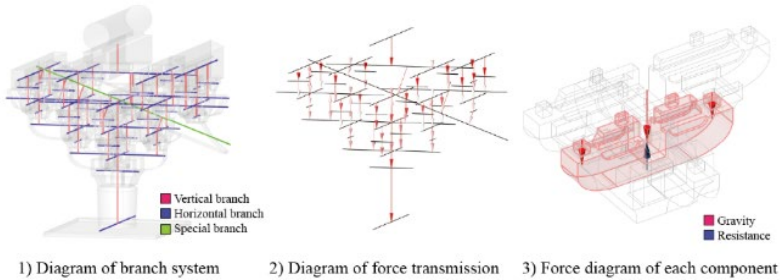


Figure 7. Visualization of Dougong's system diagram.

## 5. Alternative Applications

Besides exploring the traditional system topology of Dougong, this research also attempts to extend the wooden structure to new applications. We discuss with Dous and Gongs, which are the representative components of Dougong. Therefore, we will not discuss clamp wood. The following is the trend of the two types of Dougong. One is the discrete digital Dougong; the other is that we use Dougong in the frame system.

### 5.1. SYSTEM FRAMEWORK OF DOUGONG

Dougong is a node of a wooden structure system that can share the gravity of the roof and beams. So, this section shows the construction method of a large-scale frame system. We develop a digital tool for inputting the frame lines and finding the nodes to produce the wooden structure. This process can improve the flexibility of the angle of the Dougong and reduce its vertical and horizontal restrictions (see Figure 8).

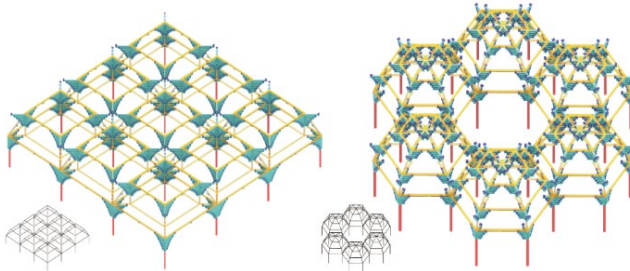


Figure 8. System modeling of Dougong.

### 5.2. DISCRETE MODELING OF DOUGONG

Discrete modeling and generative design are new possibilities for architectural design in using modular components for parametric construction. This section exhibits an assembly method based on the tenon joint relationship. We use the external expansion kit Wasp of Grasshopper to perform this operation and take the component's center point and the vector line for the discrete algorithm modeling. There are three types of attempts below: one is the discrete modeling of an individual component, the other is

the modeling of Dous and Gongs, and the last is the discrete modeling of composite components of Dougong (see Figure 9).

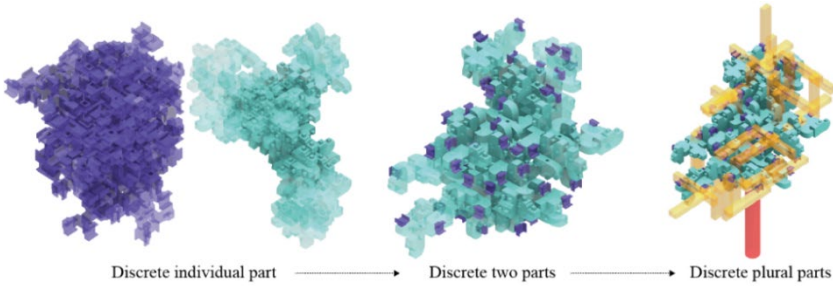


Figure 9. Discrete modeling of Dougong.

### 5.3. EXPERIMENT OF TESTING THE TOPOLOGY SYSTEM

We brought the discrete one into the automatic toolkit mentioned in Chapters 3 and 4. It can form sequence relationships, relationship id value, and depth value. The Syntactic tool can use to calculate the relationship between components. This experiment shows that the above knowledge graph system can apply to the complex assembly of Dougong components (see Figure 10).

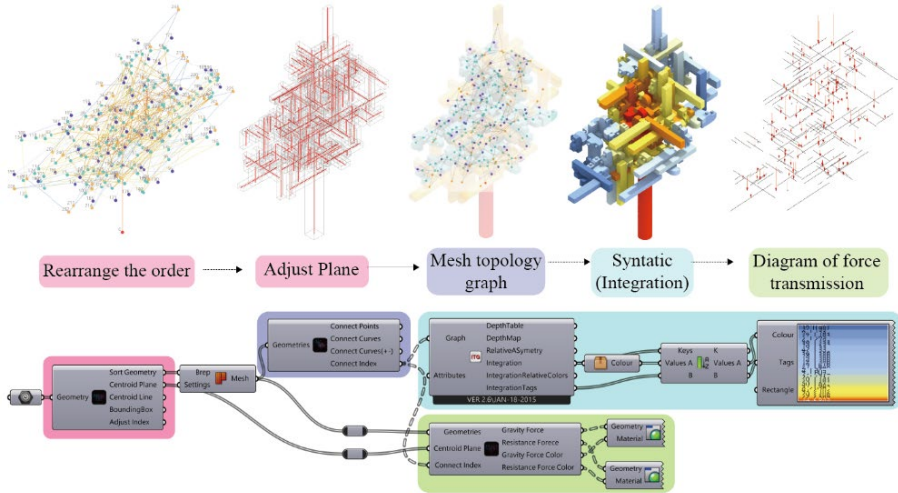


Figure 10. The flow of discrete Dougong's knowledge graph.

## 6. Conclusion and Further Research

In this research, we presented a knowledge graph that humans and computers can easily understand in Dougong through our process; in the analysis process, we developed a set of automatic sorting and classification methods for Dougong through geometric features. Then, we generated the knowledge graph and visualization of the relationship among the Dougong components and extended the application of the Syntactic toolkit to the traditional wooden systems. In addition, we explored two generative examples

and have tested one of them with the process of geometric analysis and knowledge graphing. Through the test, we found the process can use in different components' overlap. We expect the Dougong topological system as a prototype to promote the application of knowledge transfer and potentially provide a digital way of traditional construction. However, this article only studies the relationship between wood structures and has not discussed the details of wood tenoning. Therefore, presenting the knowledge graph of tenon joints is the next step that needs to be promoted.

## 7. Acknowledgments

We thank the technical support and resources provided by the GIA Architectural Informatics Lab, National Yang Ming Chiao Tung University. This research was supported by the Ministry of Science and Technology (MOST 110-2634-F-009-018).

## References

- Cai, Z. Q., Luo, Y. S., Lai, Y. C., Chan, C. S., & Tai, W. K. (2020, February). Interactive Iconized Grammar-Based Pailou Modelling. In *Computer Graphics Forum* (Vol. 39, No. 1, pp. 50-62).
- Cheng, C. L., & Hou, J. H. (2016). Biomimetic Robotic Construction Process-An approach for adapting mass irregular-shaped natural materials. In *Proceedings of 34th eCAADe Conference* (pp. 133-142).
- Di Li, M. K., & Brown, A. (2014). Digital fabrication as a tool for investigating traditional Chinese architecture-A case study of the dou gong. In *Proceedings of the 32nd eCAADe Conference* (pp. 623-632).
- Hillier, B. (2007). *Space is the machine: a configurational theory of architecture*. Space Syntax.
- Jørgensen, K. A. (2011). *Classification of Building Object Types: Misconceptions, challenges and opportunities*.
- Li, Jie. (1920). *Ying zao fa shi*.
- Li, Z., Que, Y., Zhang, X., Teng, Q., Hou, T., Liu, Y., ... & Komatsu, K. (2018). Shaking Table Tests of Dou-gong Brackets on Chinese Traditional Wooden Structure: A Case Study of Tianwang Hall, Luzhi, and Ming Dynasty. *BioResources*, 13(4), 9079-9091.
- Nourian, P. (2016). Configraphics: *Graph Theoretical Methods for Design and Analysis of Spatial Configurations*. A+BE | Architecture and the Built Environment. <https://doi.org/10.7480/abe.2016.14>.
- Prusinkiewicz, P., Lindenmayer, A., & Hanan, J. (1988, June). Development models of herbaceous plants for computer imagery purposes. In *Proceedings of the 15th annual conference on Computer graphics and interactive techniques* (pp. 141-150).
- Raner, Q., Xi, W. A. N. G., Cong, W. U., & Chengjun, B. (2020). Regularized Reconstruction of HBIM for Built Heritage—Case Study with Chinese Ancient Architecture. *Preprints* 2020, 2020010083 (doi: 10.20944/preprints202001.0083.v1).
- Rossi, A., & Tessmann, O. (2018, August). From voxels to parts: hierarchical discrete modeling for design and assembly. In *International Conference on Geometry and Graphics* (pp. 1001-1012). Springer, Cham.
- Stiny, G., & Gips, J. (1971, August). Shape grammars and the generative specification of painting and sculpture. In *IFIP congress (2)* (Vol. 2, No. 3, pp. 125-135).
- White, E. T. (1975). *Concept sourcebook: a vocabulary of architectural forms*. Architectural Media.
- Wu, Q. (2003). *Bracket study: textual, computational, and digital* (Doctoral dissertation, Massachusetts Institute of Technology).



# GEELONG DIGITAL OUTDOOR MUSEUM (GDOM) - PHOTOGRAMMETRY AS THE SURFACE FOR A PORTABLE MUSEUM

DOMENICO MAZZA<sup>1</sup>, TUBA KOCATURK<sup>2</sup> and SOFIJA  
KALJEVIC<sup>3</sup>

<sup>1,2,3</sup>*Deakin University School of Architecture.*

<sup>1</sup>*d.mazza@deakin.edu.au, 0000-0001-7880-7697*

<sup>2</sup>*kocaturk@deakin.edu.au, 0000-0001-7836-5913*

<sup>3</sup>*sofija.kaljevic@deakin.edu.au, 0000-0002-7591-4523*

**Abstract.** This paper presents the development and evaluation of the Geelong Digital Outdoor Museum (GDOM) prototype accessible at <https://gdom.mindlab.cloud>. GDOM is a portable museum—our novel adaptation of the distributed museum model (Stuedahl & Lowe, 2013) which uses mobile devices to present museum collections attached to physical sites. Our prototype defines a way for intangible heritage associated with tangible landscapes to be accessible via personal digital devices using 360° 3D scanned digital replicas of physical landscapes (photogrammetric digital models). Our work aligns with efforts set out in the UN Sustainable Development Goal 11 (SDG 11) to safeguard cultural and natural heritage, by openly disseminating the heritage of physical sites seamlessly through the landscape. Using a research by design methodology we delivered our prototype as a modular web-based platform that leveraged the Matterport digital model platform. We qualitatively evaluated the prototype's usability and future development opportunities with 32 front-end users and 13 potential stakeholders. We received a wide gamut of responses that included: users feeling empowered by the greater accessibility, users finding a welcome common ground with comparable physical experiences, and users and potential stakeholders seeing the potential to re-create physical world experiences with modifications to the digital model along with on-site activation. Our potential stakeholders suggested ways in which GDOM could be integrated into the arts, education, and tourism to widen its utility and applicability. In future we see design potential in breaking out of the static presentation of the digital model and expanding our portable museum experience to work on-site as a complement to the remote experience. However, we recognise the way in which on-site activation integrate into users' typical activities can be tangential (McGookin et al., 2019) and this would necessitate further investigation into how to best integrate the experience on-site.

**Keywords.** Cultural Heritage; Intangible Heritage; Digital Heritage; Web Platform; 3D Scanning; Photogrammetry; Digital model; Portable Museum; Distributed Museum; SDG 11.

## 1. Introduction

Geelong Digital Outdoor Museum (GDOM) is a web-based digital model platform we created for communicating the intangible heritage of significant sites in the city of Geelong, Australia, accessible at <https://gdom.mindlab.cloud>. GDOM presents photogrammetric digital models of locations as the basis for exploring museum quality collections of documentary images, texts and videos of significant cultural sites in Geelong. We recognise that physical sites themselves are the best sites to communicate intangible heritage from and that easily accessible and annotatable photogrammetric digital models can be leveraged to allow on and off-site access. The digital models we describe are spatially accurate digital 3D replicas of physical spaces made using a combination of 360° photographs and spatial information, which can be navigated by users with simple click or touch on personal computers or mobile devices.

GDOM was created in Geelong, a UNESCO City of Design, to address the UN Sustainable Development Goal 11 (SDG 11), specifically to "strengthen efforts to protect and safeguard the world's cultural and natural heritage" (United Nations, 2016). Geelong is a city with a rich industrial past that has changed dramatically over the past decades due to rapid development and the city's urban renewal. GDOM was conceived of to address the local Geelong Design Week 2021 theme 'unpredictable' which by our interpretation asked for a response to the unpredictable effects of events such as the COVID-19 pandemic and the climate emergency. GDOM was in part made to allow local audiences to connect with local places they could not visit due to local pandemic lockdowns.

Our work fits within the realm of distributed museums (Stuedahl & Lowe, 2013)—digital museum collections that are superimposed or attached to real-world sites via personal digital devices such as smartphones and tablet computers. Our contribution is to build upon the distributed museum model with the design development of what we define as a portable museum. Our portable museum engages web viewable digital models to allow the distributed museum experience to work wherever the user has an internet connected smartphone, tablet or computer.

Ahead we provide our methodology and research questions for the design and evaluation of our portable museum prototype. We delve into the rationale for creating a portable museum prototype and provide documentation of GDOM's design. This is followed with a design evaluation of GDOM from front-end users and potential stakeholders, discussion of the results and next steps for the work.

## 2. Methodology and Research Questions

In developing our portable museum prototype we have followed a research by design methodology. Research by Design (RbD), also referred to as design-led research or practice-based research, yields discoveries through the act of making, reflection and evaluation in a design-led process. What makes RbD particularly valuable for our work is the ability to produce knowledge through a design without direct commercial motivation or influence.

We investigate how to produce the portable museum prototype for a variety of common digital devices and seek feedback from front-end users and potential stakeholders to gauge their experience with the prototype and their suggestions. This

research work encompasses the following three research questions:

- How can photogrammetry be used as the surface for presenting the heritage content of significant sites through our portable museum prototype?
- How do front-end users respond to our portable museum prototype and content with respect to comparable experiences?
- What design potentials do potential stakeholders see in the portable museum that would suggest viable ways to keep the work relevant for audiences?

### **3. Design Rationale and Background**

The way we utilise digital models fits within the realm of mobile museology. Mobile museology is defined as a movement in the museum field towards digital mobility and mobilisation (Baggesen, 2018). The specific area we draw upon is using digital platforms as their own museum collections which can integrate effortlessly into everyday life, known as 'distributed museums' (Stuedahl & Lowe, 2013) or 'museum outside walls' (Arvanitis, 2005). We can trace the progeny of these platforms to Malraux's 'museum without walls' (Malraux, 1967) which sought to encourage spreading museum collections outside of formal institutions. These digital platforms seek to use landscapes as museum walls for digital museum collections such as Museum of London's Streetmuseum (Museum of London, 2010) and University of Oslo's Akerselva Digitalt (Sem et al., 2012) through standard applications, or augmented reality applications to superimpose historical images on approximate locations in real-time such as (In)tangible Heritage through eXtended reality (XR) (Kocaturk et al., 2020).

If we look at the distributed museum examples discussed, we find a gap which can be filled directly with the use of accessible digital models. Distributed museum collections can only be accessed on-site without any equivalent accessible experience grounded on that site. At present, audiences can view collections on a map through a guided tour experience, such as in Voice of Norway's mobile app for Akerselva and many other Norwegian localities (Voice of Norway, 2020). However, this approach removes a sense of being on-site by forgoing the use of the 3D landscape itself as a surface for holding collections. Accessible digital models can provide users an always on hand replica of physical sites to learn and build their understanding from regardless of where they are. We define this incremental enhancement on the distributed museum model as the portable museum.

### **4. Prototype Design**

We built our portable museum as a modular web-platform. The modular web-platform acted as a framework for holding and customising digital platforms required to meet our broad aims for digital models that were widely accessible and required relatively less time to implement. Through this approach we could leverage the Matterport digital model platform for the more difficult tasks of capturing, hosting and navigating digital models online while we built an overarching interface to suit our design aims as closely as possible.

#### 4.1. DIGITAL MODEL PLATFORM



Figure 1: 3D mesh of GDOM Eastern Beach (left) and navigation of the corresponding 360° images using the crawling cursor (right)

Matterport is a web-based 360° panoramic image viewer which creates the semblance of free movement with the use of a 3D mesh to mark the site's surfaces. Matterport aligns an invisible 3D map (navigation mesh) with the perceptible surfaces of a site in 360° panoramic images to present a freely navigable digital model, shown in Figure 1. Users can move their cursor across surfaces in the images (a crawling cursor) or tap on surfaces to be taken to the next nearest image.

A Matterport mobile application serves to join 360° images with depth information from a range of supported cameras, which are uploaded, processed and hosted by the company's servers. We used the Leica BLK360 combined LiDAR and 360° camera to merge accurate laser depth information with images in outdoor settings.

Matterport provided a flexible application programming interface (API) for interfacing with Matterport's hosted captures. The API allowed us to program features to improve the user's experience for this context.

#### 4.2. GDOM

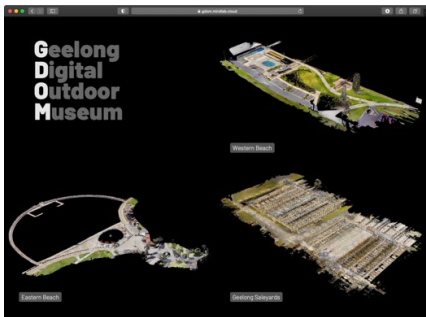


Figure 2: GDOM home page

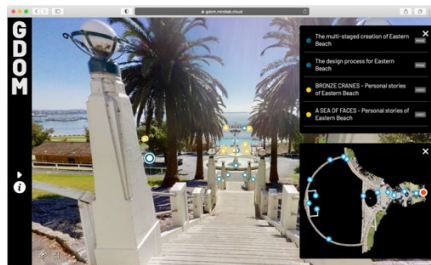


Figure 3: Western Beach site page with Mattertags (center), Minicat (top right) and Minimap (bottom left)

The design of GDOM comprises of a home page which presents all captured spaces, see Figure 2, and a site page which presents a captured site, see Figure 3. The content layout on site pages hinge off the capabilities of Mattertags. Mattertags are coloured

pin-like markers which present a tooltip with image, text or video content when hovered, tapped, or clicked on.

Within the site page we have modified Matterport's default functionality for greater ease of use with a Minimap and Minicat. The Minimap (miniature map) presents a plan view of the site with markers in key locations to allow users to teleport within the space with a marker to show where the user is in the space. The Minicat (miniature catalogue) is a catalogue of Mattertags and their content in a list, to allow quick traversal of content and teleportation to the tag location.

## 5. Evaluation

For our evaluation of GDOM we used two instruments: an online user survey and a focus group session with potential stakeholders. We sought to understand the performance of the design based on different needs and perspectives.

### 5.1. USER SURVEY

The purpose of our user evaluation was to capture the usability of and satisfaction with GDOM's design and content across captured sites. We sought to build a complete picture of the user experience. We used five-level Likert scales to ask participants to rate the quality of design features we added on top of Matterport and the different kinds of content on offer. We also asked users to explain in writing: how the experience compared to a physical museum, their experience with our added design features (the minimap and catalogue), and any additional feedback.

#### 5.1.1. *Details on participants*

Our 32 participants belonged to the following age groups: 28% aged 25–34, 28% aged 45–54, 16% aged 65–74, 13% aged 18–24, 6% aged 35–44, 6% aged 55–64 and 3% aged 75–84. They were English speaking locals (18) and overseas (14) users who were very proficient in using computers with them identifying as 53% expert skill level, 44% intermediate skill level and 3% with minimal knowledge. No participants experienced a platform like GDOM, however participants reported engaging in digital heritage through digital map-based data visualisations, search engines, newspapers, digital heritage, library, gallery and museum databases and standalone 360° image galleries. Participants used the following devices to view GDOM: 52% laptop computer, 26% desktop computer, 11% tablet and 11% smartphone.

#### 5.1.2. *Comparison to physical experience*

We asked participants to rate and explain how GDOM compares to a physical museum experience. Drawing upon participants' physical museum experiences acted as our portable museum baseline. Participants rated their feeling that there was a match as 56% satisfied, 19% moderately satisfied, 16% poorly satisfied and 9% very satisfied. An analysis of written responses shows us a spectrum of responses from preferring the physical to preferring the virtual, see Table 1.

Table 1: Coded responses of GDOM compared to a physical museum

<b>Preferred physical</b>	<ul style="list-style-type: none"> <li>• Touch and smell were missing</li> <li>• There were no real objects, tour guides or samples to take</li> <li>• The atmosphere of the places could be better captured by showing more human activity</li> <li>• It was a good alternative if physical presence were not possible</li> <li>• It would be more immersive or personal to be on-site</li> <li>• Being at a physical museum space encourages/focuses the viewer to concentrate at one thing at a time</li> </ul>
<b>Comparable / neutral</b>	<ul style="list-style-type: none"> <li>• One can move around naturally and at their own pace</li> <li>• It was like being on-site and incredibly realistic/surreal</li> <li>• It was a welcome return to the physical and outdoors where it had been restricted multiple times during the COVID-19 pandemic</li> <li>• Information was well laid out, plentiful and similar to didactic cards shown at museums</li> <li>• GDOM could be an attraction at a physical museum</li> <li>• There is a common tangible element, similar to when comparing a physical book with an eBook</li> </ul>
<b>Preferred virtual</b>	<ul style="list-style-type: none"> <li>• From a user with mobility issues as having the ability to be "in charge"</li> <li>• Navigation was less tiring / can quickly travel from place to place at any time</li> <li>• All information was provided upfront via the interface and easy to view via the artefacts (images, newspaper clippings, interviews)</li> <li>• There were useful shortcuts (in the form of the minimap and catalogue)</li> <li>• As "great for those of us who can't afford to travel, but still want to learn"</li> </ul>

For participants that found the work easier to view and navigate, the predominant highlight was the ability to self-direct exploration. We also found the juxtaposition of past and present in the virtual space to be described by users as more "coherent", "immersive", "nice" and "something you would not be able to get from a normal museum". One user's outlying feedback suggested information shown in Western Beach to be a bit overwhelming—this is something we addressed in our newer sites.

### 5.1.3. Design feature and content evaluation

Ease of navigation rated 63% satisfied, 28% very satisfied and 9% moderately satisfied. The minimap was very well received, with most participants suggesting it was helpful and made it easier to navigate. The catalogue was also well received. However several users suggested they did not notice it or did not know it was interactive. For those that did find it was interactive, it worked as intended as a useful alternative to: navigating the minimap, opening tags or virtually walking. Some users preferred virtually walking. One user suggested that removing the catalogue could make the experience more immersive by avoiding the "temptation of just scrolling through the information". In Table 2 we summarise technical issues encountered with the Matterport platform and future work suggestions.

Table 2: Matterport technical issues / limitations and GDOM future work suggestions from users

<b>Matterport technical issues</b>	<ul style="list-style-type: none"> <li>• Five users suggested Matterport capture loading performance issues, some users suggested the causation as related to their own hardware or network connection.</li> <li>• Two users cited the lurching movement when transitioning from one 360° image to the next can be unsettling.</li> <li>• One user cited some content was visually concealed by floating artefacts in the Matterport capture (an issue pertaining to image capture and stitching)</li> </ul>
<b>Suggested features</b>	<ul style="list-style-type: none"> <li>• Replace site tags with virtual objects or signage placed in situ</li> <li>• Two users wanted some form of narration to accompany the visit</li> <li>• Animate the sites with human activity</li> <li>• Incentivise visiting physical sites by hiding vouchers for local businesses and parking in GDOM</li> <li>• Add an avatar or virtual guide in period costume</li> <li>• Mark the significance of points highlighted on the minimap</li> <li>• Provide an overhead photographic view of the Geelong area to enter the sites from</li> <li>• Have GDOM as an attraction at physical spaces</li> <li>• Provide a video walkthrough with commentary of the areas as an alternative to the platform</li> </ul>

Written responses on favourite aspects of the design were predominately focused on the ease of navigation and movement, variety of content on offer and quality of the design overall. Some users suggested it felt like being there. One user suggested they would hope to see similar work for their own locality.

In asking about the quality of content overall, and specifically for images, text and video, we found participants rated 44–48% very satisfied, 35–44% satisfied, 16–11% moderately satisfied and 4–3% rating poor or very poor satisfaction.

Feedback provided regarding content mostly suggested adding more of it (specifically: sites, ambient sounds, archival films, drawings/plans and newspaper clippings) or design features to better mimic the effect of being on-site. Written responses on favourite aspects of the content were focused on variety and having access to historical images, drawings and newspaper articles.

Overall, 94% of users signified they would use a site like GDOM again and 91% would like to see more websites like GDOM.

## 5.2. FOCUS GROUP

For our stakeholder focus-group session we engaged a cross-section of professionals that would have an interest in contributing to the continued development, distribution, and content of GDOM. 13 participants were involved, with members from local government arts & culture planning and marketing, museum curators, custodians of heritage collections and academics working with relevant digital technologies and heritage.

Focus-group responses can be broadly categorised into the following categories: ways to enhance content, feedback, ideas for design features, promotional ideas and new ideas for displaying information spatially. Secondary categories concern: user-

interface improvements, content sources, suggestions & strategy, new ways to engage audiences and potential stakeholder partnerships. Table 4 illustrates the questions we asked along with answers synthesised answers from a coded analysis of all responses.

Table 4: questions provided to the focus group with answers synthesised from response analysis

Question	Synthesised answer
Q1: Can you think of any additional functionalities we could add to improve the heritage experience?	In asking about potential additional functions for GDOM we discovered that the group was interested in seeing extra guidance provided to users, in the form of: accessibility options, virtual guides and data filtering and using location-based technologies to engage users on-site
Q2: How to bring this web-platform to the attention of the public? How to increase visibility, encourage and incentivise its use?	To increase the visibility of GDOM the group collectively pointed to building layers of partnerships with local business, artists and heritage groups so the platform can work as both an outlet for them while promoting it through their own following customers, viewers, fans, etc
Q3: Can you see your organisation play a role in connecting this platform with relevant users? How?	In asking about potential collaborations from the stakeholder group, no concrete answers were supplied however ideas hinted towards engaging education, tourism and local heritage groups.
Q4.1: Imagine that you are given access to use this platform: 1) which physical location would you add (in Geelong or elsewhere)	In asking what the group would do with GDOM if they could add locations: we found the group wanted to add locations such as Barwon River, Kardinia Park, Wadawurrung (local First Nations people) meeting places, Geelong Cultural Precinct, Geelong Port railway and underwater reefs in Corio Bay—these locations all have hidden histories. We also received feedback to tie content choices with popular search terms for greater audience reach.
Q4.2: Imagine that you are given access to use this platform: 2) what kind of information / representations would you use?	In asking what additional modes of presenting information the group would like to see in GDOM, we found answers commonly referred to showing timelapses of content and spaces. It was also suggested to use a local Wadawurrung significant locations map, using virtual reality (VR) and capturing artefacts in 3D
Q5: Can you think of other uses and benefits of attaching “digital information” to “physical location”? Who else would benefit from it? (e.g. from among general public, businesses, governments)	For additional uses of digital information on physical locations the group's answers revolved around leveraging hidden spatial information (e.g. effects of tree canopy cover and hidden underground infrastructure), displaying online reviews, encouraging users to venture further than their current location, allowing choice of storytellers and connecting indigenous language to physical spaces.

## 6. Discussion and Future Work

Overall, the response to the work was positive from end-users and potential stakeholders. Users with mobility issues or who were travel restricted benefited from the ability to navigate the history of a place with great ease. We also saw desire from stakeholders to visualise site captures change across time, to dive into indigenous history, present physical artefacts and reveal hidden infrastructure and locations. A consistent request across users and stakeholders revolved around further drawing upon



the affordances of physical spaces (such as familiar physical cues, tour guides and content displayed in-situ), and to bring experiences on-site or in physical displays. We interpret this as a desire to see the portable museum ultimately bring people closer to the spaces represented.

While on-site activation is a productive way forward it's important to recognise that accessing heritage via a mobile application on-site is a tangential experience and rarely the main reason for being in a place (McGookin et al., 2019). The suggestions from our potential stakeholders to build partnerships for promotion and to engage in education and tourism applications are very useful to consider.

## 7. Conclusion

Through the execution of our portable museum prototype we created a modular web-based platform to leverage the Matterport digital model platform. This approach avoided reinventing the wheel to achieve our design objective. Our end-users revealed that the remote experience was advantageous or comparable to a physical experience. From our end-users and potential stakeholders, we discovered a desire to return the portable museum to qualities of a physical experience through: extra guidance, content layout more akin to physical experiences, spatial and temporal insights beyond those presentable by a static digital model, and the implementation of on-site activations. Our potential stakeholders helped us understand ways the work could be linked to artistic, educational and tourism activities to keep the work visible and relevant.

## Acknowledgements

This project has been supported by the City of Greater Geelong through the Arts & Culture Arts Industry Commissions. The project has been developed by MInD (Mediated Intelligence in Design) Research Lab at Deakin University, and in collaboration with Reckless Eye Productions led by artist and filmmaker Malcolm McKinnon, who contributed towards the content design of the Eastern Beach and Geelong Saleyards sites. The project team would also like to thank PhD Researcher Anahita Sal Moslehian for her contributions to improve the flow of content shown on the Western Beach site and assistance in laying out content in all our captured sites.

## References

- Arvanitis, K. (2005). *Museums outside walls: mobile phones and the museum in the everyday* (pp. 251–255).
- Baggesen, R. H. (2018). Mobile media, mobility and mobilisation in the current museum field. In *The Routledge Handbook of Museums, Media and Communication* (pp. 115–127). Routledge. doi:10.4324/9781315560168-10
- Kocaturk, T., Wang, R., Kaljevic, S., Wang, C., Argent, J., & Darkspede Technology. (2020, April). *Experiencing (In) tangible Heritage through eXtended reality (XR) – case study of Western Beach Boat Yard Reserve, Geelong Waterfront*. Retrieved July 29, 2021, from <https://mindlab.cloud/experiencing-intangible-heritage-through-extended-reality-xr-case-study-of-western-beach-boat-yard-reserve-geelong-waterfront/>
- Malraux, A. (1967). Museum without walls. On *London : Secker & Warburg*. (André Malraux, S. Gilbert, & F. Price, Eds.). London : Secker & Warburg.

- McGookin, D., Tahiroğlu, K., Vaittinen, T., Kytö, M., Monastero, B., & Vasquez, J. C. (2019). Investigating tangential access for location-based digital cultural heritage applications. *International Journal of Human-Computer Studies*, 122, 196–210. doi:10.1016/j.ijhcs.2018.09.009
- Museum of London. (2010, May). *Museum of London apps*. Retrieved July 29, 2021, from <https://www.museumoflondon.org.uk/discover/museum-london-apps>
- Sem, I., Isdal, E., Smørdal, O., Stuedahl, D., Andreassen, D., & museum, O. (2012, April 14). *Akerselva digitalt (completed) - Department of Education University of Oslo*. Retrieved July 29, 2021, from <https://www.uv.uio.no/iped/english/research/projects/contact/trango/Akerselva%20Digitalt/>
- Stuedahl, D., & Lowe, S. (2013). *Design experiments with social media and museum content in the context of the distributed museum*. Retrieved from <https://archive.nordes.org/index.php/n13/article/view/304>
- United Nations. (2016). *Envision2030 Goal 11: Sustainable Cities and Communities*. Retrieved August 2, 2021, from <https://www.un.org/development/desa/disabilities/envision2030-goal11.html>
- Voice of Norway. (2020, July 10). *Akerselva - et unikt utemuseum - nå med guide via din mobil*. Retrieved August 6, 2021, from <https://voiceofnorway.no/akerselva-et-unikt-utemuseum-na-med-guide-via-din-mobil/>

## REVISITING SHOEI YOH

*Developing a workflow for a browser-based 3D model environment to create an immersive digital archive*

NICOLE GARDNER<sup>1</sup>, M. HANK HAEUSLER<sup>2</sup>, DANIEL YU<sup>3</sup>,  
JACK BARTON<sup>4</sup>, KATE DUNN<sup>5</sup> and TRACY HUANG<sup>6</sup>

<sup>1,2,3,5</sup> UNSW / Computational Design <sup>4</sup>UNSW GRID <sup>6</sup>UNSW

<sup>1</sup>*n.gardner@unsw.edu.au, 0000-0001-6126-6716*

<sup>2</sup>*m.haeusler@unsw.edu.au, 0000-0002-8405-0819*

<sup>3</sup>*d.yu@unsw.edu.au, 0000-0002-7788-548X*

<sup>4</sup>*j.barton@unsw.edu.au, 0000-0002-4483-5179*

<sup>5</sup>*k.dunn@unsw.edu.au, 0000-0002-2753-9204*

<sup>6</sup>*t.huang@unsw.edu.au, 0000-0002-4843-572X*

**Abstract.** The digitisation of architecturally significant buildings and sites creates opportunities to innovate methods of analysis, interpretation, representation, and audience engagement. To illustrate this potential, but also examine the attendant challenges, this paper outlines a research project that has digitised archival assets and living buildings designed by the Japanese architect Shoei Yoh to create an immersive 3D Spatial Archive. It focuses particularly on the creation of a browser-based 3D environment using WebGL technology that connects to and displays a repository of digitised archival assets. This includes the use of 3D scan data of Yoh's Naiju Community Centre project to accurately model the 3D immersive environment and a Grasshopper / Rhino into the glTF. File format (graphics library Transmission Format) workflow to render Naiju's complex geometry and detailed outdoor scenery. The paper demonstrates how using the .glTF File, which is an open format specifically for transmitting processed and pre-calculated 3D models, can improve the processing efficiency of web-browser based 3D environments. Improving the stability and processing speed of 3D browser-based environments is significant to enhancing how audiences can connect with and experience culturally significant sites remotely. The digital recreation and repurposing of Naiju (which is currently unoccupied and in a state of disrepair) as an immersive archival exhibition space operates to simultaneously protect the real building from over visitation, but also raise awareness of its cultural significance to support preservation efforts. In so doing, the paper makes a further contribution to the developing field of digital cultural heritage.

**Keywords.** Digital Cultural Heritage; Browser-based Modelling; glTF File, Architectural Visualisation; Shoei Yoh; SDG 9; SDG 11.

## 1. Introduction

Digital cultural heritage is a field that seeks to creatively disrupt and transform heritage practices by engaging a range of digital technologies and techniques. An ecology of digital technologies offer new and sustainable ways to enhance the efficiency, accuracy, and scope of cultural heritage documentation (Yang & Greenop, 2020).

Within this field, much focus is given to the digitisation of architecturally significant buildings and sites through 3D spatial data capture (3D scanning). 3D laser scanning has generated opportunities to develop new methods to document and analyse culturally significant buildings and sites (Balzani, Maietti, & Köhl, 2017; Dong, Zhang, & Zhu, 2020; Uplekar Krusche, 2018). Equally, 3D scan data is useful for the virtual recreation of cultural sites (Aiello & Bolognesi, 2020; Zlot et al., 2014) and to guide physical restoration processes (Siu, 2021). The accuracy and detail of 3D scan data is also useful to processes of visual representation and the virtual (re)creation of significant buildings and sites as virtual 3D environments. Virtual recreations of lost and living buildings and sites equally facilitates new opportunities for analysis and interpretation but also new ways for global and intergenerational audiences to engage with and experience remote buildings and sites. The case example discussed in this paper involves the use of 3D scan data in a workflow to virtually recreate Japanese architect Shoen Yoh's pivotal Najju Community Centre project in Fukuoka, Japan. This forms part of a larger and ongoing research project that has been digitising assets from Yoh's architectural office that were deposited with the Kyushu University, Fukuoka, Japan in 2019.

It has also involved 3D scanning Yoh's 'living' buildings. To both preserve and make these archival and new assets more widely accessible, the project team have created the online Shoen Yoh Archive that hosts both a traditional, searchable repository of digitised assets, but also a novel 3D 'spatial archive' which displays these multi-media assets in a more immersive and narrative-driven way. The development of the spatial archive as a browser-based 3D environment and its workflow that uses WebGL technology to address issues of stability and processing efficiency forms the focus of this paper. Accordingly, the following sections of the paper introduce the larger research project and its aims and detail the workflow components used to develop the spatial archive that include Grasshopper / Rhino into the glTF. File format (graphics library Transmission Format) to render Najju's complex geometry and detailed outdoor scenery. The paper concludes by reflecting on the significance of 3D browser-based environments for enhancing how global and intergenerational audiences can connect with and experience archival materials as well as culturally significant sites remotely.

## 2. Research Project Background

The Japanese architect Shoen Yoh is an internationally recognised figure of late 20th century architecture and a pioneer of digital methods in architectural design (Lynn et al., 2013). In 2019 Yoh deposited his entire architectural office archive with the Kyushu University. Since then, a research team from the Kyushu University and University of New South Wales, Sydney, Australia have collaborated to digitise these archival assets, and 3D scan selected living buildings designed by Yoh. A core aim of the research has

been to explore how to move beyond traditional searchable online repositories of archival material by creating a more compelling, narrative-based, contextualised, and immersive archival experience. To explore this the team proposed the idea of a spatialised archive as an environment to showcase the digitised assets of the Shoei Yoh Archive in a more relational way. The team chose to 3D scan and virtually recreate the Naiju Community Centre, which was completed in 1994 in the town of Chikuho, Fukuoka, Japan, but is now unoccupied and in a state of disrepair. In the context of the recent history of digital architecture, the Naiju community centre is considered significant as Yoh combined the use of locally grown bamboo, hand weaving construction techniques, and advanced computer analysis to realise a complex geometric form in bamboo and concrete. But this project was also chosen as the spatial archive environment as it connects to the larger story of Yoh's contribution to the modernisation of timber architecture in Japan through his engagement with digital technology, and namely early structural analysis software. Consequently, within the Naiju virtual spatial archive, archival assets, as well as new parametric models and 3D scan animations related to five other key projects designed and completed by Yoh between 1979 to 1994 are displayed. As such, the spatial archive is not simply a record or virtual recreation of Naiju but the vehicle through which its history and significance is interpreted and communicated. Moreover, as the spatial archive virtually reimagines and repurposes Naiju as an archival exhibition space, it combines the tradition of heritage preservation with a kind of digital futuring that aims to catalyse new thinking about the possibilities of its restoration and reuse.

### **3. Workflow development of a browser-based 3D virtual spatial archive**

Representing architecturally significant sites in immersive and engaging ways has been aided by advances in tools such as 3D scanning along with falling costs in hardware, software, and data storage. But the dense data-capture of methods such as 3D scanning and point cloud models has equally bred a range of data management challenges. For example, practices of digital preservation that create digital records of existing sites can also be seen as fragile and precarious. As Greenop and Landorf argue, "[c]onservation in the digital era may not always involve preservation of the physical fabric itself, but *preservation* of a new, digital version of a heritage place" (Greenop & Landorf, 2017, p. 47). Equally, while the precision and scale of data that can now be readily accessed and used to virtually recreate highly detailed architectural heritage sites can also result in heavy models that can compromise audience experience and engagement through poor stability and low processing speeds. Developing a 3D browser-based environment workflow that addresses processing efficiency has been significant to realising the longer-term aims of this project which include extending the audience reach of the Shoei Yoh Archive and enhancing how people can experience culturally significant sites remotely.

#### **3.1. BROWSER-BASED 3D MODELLING**

To create the Naiju spatial archive the research team sought to leverage the possibilities of browser-based 3D modelling. The development of this work builds on prior research the team developed on web-browser-based 3D modelling and representation (Leung

et.al, 2019). This includes pioneering work in 2018 on the use of JSON (Java Script Object Notation) and GeoJSON (3D extension of JSON) that is now the foundation of online modelling tools such as spacemaker.ai; giraffe.build; and/or testfit.io.

JSON (2021) is a lightweight data-interchange format. It is easy for humans to read and write. It is easy for machines to parse and generate. It is based on a subset of the JavaScript Programming Language Standard ECMA-262 3rd Edition - December 1999. JSON is a text format that is completely language independent but uses conventions that are familiar to programmers of the C-family of languages, including C, C++, C#, Java, JavaScript, Perl, Python, and many others. These properties make JSON an ideal data-interchange language. Ibid. continues by explaining that JSON is built on two structures:

- A collection of name/value pairs. In various languages, this is realized as an object, record, struct, dictionary, hash table, keyed list, or associative array.
- An ordered list of values. In most languages, this is realized as an array, vector, list, or sequence.

As these are universal data structures virtually all modern programming languages support them in one form or another. Consequently, it makes sense, according to Ibid. that a data format that is interchangeable with programming languages also be based on these structures.

GeoJSON on the other hand is a format for encoding a variety of geographic data structures (GeoJSON, 2021). In 2015, the Internet Engineering Task Force (IETF), in conjunction with the original specification authors, formed a GeoJSON workgroup to standardize GeoJSON. The Request for Comments (RFC) 7946 was published in August 2016 and consequently GeoJSON is the new standard specification of the GeoJSON format, replacing the 2008 GeoJSON specification. GeoJSON supports the following geometry types: 'Point', 'LineString', 'Polygon', 'MultiPoint', 'MultiLineString', and 'MultiPolygon'. Geometric objects with additional properties are 'Feature' objects. Sets of features are contained by 'FeatureCollection' objects. (Ibid.)

While JSON is now commonly used for transmitting data in web applications the use of \*.obj or \*.fbx formats offer only a slow experience in particular when transmit a 3D scene data efficiently over the internet for viewing in a remote application. Thus, this research used Shoei Yoh's Naiju project with its complex geometry and its detailed outdoor scenery as an example to develop a workflow from Grasshopper / Rhino into the glTF. File format (graphics library Transmission Format).

According to Wikipedia's entry on glTF (2021) it is a standard file format for three-dimensional scenes and models. A glTF file uses one of two possible extensions, \*.gltf (JSON/ASCII) or \*.glb (binary). A \*.glTF file may be self-contained or may reference external binary and texture resources, while a \*.glb file is entirely self-contained. It is an open standard developed and maintained by the Khronos Group, an open, non-profit, member driven consortium of 170 organisations developing, publishing and maintaining royalty-free interoperability standards for 3D graphics, virtual reality, augmented reality, parallel computation, vision acceleration and machine learning. A \*.glTF file supports 3D model geometry, appearance, scene graph hierarchy, and animation. It is intended to be a streamlined, interoperable format for the delivery of

3D assets, while minimizing file size and runtime processing by apps. As such, its creators, lov.gov (2021) have described it as the "*JPEG of 3D.*"

Thus, as a \*.glTF File format is an open format specifically for transmitting processed and pre-calculated 3D models it has the potential to become the solution to the problem of the proliferation of proprietary, closed, and restrictive cloud 3D viewing formats - an interest that the development of the Shoei Yoh Archive aimed to investigate. This notion is based on the observation that a \*.glTF is more suitable for viewing models than editing them - a \*.glTF has no parametric features, or blocks, or arrays, or any other complex geometry forms. Given this it was deemed a more suitable approach for digital cultural heritage projects.

### 3.2. SET UP AND METHODOLOGY FOR DEVELOPMENT

In agreement with our web developer partner Special Design Research Studio, Sydney we proposed a scrum / agile methodology with a series of four sprints for the software development and action research methodology for collaborative researching the topic with our web developer partner. We started Sprint 1 in early May 2021, but at this time we had little processed digital data of Naiju to work with. Our team has been on site at Naiju in Japan in February 2020 and had taken photos as well as preliminary 3D scanning of the Naiju exterior with a LeicaBLK360 3D scanner. At this point the Naiju plans, sections, elevations, or detail drawings had not been digitised.

**Sprint 1 - 3D modelling a first trial.** As such, the development work began with a trial using a simplified 3D model generated in Unity, shown in Figure 1. This allowed us to explore the opportunities, constraints, and limits of bringing a 3D model in a web environment.



*Figure 1. 3D model first iteration in Unity (left), Naiju Community Centre photographed during site visit in February 2020 (right).*

The trial 3D model was examined according to key categories including topology, UV maps, pivot points and other potential artefacts along with some well-crafted attributes. The following errors were identified and required.

- **Pivot Points.** The pivot point was set to the far left to the following 3D coordinates: X 21765.302 / Y 16559.705 / Z 329.404. It was agreed to readjust the pivot point to be in the centre of the model and have the pivot placed in the accurate base position.
- **Topology.** Given the structure and shape of the 3D model, the shape does not capture the exact silhouette of the Naiju model from the external view. There are

multiple sections of the model that present some odd polygons and triangles. These are topology issue that will cause some artefacts and shadow artefacts in Unity. Furthermore, the model had numerous stretched polygons on the neck of the top section of the model and multiple jagged vertices and edges jutting into the inner creases of the model. It was agreed to clean these out to void any blur/unclean shadow lines. The results from these were also show internal as polygons were pushing inside the exhibition space, showing bad artefacts.

- **UV Map.** The overall UV (The letters "U" and "V" denote the axes of the 2D texture because "X", "Y", and "Z" are already used to denote the axes of the 3D object in model space, while "W" (in addition to XYZ) is used in calculating quaternion rotations, a common operation in computer graphics) maps have been created neatly for the most part, however there were signs of overlapping UV maps. It is important to keep these flat so that light baking and texture work is accurate and there is less chance of stretching. In addition to this there was a lot of empty space, which could be used better by scaling the UV maps/or by including more topological data.
- **Materials IDs.** It was agreed to consider assigning UV IDs for model interior and exterior and to scale each UV island to use up all the space within the UV layout.
- **Open Geometry.** The edges of the building were not enclosed and left open, resulting in the model showing paper thin like borders which can cause shadowing issues as well as limiting the use of isolating the 3D object as it is not a solid 3D object.
- **Texture work.** Texture work on the exterior environment showed a promising start. However, it was noticed to keep in mind of the scale of the environment in relation to the tile able settings on each material. For instance, if the structure is meant to be large in scale, the camera will be all the way close to the model which would show the blurry large textures if not scaled accurately. Due to the poor UV mapping, the textures on the internal dome were also very pixelated and poorly presented.

At the completion of Sprint 1, it was agreed to minimize and eliminate the risk of technical issues mindful of the following additional findings:

*Firstly*, do not use any third-party assets or applications which are primarily not built for WebGL, such as Gaia, this most likely has been the cause of the file size exceeding 4.5GB.

*Secondly*, we will export all 3D models as \*.FBX and not \*.OBJ as this will help in keeping files more readable within Unity.

*Thirdly*, to keep all texture images to the power of 2 e.g., 512x512, 256 x256 in width and height.

*Fourth and lastly*, to name all objects in your scene appropriately, as these will be called out via code in the future for optimization needs.

**Sprint 2 - Unity Package Check.** With the feedback of Sprint 1 and having gained access to digitised plans, sections, elevations as well as processed 3D scan data of Naiju in mid-June 2021 we were able to produce base model geometry correctly using GH / Rhino 3D. Yet problems occurred in the polycount of the model. Here the model, excluding landscape/environment, had a total of 12,621,597 polygons and 6,311,534



vertices. This created problems with rendering on a web browser. It was agreed to optimize the model within 40K to 50K polygons maximum and keep the shape intact.

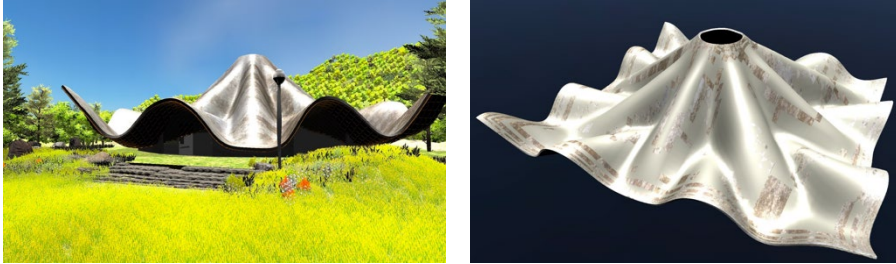


Figure 2. 3D model second iteration in Unity (left), broken model due to OBJ export (right).

This was based on research and experience that demonstrated that anything larger can cause significant lag and overly complex UV maps. If a lower polygon count could not be achieved, the alternate consideration was to create various LODs (Level of Details) in the model as this would allow the camera to render based on distance. Following corrections were identified and required in sprint 2.

- **FBX export.** Model was not exported as native 3D software \*.FBX but \*.OBJ. Due to this, the UV maps were inaccessible and showcasing inaccurate UV maps/seamlines which seem very broken and inefficient as there were multiple broken seam lines visible. While this was an oversight of an action item (see above) in Sprint 1 it demonstrated the importance to use the \*.FBX format.
- **Pivot Point.** Again, a problem raised in Sprint 1 but overlooked as now several models were placed into the environment.
- **Naming convention.** Most models were named correctly, however, not all objects were named correctly. There are a many that still are labelled as Object\_#. Causing issues.
- **Interior Mesh Gaps.** Model placement for the interior space had gaps, especially the floors. As we intended that the user will be able to walk and look all around, it was important to address these issues as one will easily see the gaps between each level.
- **Backed Lightmaps.** At this Sprint 2 the question of lightmaps were discussed for the first time. It was agreed that once the topology/segments were reduced, it would be beneficial to add lights in the 3D software of choice and bake the lighting information onto the UV maps of the model and export the data out as texture maps. This then can be then implemented into Unity and reduce the usage of Directional/spotlights in Unity.

To summarise Sprint 2. In moving forward, once aspects of the main structure were fixed, it was proposed that all vegetation and painted grass on the terrain be removed, and the terrain was converted into a mesh (See Figure 2). This was based on knowledge that Unity, and most game engines, typically become overly heavy when terrains and 'SpeedTrees' are used. SpeedTree is a 3D vegetation modelling and middleware used

to create trees and plants for films and television. For a web browser application these are not suitable and were replaced by static panoramic spheres with images of trees to encapsulate the vegetation in the following.

**Sprint 3 - Clean up model and set up of environment with lights.** In this third sprint from June to end of July 2021 the model required several aesthetic changes to meet the expectations of the team towards a near photorealistic experience on the browser. The model contained multiple vegetation assets which needed to be removed and light sources in the scene were deleted. And, while the topology on the model was already reduced to approximately 100k polygons, further reduction was required. In this sprint 3 following corrections were identified and required:



*Figure 3. Third iteration now on the web browser exterior view, and first interior views.*

- **Vegetation.** We deleted 70% of vegetation for optimization purposes.
- **Illumination.** Lights baked and positioned in areas required.
- **Size of model.** Topology was reduced to 21k polygons, texture work and materials created to compliment lights.

While the results started to look promising it was agreed to work on further refinements in terms of optimization weight, lighting/saturation, etc.

**Sprint 4 - Interior and adding interior assets such a photo, video, sounds, and 3D model.** In the fourth and final sprint, the static media content was integrated within the model. The photos and videos were split into folders with each path for the folders aligned to a display panel inside the gallery space. This integration was developed to allow the referencing link to each file within the folders to remain the same, which meant that the content was able to be changed provided the naming convention for the new files matched the referencing links in the model. For the deployment of the 3D parametric models on the plinths inside the Naiju exhibition space, each 3D model needed to be baked into the Unity model. This process of integration needed the 3D

models to be further optimised by reducing the polygon count. This optimisation process was achieved by only utilising one face of each truss structure and keeping all faces as quad polygons. With the media referencing and 3D model optimisation completed, the model needed to be transferred from its local server to an online storage and display solution. Amazon Web Services (AWS) was used to facilitate this process as its services allowed for a fast transfer speed at a global scale. The Unity model itself also needed to be integrated onto a webpage for users to access the model on a browser. With this objective, the entire website was developed with HTML, CSS, and JS. All the webpages and media content were deployed onto a public AWS S3 bucket. Using AWS Cloudfront, a distribution service was setup for the S3 bucket which allowed the entire website and Unity model to be cached on local distribution servers in each country, allowing users to access the data from their devices at the fastest speed available. Finally, using AWS Route53, the name servers for the hosting URL were pointed towards the Cloudfront distribution servers which completed the online hosting and integration process. With all sprints completed the project went live at 8pm on 2nd December 2021 and can now be accessed at <https://shoeiyoh.com/>.

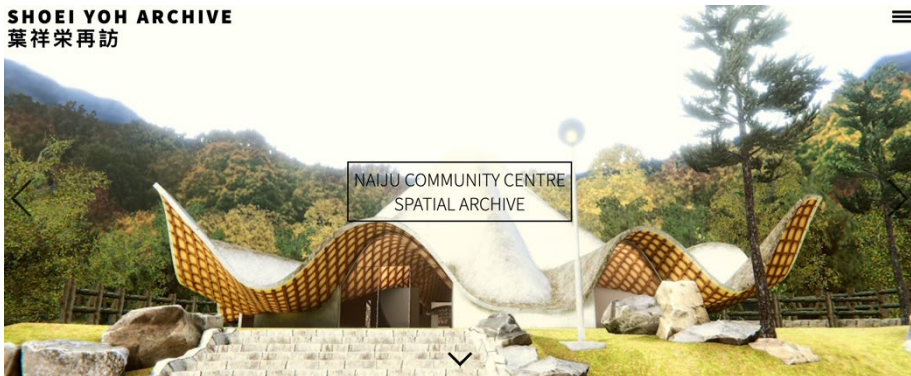


Figure 4. Naiju Community Centre on the Shohei Yoh Archive.

#### 4. Conclusion

The digitisation of architecturally significant buildings and sites creates opportunities to innovate methods of analysis, interpretation, and representation, but also extend audience reach and engagement. The falling cost of 3D laser scanning equipment and data storage, along with the use of technologies such as drones and sensors make the collection of data about living buildings and sites far more accessible and continual. But the high precision and granular nature of this data, while useful for a range of analytical and interpretive purposes, brings forth attendant issues of data handling and management. In the context of virtual recreations designed for online public outreach, the richness of data can work against the goal of creating accessible, high quality and immersive modes of remote audience engagement. To explore these issues, this paper has drawn on the case example of virtually recreating Japanese architect Shoji Yoh's pivotal Naiju Community Centre as a novel browser-based 3D spatial archive. It has shown that a browser-based 3D model environment and data management methods adopted are a suitable to the virtual recreation of significant buildings and sites for the

purposes of public engagement in digital cultural heritage projects.

### Acknowledgements

The authors would like to thank Special Research Design Studio for partnering on the project and our collaboration at Kyushu university, Fukuoka, here in particular our collaborators Masaaki Iwamoto, Tomo Inoue and Momoeda Yu. Funding by the Australian Government Department of Foreign Affairs and Trade (DFAT) Australia-Japan Foundation Grant Round 2020-2021, project number and title: AKF2020045 – Revisiting Shoen Yoh: Digital Preservation and Architectural Archiving.

### References

- Aiello, D., & Bolognesi, C. (2020). Reliving history: the digital reconstruction of the convent of Santa Maria delle Grazie in Milan. *Virtual Archaeology Review*, 11, 106. doi:10.4995/var.2020.13706
- Balzani, M., Maietti, F., & Köhl, B. (2017). Point cloud analysis for conservation and enhancement of modernist architecture. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W3, 71-77. doi:10.5194/isprs-archives-XLII-2-W3-71-2017
- Dong, Q., Zhang, Q., & Zhu, L. (2020). 3D scanning, modeling, and printing of Chinese classical garden rockeries: Zhanyuan's South Rockery. *Heritage Science*, 8(1), 61. doi:10.1186/s40494-020-00405-z
- Greenop, K., & Landorf, C. (2017). Grave-to-cradle: A paradigm shift for heritage conservation and interpretation in the era of 3D laser scanning. *Historic Environment*. doi:10.3316/ielapa.958198990712870
- JSON (2021, December 6) *Introducing JSON*. Retrieved December 7, 2021, from <https://www.json.org/json-en.html>.
- GeoJSON (2021, December 6) *GeoJSON*. Retrieved December 7, 2021, from <https://geojson.org/>.
- Leung, E., Butler, A., Asher, R., Gardner, N., Haeusler, H.M. (2019) Redback BIM - Developing a Browser-based Modeling Application Software Taxonomy, in: *Proceedings of 24th CAADRIA Conference*, pp. 775-784.
- Loc.gov (2021, December 6) *gITF (GL Transmission Format) Family*. Retrieved December 7, 2021, from <https://www.loc.gov/preservation/digital/formats/fdd/fdd000498.shtml>.
- Lynn, G., Eisenman, P., Gehry, F. O., Hoberman, C., Yoh, S., & Zardini, M. (2013). *Archaeology of the digital* : Peter Eisenman, Frank Gehry, Chuck Hoberman, Shoen Yoh. Montréal, Québec, Canadian Centre for Architecture, Sternberg Press.
- Siu, S. (2021). APPLICATION OF 3D SCANNING TECHNOLOGY IN RESTORATION OF HERITAGE SITE DAMAGED BY NATURAL DISASTER. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVI-M-1-2021, 685-691. doi:10.5194/isprs-archives-XLVI-M-1-2021-685-2021
- Uplekar Krusche, K. (2018). 3D Documentation and Visualization of the Forum Romanum: 7th International Conference, EuroMed 2018, Nicosia, Cyprus, October 29–November 3, 2018, Proceedings, Part I. In (pp. 281-300).
- Yang, C., & Greenop, K. (2020). Introduction: applying a landscape perspective to digital cultural heritage. *Built Heritage*, 4(1). doi:10.1186/s43238-020-00002-w
- Zlot, R., Bosse, M., Greenop, K., Jarzab, Z., Jukes, E., & Roberts, J. (2014). Efficiently capturing large, complex cultural heritage sites with a handheld mobile 3D laser mappingsystem. *Journal of Cultural Heritage*, 15(6), 670-678. doi:<https://doi.org/10.1016/j.culher.2013.11.009>

# LEARNING FROM HALE

## *An Educational Augmented Reality for an Indigenous Hawaiian Architecture*

RICHARD ROBINSON<sup>1</sup> and HYOUNG-JUNE PARK<sup>2</sup>

<sup>1,2</sup>*School of Architecture, University of Hawaii at Manoa*

<sup>1</sup> *rrrobin@hawaii.edu* <sup>2</sup> *hjpark@hawaii.edu*

**Abstract.** An educational Augmented Reality (AR) application with Head Mount Display (HMD) is developed for the revitalization of the Hales. The proposed application allows a user to have a dynamic learning experience of the Hale by 1) full immersion into an extended reality, 2) enabling the hands-on construction & assembly process with real-time feedback, and 3) visualizing context-specific information and concepts. Through this intact experience, tacit knowledge embedded in the Hawaiian Hale design is delivered. In this paper, the implementation of the proposed application is explained, and the usage of the application is also demonstrated.

**Keywords.** Augmented Reality; Tacit Knowledge; Cultural Heritage; Hale; SDG 4.

### 1. Introduction

Hale, a traditional Hawaiian house, has diminished in and lost its ground from the landscape in Hawaii. The revitalization of traditional hale is an opportunity for community gathering, and the restored structure itself has become a place for teaching (Abermathy et al, 1978; Bryan, 1950). Tacit knowledge embedded in Hale design serves to revive and restore the traditional building practices and instil pride that comes from living traditional Hawaiian values (Edwards 2019; Apple, 1971; Buck, 1957). However, such tacit knowledge has been an elusive subject due to its difficulty to be articulated, recorded, and communicated (Dampney et al, 2002). Furthermore, the oral tradition for transferring the tacit knowledge in Hawaii makes more difficult to convert it to explicit knowledge (Nonaka and Takeuchi, 1995) for the future generation of Hawaii.

Augmented Reality (AR) serves to create a reality that is supplemental to the physical environment (Caudell and Mizell, 1992). AR provides new possibilities for innovative education, and they have been increasingly recognized by educational researchers (Hashimoto and Park, 2020). By adding an enhanced layer of computer-generated information to the real-world environment, AR allows a user to 1) interact with two- and three-dimensional synthetic objects in real-time, 2) visualize context-specific complex spatial relationships and abstract concepts, and 3) experience phenomena with a full-scale immersion (Hashimoto and Park, 2020; Arvanitis et al.,

2007). In the area of Cultural Heritage (CH) education, the implementation of AR is already acknowledged to improve a student's learning experience and motivation (Gonzales Vargas et al., 2020). AR allows users to experience how people used to live within Cultural Heritage (CH) rather than structures or objects of the past (Voinea et al., 2018). With the help of AR, many breakthroughs of preservation, documentation, and exploration of CH have been made by different professions such as archaeologists, researchers, or museum curators (Bekele et al, 2018; Cozzani et al, 2019).

In this paper, an educational Augmented Reality (AR) application with Head Mount Display (HMD) for learning Hale, an indigenous Hawaiian architecture, is proposed for reviving Hale as a place for community gathering and teaching. Including Hale Wa'a, four different types based upon the specific function and purpose for its residents have been studied for the development of the application. Rhinoceros 3D and its visual programming language Grasshopper 3D are employed with Fologram plugin for AR environment. This proposed application consists of 1) initiation of AR + assembly procedures, and 2) Hale in AR: geometric model, text tag, and diagrams. The assembly process of Hale Wa'a is demonstrated within the proposed application.

## 2. Hale

Between 124 and 1120 AD, the Hawaiian Islands were first settled by Polynesians, and the Hawaiian civilization was born (Buck, 1957). For the next 500 years, this civilization was isolated from outside contact, and from this isolation, Hawaiian culture was created. Due to the warm climate of the Hawaiian Islands, the shelters created by the Hawaiian people were well adapted to the tropics and provided storage and/or protection against rough weather. Hawaiian Hales are either single houses or complex houses for chiefs and nobles of the island. The Hale is made from grass and is a wooden framework consisting of a ridgepole, rafters, and purlins. Hawaiian Hales typically had a small door and no windows. The building practices of Hawaiian Hale's also incorporate tacit knowledge such as building materials, techniques, and environmental design (Edwards 2019).

The Hawaiian Hale adapted to change after the islands made first contact with missionaries, by rearranging and adopting a new belief system that affected how the hale was built and used. The changes introduced concepts of domestic privacy in contrast with Hawaiian household ideas of "freedom of access" (Aikau and Gonzales, 2019). Similarly, when the Hawaiian language was banned in the late 1800s the culture behind the language was almost lost due to the influence of colonization (Nakata, 2017). The language was revitalized over a hundred years later in the late 1900s and preserved by the native people so it can be passed down and taught to the next generations (Warschauer and Kuamo'yo, 1997). Augmented Reality is a tool that can be used as a means to preserve the culture behind Hales so the information can be passed on to later generations, similar to the Hawaiian language. By preserving tacit knowledge of Hale through AR, institutions like schools and cultural museums can teach the cultural significance behind the Hale of the Hawaiian settlers.

Traditional Hawaiian Hale can be categorized by the ROH (the Revision Ordinance of Honolulu) into four different types of classic Hale styles. These unique styles include A-Frame (Hale Wa'a), open wall (Halawai), lean-to (Ku'ai), or fully enclosed (Noa). In addition, each variety of Hale are used for specific functions including eating,

sleeping, assembling, retailing, and storage (Kilolani and Middlesworth, 1992). The following Hale Wa'a was used for eating, storage, retailing and assembly.

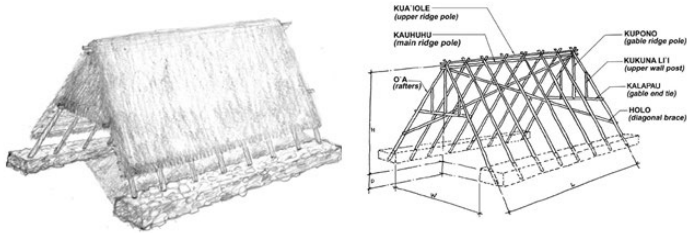


Figure 1. Diagram of Hale Wa'a and Framework (City and County of Honolulu, 2021)

All the four different types of the Hale have been studied for 1) creating the historical and cultural contents, 2) developing geometric models, and 3) translating the assembly process of the Hale into a sequence of design actions with the models as shown in Figure 6.

### 3. Augmented Reality (AR) with Head Mount Display (HMD)

An extended reality (Kaplan et al, 2020) for learning the Hale through its assembly process is implemented with 1) Rhinoceros3D, 2) Grasshopper3D, and 3) Fologram.

#### 3.1. INITIATION OF AR + ASSMBLY PROCEDURES

Rhinoceros3D is used for the geometric modelling of the Hale. Its geometric data are translated into Grasshopper3D for the synchronization of the geometric model to the layer of augmented reality through Fologram.

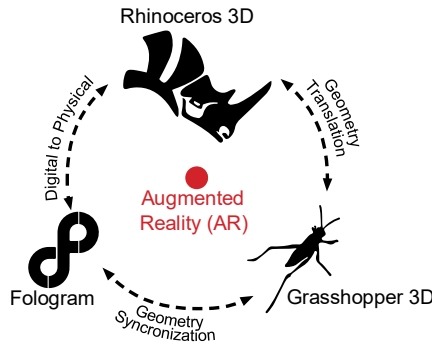


Figure 2. AR framework with Rhinoceros 3D, Grasshopper 3D, and Fologram

In setting up the application, the geometric models of Hale necessary for its construction are developed in Rhinoceros 3d. The geometric models reflect what the physical structure should closely resemble in AR. The first part of the script is used to display the wireframe component of the digital model. Used in unison with Fologram, the wireframe of the model is displayed to the user first indicating the size and scale of the model prior to assembly.

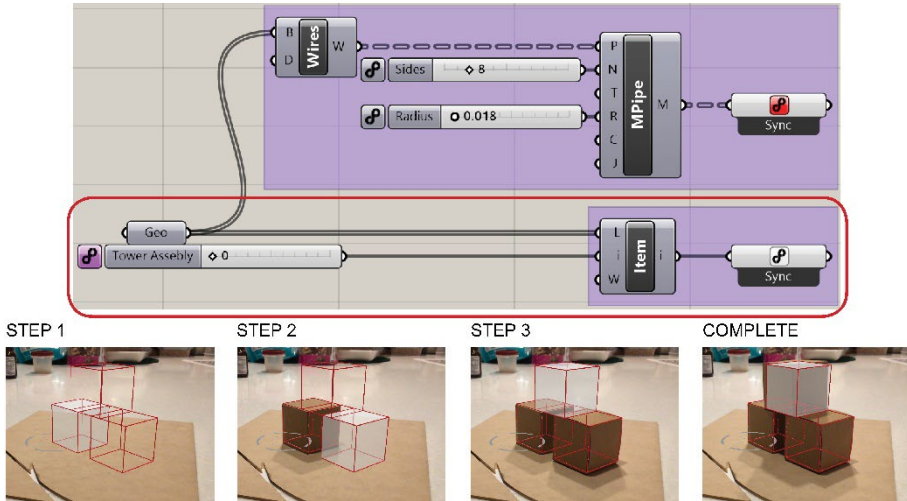


Figure 3. Wireframes of geometric models and Assembly Steps in AR

The second part of the script involves a “parameter” script that controls the assembly steps as shown in figure 5. The script of the digital model in wireframe and its parameter controller are connected to a “list index” script. The “parameter” script then allows the user to cycle through all the connected geometries in the “list index”. By utilizing the “parameter” script, the portion of the model will be highlighted for notifying the targeted object and its placement before moving to the next step in the assembly procedures as shown in Figure 3. The two-way arrow between Fologram and Grasshopper 3D specifically affects the instructional parameters that get ported from grasshopper to the Fologram app, so if the user goes from step 1 to step 2 in Fologram, Grasshopper updates the digital model in Rhinoceros to reflect the assembly step that the user is working on. The user would use AR as a guiding system for the construction of a Hale. Also, the proposed AR allow the user to deviate from the instructions for further exploration.

### 3.2. HALE IN AR: GEOMETRIC MODEL, TEXT TAG, AND DIAGRAM

To create the assembly process in AR, a digital collection of all the different Hale styles and variations needs to be established. By having each model, the assembly can be broken down into pieces and the Grasshopper3D script can be utilized to create the assembly steps. The collection allows a user to cycle through each holographic model prior to engage in its assembly process. The construction of Hale Wa'a is demonstrated with AR in this paper. The Fologram add-on “text tag” becomes movable texts including historic and cultural contents of the Hale. The text tag follows the user’s visual focus through the assembly build as its guidance as shown in Figure 4. In addition to the geometric modelling of the Hale and the text tag, various skill sets embedded in Hale design including lashing, thatching, stacking, and securing are illustrated with any diagrams and images through the panels in AR environment. The diagrams and images had to be saved as images files such as JPEG and added to the



Grasshopper script as textures. The diagram of tie lashings to the Hale is shown in Figure 5.

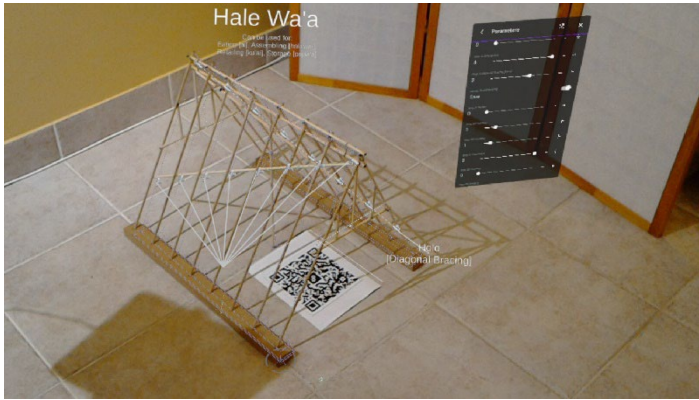


Figure 4. Text tag within AR

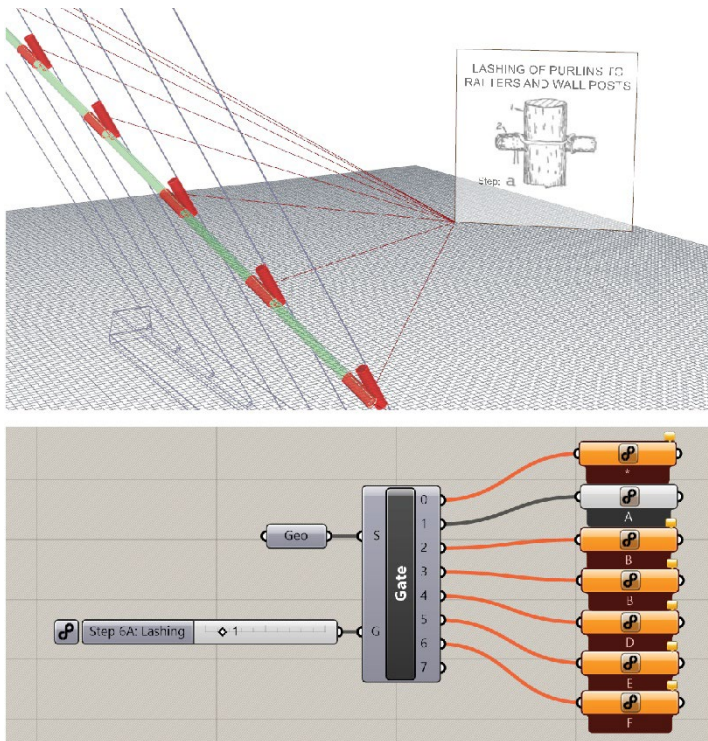


Figure 5. Overlaid Diagrams for Assembly Instructions

After implementing the visualization of the text and diagrams within the AR environment, toggle options are developed for hiding and showing the layers of four different information: master text, titles, parts, and diagrams. This layering system with the toggle options controls the four types of the visualized information according to a user's need.

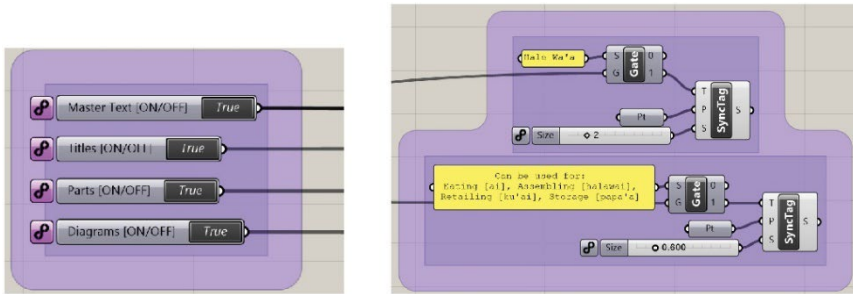


Figure 6. Controlling the information layers with toggle options

From compiling each part of the digital model in Rhinoceros3D and finalizing the Grasshopper3D script with the Fologram tags and diagrams, the AR-based Hale assembly is ready for proceeding with Microsoft HoloLens, a head mount display.

#### 4. Assembly of Hale Wa'a

The Hale Wa'a, shown in figure 1, involved seven steps to the assembly from start to finish. As additional steps to the assembly, sub-steps were added to include visual instructions of the lashing diagrams. the seven steps are 1) Pa Pohaku (Foundation Walls), 2) O'a (Rafters), 3) Pou Hana (Ridge Post) & Kauhuhu (Ridge Pole), 4) Gable Walls, 5) Holo (Diagonal Bracing), 6) 'Aho Pueo (Purlins), and 7) Ako (thatching). Each step is initiated by the parameter changes in the list index of all the assembly procedures as explained in Section 3.1 Initiation of AR + Assembly Procedures.

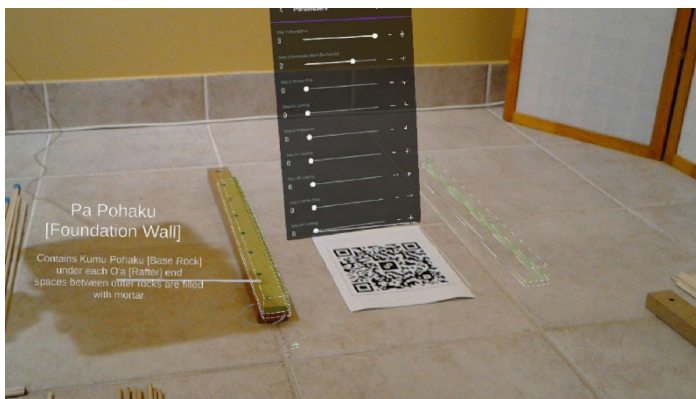


Figure 7. Placement of Pa Pohaku

#### 4.1. PA POHAKU

This step shares QR code location as the origin through AR to find where the pa pohaku would be built and placed upon. Pa Pohaku is a collection of local stones that are used to build up the foundation wall for a Hale. The pa pohaku is placed where the hologram is shown “green” as an indication of object placement. The completion of this step leaves the outline of the digital pa pohaku overlaying on the Pa Pohaku that was placed.

#### 4.2. O'A (RAFTER)

O'a, rafter, is highlighted in the assembly steps, one by one, additionally giving the user visual instructions on how to tie the lashing around each O'a.



Figure 8. Diagram of Lashing Instruction of O'a

#### 4.3. POU HANA (RIDGE POST) + KAUHuhu (RIDGE POLE)

In the assembly, a Pou Hana (ridge post) is erected under the Kauhuhu (ridge pole) then lashed together. Following the lashing of the ridge post to the ridge pole, the Kua'iole (upper ridge pole) is then placed on top between the O'a and lashed together with the Kauhuhu. The lashing of the two ridgepoles is explained with the diagrams.

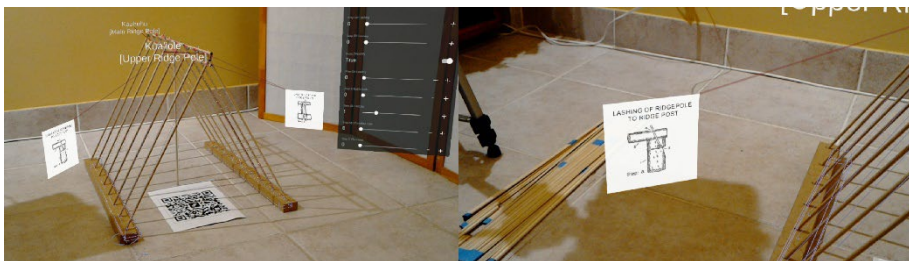


Figure 9. Lashing Instructions of Pou Hana (Ridge Post) and Kauhuhu (Ridge Pole)

#### 4.4. GABLE WALLS

The gable is attached to the O'a and Kauhuhu on the ends of the model. Each gable wall is comprised of a Kalapau (gable end tie), two Kukuna li'i (upper wall post), and a Kupono (gable ridge pole). After the Kalapau and Kupono are lashed on to the model, the Kukuna li'i gets lashed to the Kalapau and the O'a.

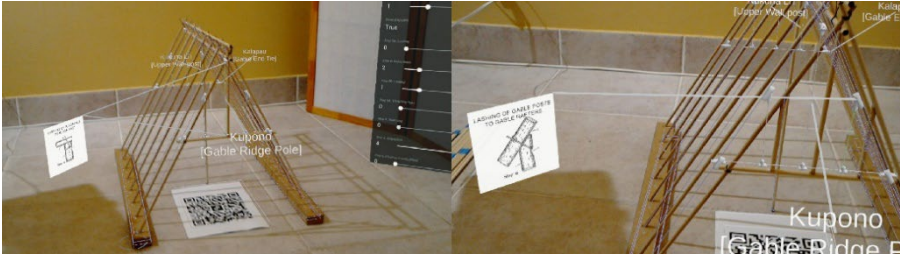


Figure 10. Lashing Instructions of Pou Hana (Ridge Post) and Kauhuhu (Ridge Pole)

#### 4.5. HOLO (DIAGONAL BRACES)

The Holo means the diagonal braces that are lashed to the O'a starting from one end up near the Kauhuhu down to the Pa Pohaku. The Holo provides the structural stability with keeping the O'a from shifting over time.

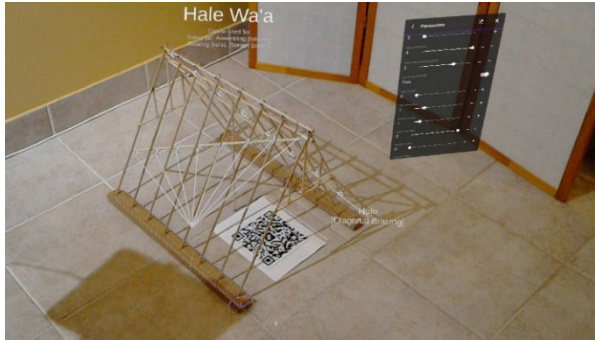


Figure 11. Holo (Diagonal Bracing) in the AR Assembly Process

#### 4.6. 'AHO PUEO + AKO (THATCHING)

The 'Aho Pueo (Purlins) are used to attach thatching on to the Hale. Purlins span horizontally across the O'a, then lashed vertically lined up with the O'a with the horizontal purlins in between. This spacing additionally benefits the Hale by keeping the thatching from making direct contact with the O'a.

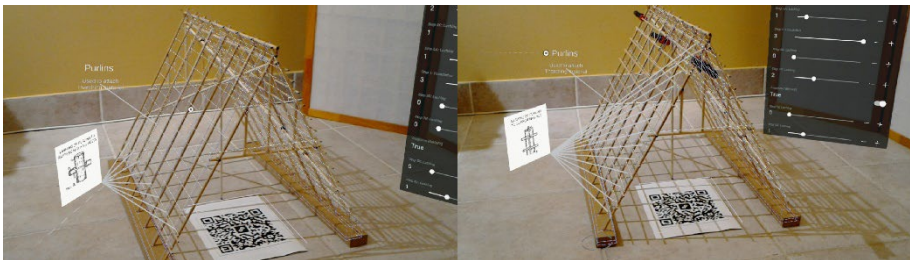


Figure 12. Assembly Placement of Horizontal and Vertical Purlins Against O'a

Depending on the specific thatching material there are a few different ways the thatching can be attached to the Hale. Thatching could be made with pili grass bundle, pandanus leaves, ti leaves, and more depending on the resources of the location where the Hale is being built. According to the selection of the material, the instructional diagram is provided for lashing the thatching to the purlins. The upper thatching is joined by overlapping with the purlin above it. The Thatching is then attached to the purlins on the side of the Hale.



*Figure 13. User Assessing the Variety of Thatching Options of the Hale*

## 5. Discussion

In the assembly of the Hale model, real-time feedback was achieved through the procedural steps toggled within the view of the Microsoft HoloLens. With cycling through each style from the library of digital Hale, a user can initiate the target of his/her investigation. The list of the selected Hale's assembly procedures allows the user to proceed. At each procedure, the wireframe of the partial components of the Hale is overlaid as a holographic guidance in an extended reality. Through the guidance, the user's hands-on interaction with two-dimensional and three-dimensional objects in between physical and virtual was intact and dynamic. The visualization of context-specific complex spatial relationships and abstract concepts is supported by text tags and diagrams generated within Fologram. The text tags provided the explanation of historical and cultural contents of each procedure. Furthermore, the diagrams overlaid with a given holographic wireframe of each procedure visualize more complicated steps such as join assembly, lashing, and thatching of the Hale. The learning experience by the proposed AR application using Microsoft HoloLens, a Head Mount Display, allows the user to get fully immersed into the assembly process of the Hale with sharing its cultural significance. The texture of Hale was not applied to the current application for providing an ease of visual access for the assembly of Hale. A toggle function for the texture is being developed for the representation of the Hale as a whole.

In this paper, the proposed AR based construction process provides a way of

learning through making and without being told verbally how to do so as a tool for teaching tacit knowledge embedded in the Hale. This hands-on learning experience would hopefully foster a deeper appreciation and understanding of Native Hawaiian culture and allow for further conservation of this indigenous knowledge.

## References

- Abernathy, J. E. and Suelyn C. T. (1983). *Made in Hawai'i*. University of Hawaii Press.
- Aikau, H. K., & Gonzalez, V. V. (2019). Kūkulu Hale in Hāna, East Maui: Reviving Traditional Hawaiian House and Heiau Building. In *Detours: A decolonial guide to hawai'i* (pp. 211-219). Duke University Press, NC: Durham.
- Apple, R. A. (1971). *The Hawaiian Thatched House*. Island Heritage.
- Arvanitis, T. N., Petrou, A., Knight, J. F., Savas, S., Sotiriou, S., Gargalakos, M. (2007). Human factors and qualitative pedagogical evaluation of a mobile augmented reality system for science education used by learners with physical disabilities. *Personal and Ubiquitous Computing*, 13(3), 243–250.
- Bekele, M. K., Pierdicca, R., Frontoni, E., Malinverni, E. S., & Gain, J. (2018). A survey of augmented, virtual, and mixed reality for Cultural Heritage. *Journal on Computing and Cultural Heritage*, 11(2), 1-36. doi:10.1145/3145534
- Bryan, E. H. (1950). "6. The Home." in *Ancient Hawaiian Life*. Books about Hawaii.
- Buck, P. H. (1957). *Arts and Crafts of Hawaii*. Bishop Museum Special Publication.
- Caudell, T.P. and Mizell, D.W. (1992). Augmented Reality: an Application of Heads-up Display Technology to Manual Manufacturing Processes. In *Proceedings of the 25th Hawaii International Conference on System Sciences* (pp. 659-669).
- Cozzani, G., Pozzi, F., Dagnino, F. M., Katos, A. V., & Katsouli, E. F. (2016). Innovative Technologies for Intangible Cultural Heritage Education and preservation: The case of I-treasures. *Personal and Ubiquitous Computing*, 21(2), 253-265. doi:10.1007/s00779-016-0991-z
- Edwards, J. (2019, March 25). *Social Structure: Hawaiian Hale*. Hawaiian Science and Technology. Maui Magazine. Retrieved December 12, 2019, from <https://www.maumagazine.net/hawaiian-hale/>
- Dampney, K., Busch, P. and Richards, D. (2002). The Meaning of Tacit Knowledge. *Australasian Journal of Information Systems*, 10(1), 3-13
- Gonzales Vargas, J.C., Fabregat, R., Carrillo-Ramos, A., Jové, T. (2020). Survey: Using Augmented Reality to Improve Learning Motivation in Cultural Heritage Studies. *Applied Sciences*, 10(3):897
- Hashimoto, J., and Park, H-J. (2020). Dance with Shadows: Capturing tacit knowledge with smart device augmented reality (SDAR). In *Proceedings of the 38th eCAADe: Anthropologic: Architecture and Fabrication in the cognitive age* (Vol 2, pp. 165-172).
- Kilolani, M. D., and Middlesworth, N. (1992). "Unit 14: Thatched Houses and Other Structures." *Essay. In Resource Units in Hawaiian Culture*, The Kamehameha Schools Press, 197-212
- Nakata, S. (2017). Language Suppression, Revitalization, and Native Hawaiian Identity. *Diversity & Social Justice Forum*, Volume 3, 14-27.
- Nonaka, I. and Takeuchi, H. (1995). *The Knowledge-Creating Company How Japanese Companies Create the Dynamics of Innovation*. Oxford University Press.
- Voinea, G., Gîrbacia, F., Postelnicu, C. C., & Marto, A. (2018). Exploring cultural heritage using augmented reality through Google's Project Tango and arcore. *Communications in Computer and Information Science*, 93-106. doi:10.1007/978-3-030-05819-7\_8
- Warschauer, M., Donaghy, K., & Kuamo'yo, H. (1997). Leoki: A powerful voice of Hawaiian language revitalization. *Computer Assisted Language Learning*, 10(4), 349-361. doi:10.1080/0958822970100405

# CONFLICT AND RECONCILIATION BETWEEN ARCHITECTURAL HERITAGE VALUES AND ENERGY SUSTAINABILITY

*A Case Study of Xidi Village, Anhui Province*

ZHIXIAN LI<sup>1</sup>, XIAORAN HUANG<sup>2</sup> and SZYMON RUSZCZEWSKI<sup>3</sup>

<sup>1,3</sup>*University of Sheffield*

<sup>2</sup>*North China University of Technology*

<sup>2</sup>*Swinburne University of Technology*

<sup>1</sup>*zli230@sheffield.ac.uk, 0000-0003-2649-8137*

<sup>2</sup>*Xiaoran.huang@ncut.edu.cn, 0000-0002-8702-2805*

<sup>3</sup>*s.ruszczewski@sheffield.ac.uk*

**Abstract.** Energy consumption in buildings has increased dramatically over the last decade due to population growth, increased demands for indoor environmental quality, and global climate change, resulting in a growing awareness of the importance of sustainable development. Amongst many cases, architectural heritage faces a unique challenge in striking the delicate balance between its value and energy sustainability. The key aim of this paper is to reconcile the impact of sustainable design interventions with the values of architectural heritage by proposing a pragmatic and site-sensitive design approach. This article uses the case of the village of Xidi to simulate the energy consumption of traditional South-eastern Chinese residential buildings using the Designer's Simulation Toolkit (DeST) considering different design options. The material selections and simulation results are further analysed by considering essential heritage building conservation documents and charters to accommodate the need for reducing energy consumption without compromising the value of heritage buildings.

**Keywords.** Architectural Heritage; Energy Consumption; Designer's Simulation Toolkit; DeST; Xidi Village; Huizhou; SDG 11.

## 1. Introduction

Huizhou, an ancient region of China, is located near the Xin'an riverbank at the base of the Huangshan Mountains in southern Anhui Province, and its characteristic buildings represent one of the pivotal styles in traditional Chinese architecture. Their recognisability relies in the use of such characteristic elements as grey and black tiles and white walls, as well as in having four dwellings enclosed in a courtyard forming a patio (Figure 1). Within this context, this article focuses on the village of Xidi due to its recognition by United Nations Educational, Scientific and Cultural Organization

and listing in the World Heritage List in 2000.

In recent years, most research on Huizhou architecture has focused on the sustainability of energy consumption, aesthetic features, and culture of traditional Huizhou architecture (Sun & Guan 2012; Lu 2016). What emerges from this research is the focus on experimental testing, preferring to quantitatively analyse the impact of building energy sustainability, without however connecting the sustainability of energy consumption with the sustainability of historical values. In this paper, the issue of sustainability is considered based on its twofold context: on the one hand, it refers to the issue of energy consumption, and on the other hand, it evaluates the sustainability of heritage as a resource to be preserved for the future generations. How to reconcile these two aspects is the core argument of this paper. Firstly, the paper discusses architectural heritage values and energy sustainability in the context of the village Xidi, and it proposes a theoretical framework for interventions through the analysis of key literature, legislation, and charters related to both fields. Based on this discussion, existing problems related to the value and the meaning of the buildings, to the environmental sustainability, and to the needs of the present users have been discussed through a series of online interviews. Possible interventions, identified as possible solutions to these problems, have been thus simulated and discussed against their feasibility both in relation to their efficiency and respect of the material and cultural values of the buildings. The aim of this approach is to use digital simulations to plan heritage renovation without harming the buildings' value and to provide a reference for similar sustainable renovation projects.

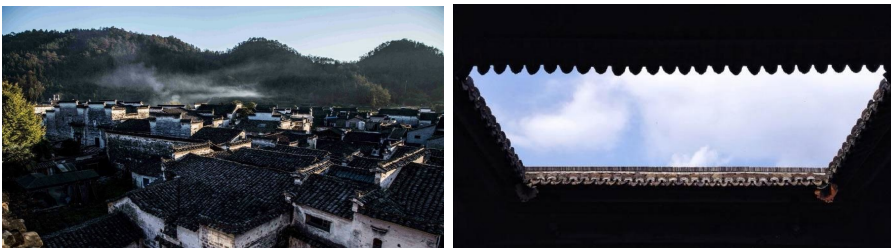


Figure 1. Left: Typical features of Huizhou style architecture (source: <https://weibo.com/>), with white walls and black tiles; Right: a patio formed by the walls and houses (source: <https://k.sina.cn/article>)

## 2. Research Framework

This research establishes a theoretical framework by investigating key academic literature, charters, and legislation documents, suggesting a series of principles that a sustainable renovation project should practise. Secondly, through site investigation, the problems of Xidi residential buildings are identified and proposed renovation strategies are tested and verified through digital simulation (Figure 2). This research is based on a design-oriented simulation system, Designer's Simulation Toolkit 2.0 (abbreviated as "DeST") with the latest updates (9th of July 2021 release). DeST can be used to simulate and analyse building energy consumption and HVAC (heating, ventilation and air conditioning) systems, as it aims to improve the reliability of system design, to ensure the quality of system performance, and to reduce the energy



consumption of buildings (Yan et al., 2008). This paper uses DeST to simulate the year-round energy consumption of a typical building in Xidi village to optimise passive strategies and provide support for subsequent construction.

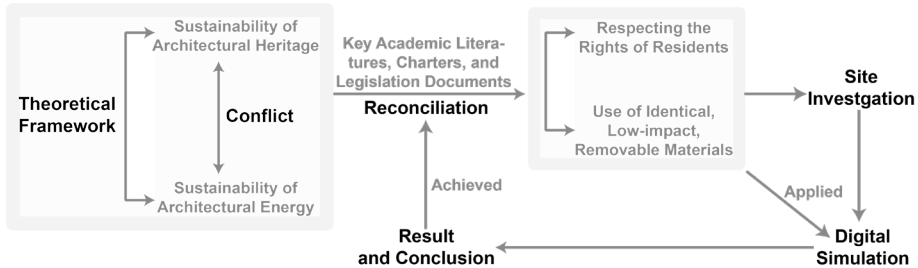


Figure 2. Project's research framework

### 3. Theoretical Framework

#### 3.1. SUSTAINABILITY OF ARCHITECTURAL HERITAGE

The core issue of this paper is how to reconcile sustainability in terms of energy consumption in buildings with sustainability in terms of architectural heritage. It is difficult to discuss architectural heritage without reference to one word: authenticity. This is one of the key criteria for eligibility for listing as a World Heritage Site, and this article discusses the conflict and reconciliation between the authenticity of buildings and the sustainability of buildings' energy consumption. As one of the key concepts of the Venice Charter (1964, p. 1) for the conservation of architectural heritage, authenticity appears for the first time in a dignified way in the first paragraph of the text. "People are becoming more and more conscious of the unity of human values and regard ancient monuments as a common heritage. The common responsibility to safeguard them for future generations is recognized. It is our duty to hand them on in the full richness of their authenticity". Based on the Venice Charter, the World Heritage Committee developed subsequently the Nara Document on Authenticity (1994) which declares for the first time the sources of information for authenticity, which include: form and design, materials and substance, use and function, traditions and techniques, location and setting, spirit and feeling, and other internal and external factors. These can be clearly related to the values recognised in the document nominating Xidi as a World Heritage Site, stating "Xidi and Hongcun have faithfully preserved the typical elements of a former traditional village, including the surroundings, the man-made waterways, the layout of the village, the architectural style, the decorative arts, the building methods and materials, the traditional techniques and the overall appearance of the village" (2000, p. 10). In this sense, the value of this village lies not only in the style and historical aspects, but also as a testimony of the use of materials and craftsmanship, pointing to potential problems when dealing with new interventions.

#### 3.2. SUSTAINABILITY OF ARCHITECTURAL ENERGY

In the report of the World Commission on the Environment and Development, United Nations Brundtland Commission (1987) defined "sustainability" as meeting the needs

of the present without compromising the ability of the future generations to meet their own needs. Similarly, Charles Keibert stated in 1994 that sustainable building is the creation and responsible management of a healthy built environment based on resource efficient and ecological principles. Emerging from these definitions, one understands clearly that sustainable architecture intends to minimise the negative impact of buildings on the environment through the efficient and proportionate use of materials, energy and the ecosystem as a whole, and to adopt a conscious approach to energy and ecological conservation in the design of the built environment. An approach, which is a different field could be compared to the concept of “preservation for future generations” mentioned in the Venice Charter.

#### **4. Conflict and Reconciliation**

That said, heritage protection and sustainable architecture, as different focuses in the field of architecture, would require finding the right balance between reaping the benefits of a sustainable building and altering architectural heritage. Consequences of energy efficiency interventions, when applied inappropriately and with no understanding of heritage, can easily damage not only meaning, value, and stability of the buildings, but also users’ health, and eventually, they could fail to achieve the desired savings or reduction of the environmental impact (Historic England, 2018). Achieving the balance and avoiding unintended consequences requires a multitude of considerations, but it would be important to focus here on two key aspects of the context of the village of Xidi: the needs of the present-day residents and the use of specific materials.

##### **4.1. RESPECTING THE RIGHTS OF RESIDENTS**

When regenerating heritage buildings, which are still in use, the comfort of the original inhabitants and the “vitality” of the heritage building need to be respected. Already in 1931, the Athens Charter states that the conservation of heritage buildings should not compromise the inhabitants’ rights to a healthy living environment. In any case, the value of heritage must not override the interests of the living environment, which is directly related to the welfare and physical and mental health of the individual. The treaties mentioned above indicate that the habitability of buildings is also part of the conservation of architectural heritage. The Washington Charter (1987) also emphasises the inseparable relationship between historic cities or regions and those who inhabit them, suggesting that new functions and activities should be adapted to the characteristics of historic towns and urban areas. It has been also claimed that the living environment of their inhabitants is a priority in the conservation of historic cities or regions, while many existing and potential uses of the World Heritage sites can contribute to the quality of life of the communities (International Council on Monuments and Sites, 1964). Similarly, the Machu Picchu Charter (1977) states that conservation, renovation, and reuse of existing heritage sites and architecture should be integrated into the urban development process in order to ensure the economic significance and continued viability of these objects. Keeping these buildings vibrant means to optimise the living environment for the residents. As a result, in contrast to uninhabited buildings, heritage-protected dwellings and villages are concerned both

with the cultural values and the importance of the living environment for the inhabitants. In addition, new interventions would need to consider not only conservation-based approaches, but also regeneration and development: the primary function of a building is to be used and inhabited.

#### 4.2. USE OF IDENTICAL, LOW-IMPACT, REMOVABLE MATERIALS

Regeneration with the same materials as the original building is also a way of reducing damage to the heritage value of the building. Alterations must impact the existing forms and materials as little as possible, whether dealing with regeneration of an isolated building or of an entire village. Choosing those technical measures that do not compromise the authenticity of heritage or are demountable, low-impact, and easily recognisable (as suggested by the Venice Charter) could be therefore a possible solution. For example, in the Hoi An Declaration (2003), it is suggested that new materials with the same or lower strength and hardness than the original ones should be used in the renovation and conservation of old buildings. However, whereas the very ideas of removability and different characteristics of materials are appealing when dealing with the importance of heritage, such solutions need to be assessed also from the point of view of sustainability.

### 5. Present-day Issues in Xidi

The current research on building energy consumption in Xidi is aligned with the general research in Huizhou buildings. It focuses mainly on summer insulation, patio lighting, window lighting, and winter weather resistance. These issues are also aligned with the problems pointed by the present-day inhabitants. A series of six online interviews about the satisfaction of the living environment in Xidi were conducted with local residents. The interviews include the candidates' personal information, average time spent indoors per day, available room area, frequency of window ventilation per day, satisfaction with indoor temperature, satisfaction with indoor humidity, satisfaction with lighting, satisfaction with air quality and satisfaction with noise. The results indicated that all of the candidates indicated that the room temperature was too low in winter. Since the area is not in the far north of China, only few contemporary buildings in Anhui Province are equipped with heating devices, and the very layout of the dwellings in Xidi favours flow of cool air in the winter, mainly due to the additional airflow from the patio. In addition, four of the residents indicated that the indoor temperature of the building was also too high in summer, and that cooling equipment was used to achieve a more comfortable indoor environment, making effective insulation an important issue to be solved, which is discussed later.

Furthermore, half of the residents indicated that rooms were insufficiently illuminated, and this problem is also taken into account in the renovation measures. This issue is related to the patio-centred layout of the dwellings in Xidi, most of which are very similar, even though buildings have some minor differences. Amongst these, some typical layouts became prototypes for others, as a number of studies already summarise (Lv, 2005; Zhang et al., 2013; Che, 2015). According to them, most of dwellings in Xidi can be formed from two basic architectural layouts through a number of combinations, which has led to modelling the two combinations in DeST (Figure 3).

These models follow a similar pattern, where the patio remains the central part of a house. Based on the above, a renovation method is proposed for the Xidi residence, in order to meet the needs of local residents: the installation of a shading structure on the patio that can be opened or closed (Figure 4).

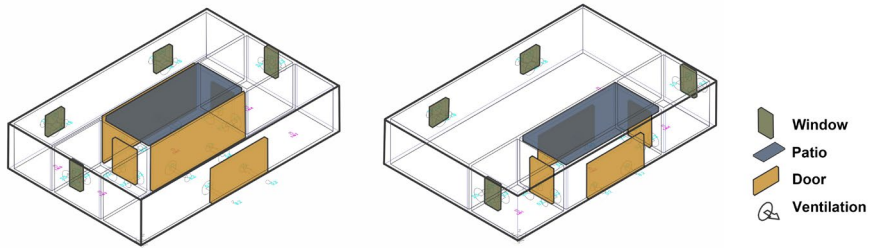


Figure 3. Left: Typical layout 1; Right: Typical layout 2

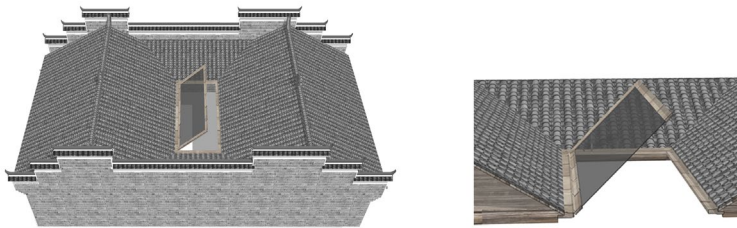


Figure 4. A possible renovation measure: the addition of removable shelters on the patio to keep the room warm in winter

## 6. Simulation Result and Discussion

This modelling process included adding the following data to the simulation: geographical location, building materials, opening sizes in order to evaluate the effects of ventilation, and heat generated by human activity in each room (which, accordingly with the sixth census of China in Anhui Province, has been set to the heat generated by 1-2 inhabitants). Furthermore, based on interviews with local residents, it has been assumed that the buildings' users would turn on cooling or heating equipment when the room temperature makes them uncomfortable in order to save the electricity bill. Taking into account Anhui Province's monsoon and subtropical monsoon climate, with the summer season from June to August and the winter season from November to January, it has been assumed that with the present-day situation, heating would be necessary on ninety days (2 hours per day) and cooling on ninety days (2 hours per day) throughout the year. For example, outdoor temperatures in Xidi range from  $-5^{\circ}\text{C}$  to  $16^{\circ}\text{C}$  in the coldest months, with an average temperature of around  $5^{\circ}\text{C}$  (Figure 5), which is below the World Health Organisation's standards for human habitable temperatures.

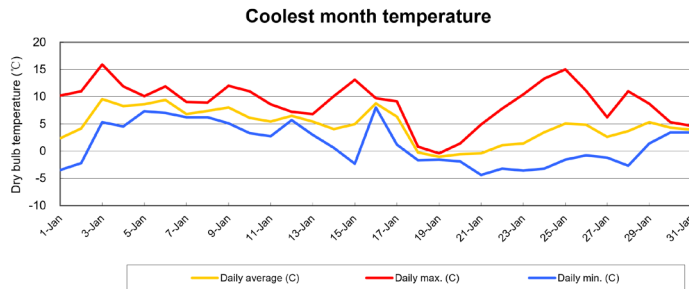


Figure 5. Temperature in Xidi Village, Anhui in the coldest month

Within this context, the simulation process proceeded taking into account the concept of “vitality”, i.e. “the rights of residents with less damage to residential heritage buildings”, which have been previously mentioned. However, at the same time, it is worth noting that common measures, such as adding thermal insulation to walls or patios, cannot be considered due to the heritage value of the buildings and their typical “white-walled, black-tiled” character. Given that, the only insulation possibility could be adding such a layer to the roof.

During the interviews, the residents also mentioned adding glazing to the patio, which was in fact considered in the DeST. It would have increased the internal temperature control without impairing the already difficult natural sunlight access to the inhabitable areas of the dwellings. Indeed, based on the existing research, one of the more significant effects on thermal insulation is insulating glass, particularly low-e film insulating glass. Glass, insulated glass, and low-e film insulating glass were finally selected for testing and the data from the simulations are shown in Figure 6 to left of Figure 7. Besides, since the patios in Xidi are mostly wooden, based on the idea of using similar materials expressed before, it has been considered to install removable timber structures in the patio, using the same spruce timber as in the traditional local constructions (right of Figure 7) (1-N-2 is the patio room and 1-N-1, 1-N-3, 1-N-4 and 1-N-5 are the normal rooms).

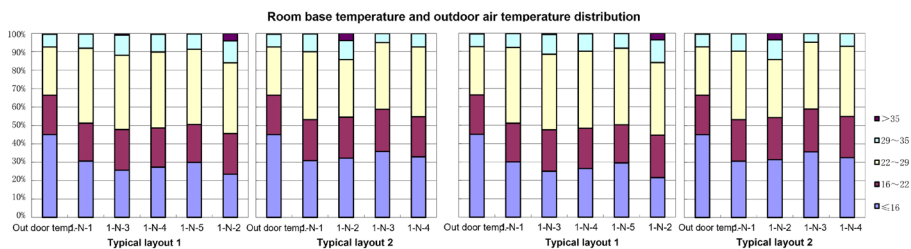


Figure 6. Left: Temperature distribution throughout the year in patio with ordinary glass(12mm); Right: Temperature distribution throughout the year in patio with insulated glass (6mm glass+air space+6mm glass)

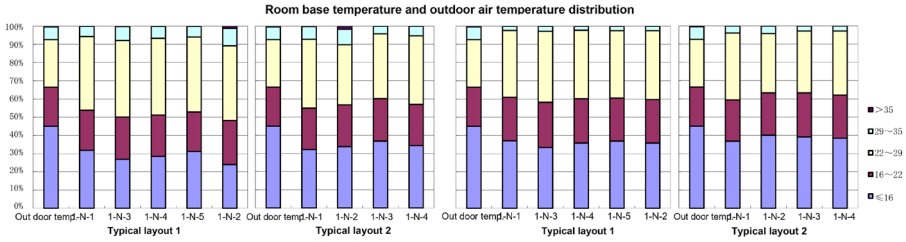


Figure 7. Temperature distribution throughout the year in patio. Left: With low-e film insulating glass (6mm glass+air space+6mm glass); Right: With timber installation (12mm)

According to these four comparison experiments, it can be demonstrated that the percentage of extreme temperatures (less than 16°C and more than 29°C) is the smallest of the three types of glass when using low-e film insulating glass. This situation is particularly noticeable in summer. Therefore, the installation of low-e film insulating glass is more effective than ordinary glass and insulated glass. Moreover, as shown in the right of Figure 8, the installation of timber is more effective than plain glass, insulated glass and low-e film insulating glass in summer. However, timber shading is the least effective in winter insulation, and one of the issues raised by the residents of Xidi is the lack of light and timber interferes with the lighting. As mentioned, one of the principles of reconciling architectural heritage values with energy sustainability is to respect the rights of the inhabitants, thus low-e film insulating glass considered to be more suitable for the dwellings. In order to manage the potential negative effects of low-e film insulating glass in summer, an additional removable shading could be provided (such as locally available dry bamboo canes), as well as regular ventilation by opening the windows could be advisable.

As the last part of this analysis, this paper simulates the thermal and cooling loads of a building with a typical Xidi layout equipped with low-e film insulating glass and without low-e film insulating glass (fifty times ventilation per hour) in order to demonstrate the positive impact of adding patio glass on building energy consumption (Figure 8 and Figure 9). By comparing these diagrams, it is clear that the installation of the low emissivity coating glass results in significant energy savings in the winter months in both typical layouts. This last piece of information is important as it suggests that such an intervention can have a broader application in a number of cases of dwellings, which – as stated before – follow a similar layout pattern or are modifications of these two basic plans.

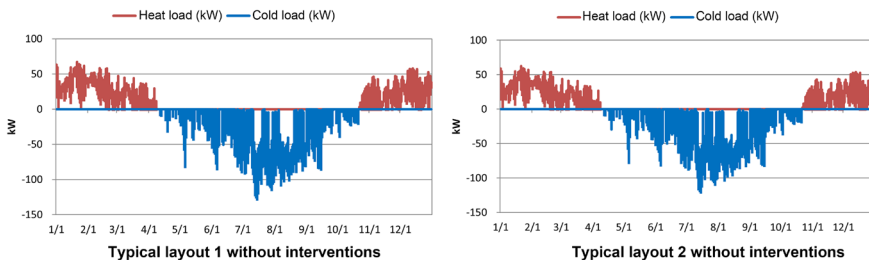


Figure 8. Typical layout 1 & 2, heat and cold loads in buildings without installation of low-e film insulating glass during one year

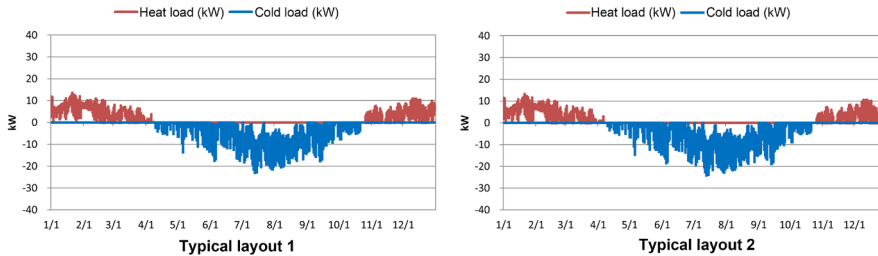


Figure 9. Typical layout 1 & 2, heat and cold loads in buildings after installation of low-e film insulating glass during one year

## 7. Conclusion

While this paper concentrates on Xidi as a case study for analysing how to effectively reduce energy consumption in heritage buildings, some of the conclusions here have broader implications. The basic theoretical framework discussed throughout the article stands as a valid common ground for a number of interventions, which engage in both fields – sustainability and heritage conservation. The issue of authenticity in relation to the present-day users and inhabitants stands as a core problem in all the buildings that are constantly and continuously used. The simulation results indicate that, albeit the timber is the most suitable option for this project in the summer, its insulation effect in winter is worse than the other three types of materials, and the principle of respecting the residents' rights would be unattended due to the interfered lighting. Therefore, the low radiation coating glass option is suggested to be the most suitable solution for the patios renovation in houses in Xidi. It is important to note that the change of buildings' layout had a relatively small impact on the overall simulation, while the change of materials was crucial. It leads to the important assumptions that the above simulation would remain valid in a number of other cases and that further research in sustainability would certainly benefit from the historical and cultural background, which would point to the typical and traditional solutions, which would help to draw the common ground for the possible interventions. Future work will require other cases to explore the relationship between other aspects of building energy efficiency and architectural heritage and develop rulemaking in procedural, systematic, and more in-depth simulations.

## References

- Che, X. (2015). *A study on the renewal of the internal spatial form of traditional Huizhou houses* [Master's thesis, Xi'an University of Architecture and Technology]. Cnki. <https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD201601&filename=1015995737.nh>
- Congrès Internationaux d'Architecture Moderne. (1931). The Athens Charter for the Restoration of Historic Monuments. In *4th International Congresses of Modern Architecture s.* Athens, Greece.
- Historic England. (2018). *Energy efficiency and historic buildings*. Historic England. Retrieved November 2, 2021, from <https://historicengland.org.uk/>
- International Council on Monuments and Sites. (1964). International Charter for the Conservation and Restoration of Monuments and Sites (The Venice Charter 1964). In *2nd*

- International Congress of Architects and Technicians of Historic Monuments*. Venice, Italy.
- International Council on Monuments and Sites. (1987). Charter for the Conservation of Historic Towns and Urban Areas. In *ICOMOS General Assembly in Washington*, DC October 1987. Washington, DC, USA.
- International Council on Monuments and Sites. (1994). The Nara Document on Authenticity. In *Nara Conference on Authenticity in Relation to the World Heritage Convention*. Nara, Japan.
- International Council on Monuments and Sites. (2003). The Hoi An Declaration on Conservation of Historic Districts of Asia. In *International Symposium on the Conservation of Cultural Heritage Sites and International Cooperation*. Hoi An, Vietnam.
- World Architecture Construction Federation. (1977). The Charter of Machu Picchu. In *The Universidad Nacional Federico Villarreal*. Lima, Peru.
- Kilbert, C. J. (1994). *Sustainable Construction: Proceedings of the 1st International Conference*, 1994. International Council for Building Research, Studies and Documentation.
- Lv, H. (2005). *Sustainable development pattern and its experimental planning in Chinese case study villages* [Doctoral dissertation, Xi'an University of Architecture and Technology].  
Cnki.<https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CDFD9908&filename=2005085031.nh>
- Lu, J. (2016). Study on visual form design of Huizhou folk house. In *Proceedings of the 2016 4th International Conference on Management Science, Education Technology, Arts, Social Science and Economics* (msetasse-16) (pp. 307–311). Atlantis Press.
- National Bureau of Statistics of China. (2010). *Tabulation on the 2010 population census of the People's Republic of China*. National Bureau of Statistics of China. Retrieved August 2, 2021, from <http://www.stats.gov.cn/tjsj/pcsj/rkpc/6rp/index.htm>
- Sun, Q., & Guan, C. C. (2012). Existing condition analysis and countermeasures of Huizhou traditional architecture construction. *Applied Mechanics and Materials*, 209–211, 18–22. <https://doi.org/10.4028/www.scientific.net/AMM.209-211.18>
- United Nations. (1987). *Sustainability*. United Nations. Retrieved August 12, 2021, from <https://www.un.org/en/academic-impact/sustainability>
- United Nations Educational, Scientific and Cultural Organization. (1972). *Convention concerning the protection of the world cultural and natural heritage*. United Nations Educational, Scientific and Cultural Organization.
- United Nations Educational, Scientific and Cultural Organization. (1987). *Operational guidelines for the implementation of the World Heritage Convention*. United Nations Educational, Scientific and Cultural Organization.
- United Nations Educational, Scientific and Cultural Organization. (2020). *Document nominating Xidi*. United Nations Educational, Scientific and Cultural Organization. Retrieved July 17, 2021, from <https://whc.unesco.org/en/list/1002/documents/>
- Yan, D., Xia, J., Tang, W., Song, F., Zhang, X., & Jiang, Y. (2008). DeST – An integrated building simulation toolkit Part I: Fundamentals. *Building Simulation*, 1(2), 95–110. <https://doi.org/10.1007/s12273-008-8118-8>
- Zhang, R. X., & Liu, T. J. (2013). The Development of An Ancient Village - Xidi, Wannan. *Advanced Materials Research*, 671–674.



# USE OF OBJECT RECOGNITION AI IN COMMUNITY AND HERITAGE MAPPING FOR THE DRAFTING OF SUSTAINABLE DEVELOPMENT STRATEGIES SUITABLE FOR INDIVIDUAL COMMUNITIES, WITH CASE STUDIES IN CHINA, ALBANIA AND ITALY

FRANCESCO GRUGNI<sup>1</sup>, MARCO VOLTOLINA<sup>2</sup> and TIZIANO CATTANEO<sup>3</sup>

<sup>1,3</sup>*Università degli Studi di Pavia, <sup>2</sup> Politecnico di Milano.*

<sup>1</sup>*grugnifrancesco@gmail.com, 0000-0003-3198-9950*

<sup>2</sup>*marcovoltolina@outlook.it, 0000-0002-4786-7475*

<sup>3</sup>*tiziano.cattaneo@unipv.it, 0000-0003-1604-6831*

**Abstract.** In order to plan effective strategies for the sustainable development of individual communities, as prescribed by the United Nations' Sustainable Development Goal 11, it is necessary for designers and policy makers to gain a deep awareness of the bond that connects people to their territory. AI-driven technologies, and specifically Object Recognition algorithms, are powerful tools that can be used to this end, as they make it possible to analyse huge amounts of pictures shared on social media by residents and visitors of a specific area. A model of the emotional, subjective point of view of the members of the community is thus generated, giving new insights that can support traditional techniques such as surveys and interviews. For the purposes of this research, three case studies have been considered: the neighbourhood around Siping Road in Shanghai, China; the village of Moscopole in southeastern Albania; the rural area of Oltrepò Pavese in northern Italy. The results demonstrate that a conscious use of AI-driven technologies does not necessarily imply homogenisation and flattening of individual differences: on the contrary, in all three cases diversities tend to emerge, making it possible to recognise and enhance the individuality of each community and the genius loci of each place.

**Keywords.** Sustainable Communities; Artificial Intelligence; Object Recognition; Social Media; SDG 11.

## 1. Introduction

According to the United Nations, “sustainable cities and communities” are a crucial goal to achieve before the year 2030, as prescribed by Sustainable Development Goal 11 (SDG 11). Specifically, some of the targets outlined in this goal are: to “enhance inclusive and sustainable urbanisation and capacity for participatory, integrated and sustainable human settlement planning and management”, to “strengthen efforts to protect and safeguard the world’s cultural and natural heritage”, and to “support positive economic, social and environmental links between urban, peri-urban and rural

areas by strengthening national and regional development planning” (United Nations General Assembly, 2015, pp. 21-22).

In order to achieve such targets, designers and policy makers need operative tools to analyse the necessities of local communities and the unique relationship that they entertain with the place they inhabit. This research argues that AI-driven technologies, and specifically Object Recognition algorithms, can be effectively used for this purpose.

In recent years, researchers have applied this kind of algorithms to the analysis of cities and neighbourhoods, focusing on parameters related to inequality, safety and liveliness (see for example Salesses et al., 2013, and De Nadai et al., 2016). In these studies, the image dataset was composed of geo-tagged pictures taken from Google Street View or collected manually by the authors. A study by MIT researchers, however, points out that this choice presents some limitations, as “a lot of content related to city perceptions, such as mountains and crowded indoor scenes are missing”, and proposes instead to use images taken from photo-sharing platforms such as Flickr and Panoramio (Zhou et al., 2014, p. 522). Through this different strategy, the researchers try to define the *identity* of cities, and analyse millions of geo-tagged images according to seven pre-determined attributes: green space, water coverage, transportation, architecture, vertical building, athletic activity, and social activity (pp. 524-525).

Rather than the *identity* of a city, the present research aims at defining what Kevin Lynch called the *image* of the city (Lynch, 1960), that is how the community lives and perceives the place. The purpose here is to understand the relationship that a community entertains with a place through the lens of the emotional, subjective point of view of its members. In order to do so, this study proposes a new approach, based on:

- A multi-scale strategy through which it is possible to consider a specific neighbourhood, a city or even an entire region, based on the local context, the features of the community and the purpose of the analysis;
- A dataset composed of geo-tagged images taken from a popular social network such as Instagram, as it reflects everyday life and subjective perceptions more accurately than Google Street View or photo-sharing platforms;
- Flexibility and absence of pre-determined attributes, as the purpose in this case is not to compare different places according to fixed parameters, but to allow their specificity and uniqueness to emerge;
- Versatility, as the approach can be successfully applied to places in different parts of the world, with different cultures, demographics and economies (this is demonstrated by the diversity of the proposed case studies).

## 2. Method

The method followed for this research is composed of five main steps. The first one consists in defining the object of the analysis. As mentioned above, this study follows a multi-scale strategy, so it is possible to select, for example, a specific neighbourhood,

a village, a city or even a larger territory. Another possibility is to work on two different scales at the same time, as experimented in the third case study. What is important in this phase is to define an area that is meaningfully associated with one or more communities of people.

The second step is to define the dataset. All the pictures of this research have been taken from Instagram, as it is a popular image-based social media. The images have been collected through a python script, based on their geo-tag and hashtags. The number of pictures collected for each case study depends on the local context and the scale of the considered area: for the first two case studies (a neighbourhood and a village), 500 images have been selected, while for the third one (a larger-scale territory) the number increased to 1,200. The date range is the same for all the case studies, as in this research have only been used pictures shared on Instagram during the year 2021.

The third step is to apply to the dataset an Object Recognition algorithm that includes the following tools: the SIFT (Scale-Invariant Feature Transform) system and the Lowe method (Lowe, 1999), the R-CNN method (Girshick et al., 2014), the Spatial Pyramid Pooling strategy (He et al., 2014), and the SSD model (Liu et al., 2016). For this study, the algorithm has been pre-trained with millions of images and has learnt to recognise around 10,000 different objects (“picturable nouns”). It has also been programmed to select, for each picture, only the five most prominent objects, so as to avoid giving irrelevant or redundant information as output. As the research progressed, the algorithm has been trained to recognise new objects that were peculiar to the analysed areas, thus enabling it to better highlight the uniqueness of each place. This process gives as a result a graph that shows the frequency with which different objects appear in the image dataset (the graphs in this paper include only a selection of objects, but the algorithm actually recognised many more).

For the fourth step, it is necessary to select a limited number of significant objects: in this research, the focus has been put on five of the objects that appeared more frequently in each case study. Then, the positions of all the pictures in which these significant objects appear are located on a map. This is done through the geo-tags and metadata of the pictures or, when the geo-tag is generic, through a software that defines the position of the image relative to a set of known points (e.g. famous buildings, monuments, infrastructures).

Finally, the results are discussed in order to understand how the output data (graphs and maps) can be used as a basis for drafting strategies for the sustainable development of the local communities.

## 2.1. LIMITATIONS

The effectiveness of strategies based on AI-driven technologies always depends on the quality of input data. As the method proposed in this study takes its dataset from social media, it is important to be aware that these platforms are used by certain demographic groups more than others.

In fact, a research carried out in 2021 by Auxier and Anderson shows that most Instagram users in the United States of America are young, have studied in college and live in urban areas, while categories such as elderly people and residents in rural areas are much less present on the platform.

For this reason, it is important to stress that the method proposed in this research is not intended to be used alone, but rather as part of a wider strategy that also involves other tools and techniques. In fact, it is only by using an adequate array of methods that it is possible to carry out an analysis that reflects the feelings and perceptions of the whole community.

### 3. The Neighbourhood of Siping Road, Shanghai, China

The first case study that will be discussed is the analysis of the neighbourhood around Siping Road, in the Yangpu district of Shanghai, China. The area is part of one of the most populous cities in the world, with more than 24 million inhabitants. Specifically, the segment of Siping Road on which the analysis has focused is the one close to Tongji University's main campus. A dataset of 500 pictures has been collected from Instagram pictures with geo-tags located in the area and hashtags related to the neighbourhood.

#### 3.1. OUTPUT

The output of the Object Recognition algorithm (fig. 1) shows that many of the elements that appear in the pictures are connected to the presence of the university: “graduation” (42 appearances), “student” (32), “graduation photo” (30), “ceremony” (12), and “lecture” (10). The other themes that emerge from the analysis are related to social activities (“indoor”, 36 appearances; “food”, 27; “get-together”, 15; “night”, 12), natural elements (“tree”, 33), urban landscape (“architecture”, 24; “building”, 12) and art (“art installation”, 10).

Then, the distribution of five elements has been located on a map (fig. 2): “food”, “student”, “tree”, “graduation” (also including “graduation photo”), and “indoor”. The result is a map that shows the areas usually frequented by students, and the places that they consider important enough to have their graduation photos taken there, but also the location of popular bars and restaurants, of the most frequented public buildings (e.g. the university library), and the spots where trees are found.

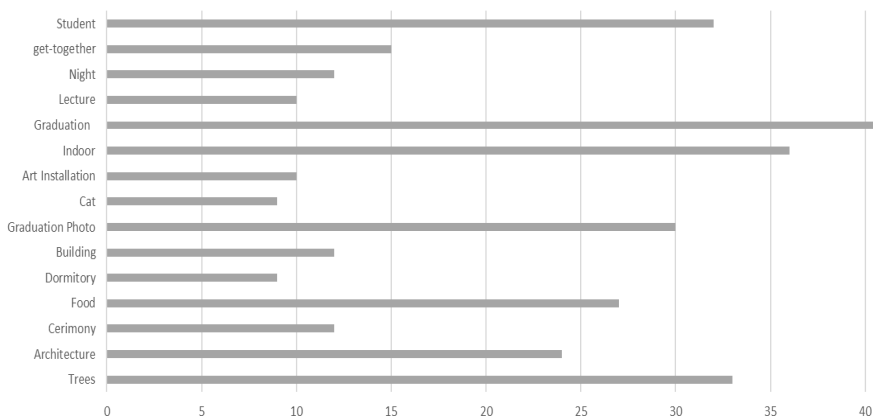
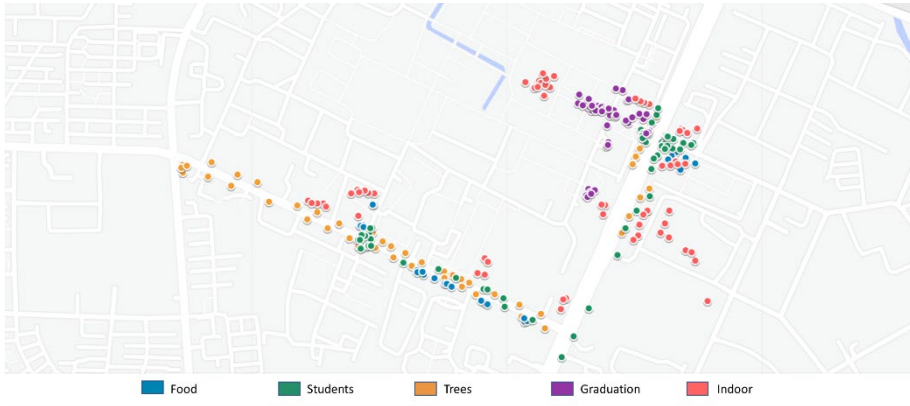


Figure 1. Output of the Object Recognition algorithm, Siping Road.



*Figure 2. Distribution of five frequent elements in the neighbourhood of Siping Road.*

### 3.2. CONSIDERATIONS

The experiment shows the potential of the Siping Road neighbourhood, especially from the point of view of its student community. Its output describes what kind of social activities are practised in the area and where exactly they are practised. Based on the results, it is possible to imagine strategies for establishing beneficial relationships between students, residents and business owners, and for improving the public spaces of the neighbourhood: on one hand, by enhancing the most frequented areas, adapting them to the activities that they already host; and on the other hand by regenerating less frequented areas, reproducing in them the traits of more successful parts of the same neighbourhood.

## 4. The Village of Moscopole, Albania

The second case study focuses on Moscopole, a village of little more than 1,000 inhabitants in Korçë County, in southeastern Albania. Despite its small population, however, Moscopole is slowly gaining popularity as a tourist destination, thanks to its mountain landscape and its cultural heritage. The dataset is once again made of 500 Instagram pictures with geo-tags located in the village and its immediate surroundings, and hashtags related to the area.

### 4.1. OUTPUT

In this case, in the output data (fig. 3) it is possible to clearly recognise two main themes: on one hand nature and landscape, and on the other architecture and history. The former includes “outdoor” (43 appearances), “nature” (37), “forest” (20), “hills” (18), “mountain” (14), “snow” (14), “treehouse” (12), “wild” (11), and “quad” (11). The latter, instead, is related to “mediaeval” (47), “church” (24), “city” (13), and “house” (13).

These two themes also define the map (fig. 4) that shows the position of the following five elements: “forest”, “nature”, “outdoor”, “mediaeval” and “church”. The

map clearly suggests what are the most frequented naturalistic paths and cultural heritage in Moscopole. It also demonstrates that people in the village tend to spend time and perform activities outdoors.

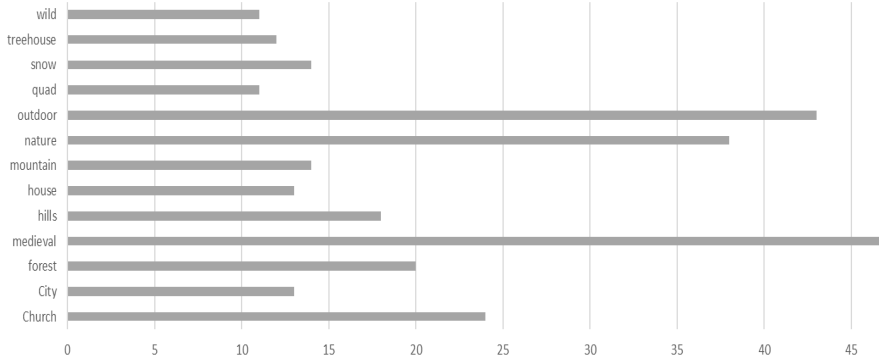


Figure 3. Output of the Object Recognition algorithm, Moscopole.

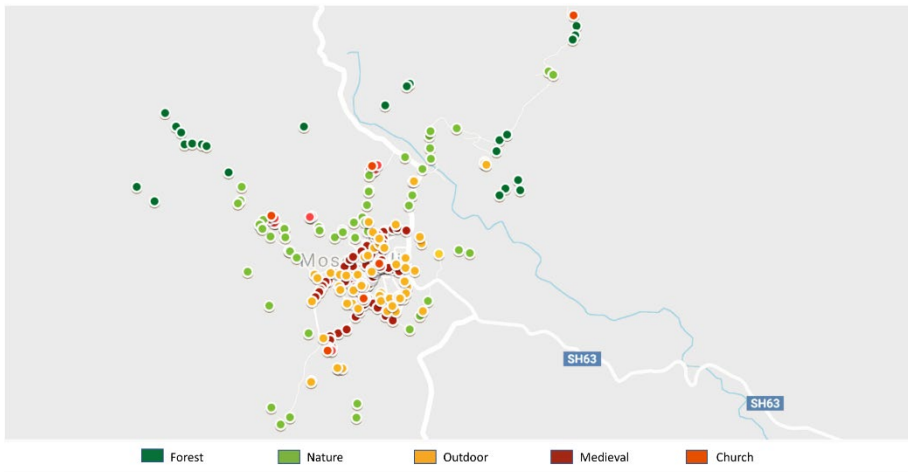


Figure 4. Distribution of five frequent elements in the village of Moscopole.

#### 4.2. CONSIDERATIONS

In this case, the experiment provides a good basis for implementing strategies for sustainable tourism. On one hand, it highlights the potential for naturalistic excursions and sports, also showing existing trails that could be enhanced and promoted. On the other hand, it suggests that Moscopole could also attract tourists interested in its history and architecture, and this could be achieved, for example, by renovating outdoor public spaces, enhancing tourist facilities and preserving its cultural heritage, especially its ancient churches. As these two aspects are interrelated, improvements in any of them would end up being beneficial for the whole tourism sector of Moscopole.

## 5. The Territory of Oltrepò Pavese, Italy

Finally, the last case study takes into consideration a large rural territory, made up of four municipalities (Corvino San Quirico, Mornico Losana, Oliva Gessi and Torricella Verzate) in an area known as Oltrepò Pavese, in the province of Pavia, in northern Italy. The area is known for its wines and castles, and it is home to around 2,600 people. In this case, the analysis has been carried out at two different scales: that of the whole territory and that of each specific municipality. For this reason, the dataset is larger, as it includes 300 Instagram pictures for each municipality, for a total of 1,200 images, always based on geo-tags and hashtags related to the area.

### 5.1. OUTPUT

Compared with the previous case studies, the output data for the Oltrepò territory present a greater variety (fig. 5). In fact, the list of objects that appear in the pictures include elements related to many different semantic fields: nature (“landscape”, 118 appearances; “flowers”, 11, “apiary”, 11), wine (“wine”, 68; “Defilippi winery”, 53; “Monsupello winery”, 33; “Cavallini winery”, 11; “Tenuta di Oliva winery”, 11), historical heritage (“castle”, 61; “church”, 23), sports (“sport”, 42; bike, 27; “swimming pool”, 22; “walk”, 12; “motocross”, 10), weddings (“bride and groom”, 28; “wedding”, 27), social activities (“hamburger”, 13; “show”, 12).

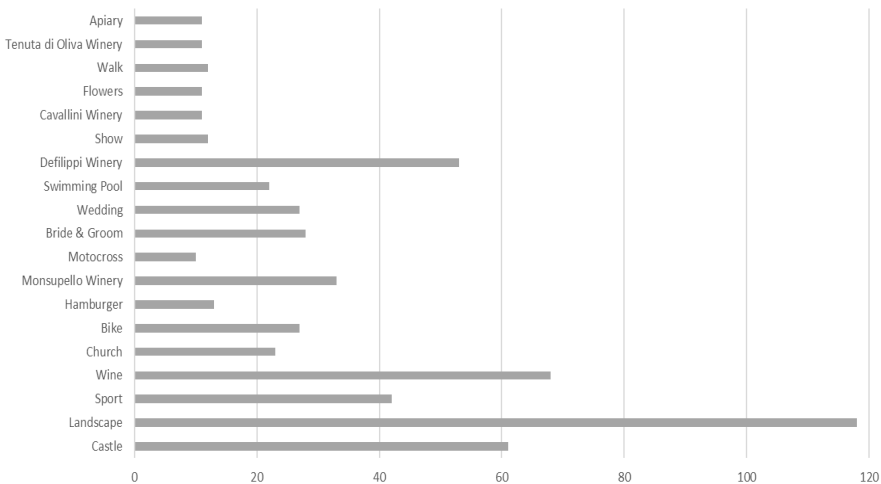


Figure 5. Output of the Object Recognition algorithm, Oltrepò Pavese.

The map with the positions of the most frequent elements (fig. 6) clearly shows how some of them are distributed quite homogeneously on the whole area (“landscape” and “sport”), while other ones are found only in some places (“wine”, that also includes all the different wineries), and still others are specific to certain locations (“castle” and “church”).

For this reason, in this case it is useful to zoom in and consider the territory at a different scale (fig. 7). In other words, after a first general analysis that highlights the

main vocations of the area, a second step is needed to better identify the unique features of each location. By analysing the data of each of the four municipalities, what emerges is that Mornico Losana is a coveted destination for weddings, but it also hosts a historical castle and a popular swimming pool; Corvino San Quirico and Oliva Gessi are frequented mostly because of their landscape and wine, but also for their apiaries; and Torricella Verzate has a popular winery and a historical church, but also good restaurants and places to practice motocross.

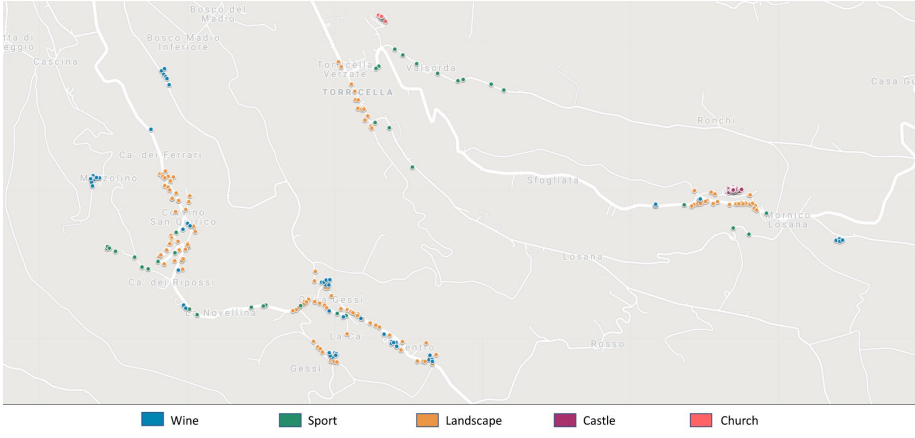


Figure 6. Distribution of five frequent elements in the territory of Oltrepò Pavese.

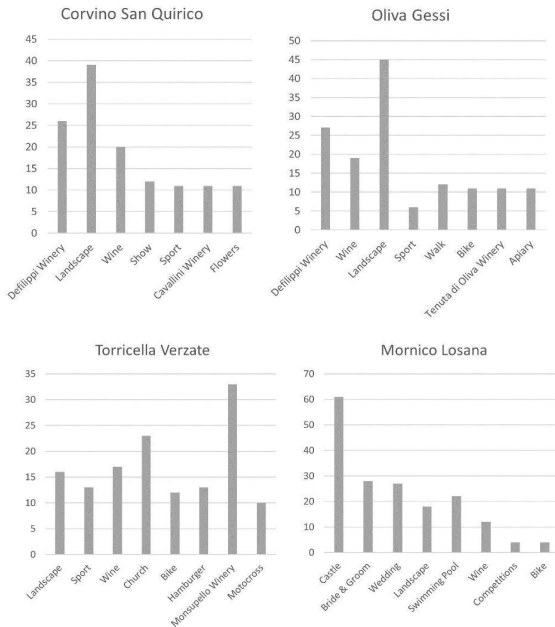


Figure 7. Output of the Object Recognition algorithm: Torricella Verzate, Mornico Losana, Corvino San Quirico, and Oliva Gessi.



## 5.2. CONSIDERATIONS

As the analysis for this case study has been carried out at two different scales, it is possible to outline two sets of strategies. On one hand, the experiment suggests general strategies for the whole territory: enhancing viticulture, food and wine tourism and naturalistic trails. On the other hand, it highlights the need for specific strategies to implement in selected locations: preserving and promoting the cultural heritage (the castle of Mornico Losana and the church of Torricella Verzate), encouraging certain sports (swimming in Mornico Losana and motocross in Torricella Verzate) and enhancing beekeeping in Oliva Gessi.

## 6. Conclusions

This research proposes a method for understanding the subjective relationship between a community and its territory. Three different case studies have been analysed, demonstrating how Object Recognition algorithms can be used to recognise diversity, enhancing the individuality of each community and the genius loci of each place (fig. 8).

Despite its limitations, the proposed method can offer a good understanding of the potential of an individual community and the opportunities offered by its territory. It can thus contribute to creating a basis for drafting effective strategies for the sustainable development of the community, in harmony with its feelings and sensitivity: for example, it suggests to enhance the public spaces for the students of Siping Road, to develop sustainable tourism in Moscopole, and to promote viticulture, apiculture, sustainable tourism and the preservation of natural and cultural heritage in Oltrepò Pavese.

This study demonstrates the great potential of Object Recognition algorithms and how they can be used for the purpose of achieving the UN's SDG 11. In particular, they can play a key role in improving indicator 11.a.1, giving local administrators the tools to draft "development plans that (a) respond to population dynamics; (b) ensure balanced territorial development" (United Nations, n.d.). Furthermore, they can indirectly contribute to enhancing the participation of citizens to the planning process (indicator 11.3.2) and to improving the efficiency of public funding for the preservation of natural and cultural heritage (indicator 11.4.1).

Further research will probably expand even more this potential: some possibilities include combining data from different sources (other social media, but also review platforms such as TripAdvisor), combining different datasets (not only images, but also text or music), integrating algorithms for sentiment analysis, or working at greater scales (metropolitan areas, provinces or regions).

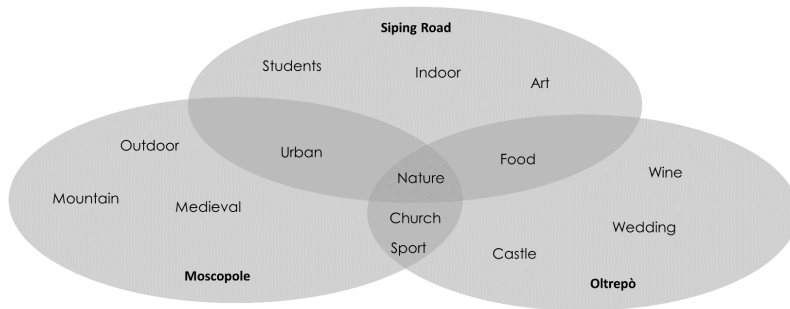


Figure 8. The use of Object Recognition algorithms recognises diversity: the graph shows the main semantic fields associated with each case study.

## References

- Auxier, B., & Anderson, M. (2021, April 7). *Social Media Use in 2021*. Pew Research Center. Retrieved November 17, 2021, from <https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/>.
- De Nadai, M., Dragicevic, S., Hidalgo, C.A., Vieriu, R. L., Naik, N., Sebe, N., Zen, G., Caraviello, M., & Lepri, B. (2016, October). Are Safer Looking Neighborhoods More Lively? A Multimodal Investigation into Urban Life. In *MM '16: Proceedings of the 24th ACM International Conference on Multimedia* (pp. 1127-1135). <https://doi.org/10.1145/2964284.2964312>.
- Girshick, R.B., Donahue, J., Darrell, T., & Malik, J. (2014). Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. *2014 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 580-587). Retrieved February 4, 2022, from <https://arxiv.org/pdf/1311.2524.pdf>.
- He, K., Zhang, X., Ren, S., & Sun, J. (2014) Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. In *Computer Vision – ECCV 2014. Lecture Notes in Computer Science*, vol. 8691. Springer, Cham. [https://doi.org/10.1007/978-3-319-10578-9\\_23](https://doi.org/10.1007/978-3-319-10578-9_23)
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C., & Berg, A. (2016). SSD: Single Shot Multi-Box Detector. In *Computer Vision – ECCV 2014. Lecture Notes in Computer Science*, vol. 9905. Springer, Cham. [https://doi.org/10.1007/978-3-319-46448-0\\_2](https://doi.org/10.1007/978-3-319-46448-0_2)
- Lowe, D. G. (1999, September). Object recognition from local scale-invariant features. In *Proceedings of the Seventh IEE International Conference on Computer Vision*. <https://doi.org/10.1109/ICCV.1999.790410>.
- Lynch, K. (1960). *The Image of the City*. MIT Press.
- Salasses, P., Schechter, K., & Hidalgo, C. A. (2013, July 24). The Collaborative Image of the City: Mapping the Inequality of Urban Perception. *PLOS ONE*, 10(3), <https://doi.org/10.1371/journal.pone.0119352>.
- United Nations (n.d.). *Goal 11*. United Nations Department of Social and Economic Affairs. Retrieved February 4, 2022, from <https://sdgs.un.org/goals/goal11>.
- United Nations General Assembly (2015, September 25). *Transforming our world: the 2030 Agenda for Sustainable Development. Resolution 70/1*, UNGAOR, 70th Sess, UN Doc A/RES/70/1. Retrieved from [undocs.org/A/RES/70/1](https://undocs.org/A/RES/70/1).
- Zhou, B., Liu, L., Oliva, A., & Torralba, A. (2014). Recognizing City Identity via Attribute Analysis of Geo-tagged Images. In *European Conference of Computer Vision 2014* (pp. 519-534). [https://doi.org/10.1007/978-3-319-10578-9\\_34](https://doi.org/10.1007/978-3-319-10578-9_34).

# DEEP ARCHITECTURAL ARCHIVING (DAA)

*Towards a Machine Understanding of Architectural Form*

FREDERICK CHANDO KIM<sup>1</sup> and JEFFREY HUANG<sup>2</sup>

<sup>1,2</sup> *École Polytechnique Fédérale de Lausanne (EPFL)*

<sup>1</sup>*frederick.kim@epfl.ch, 0000-0001-9311-5322*

<sup>2</sup>*jeffrey.huang@epfl.ch, 0000-0003-0501-2791*

**Abstract.** With the ‘digital turn’, machines now have the intrinsic capacity to learn from big data in order to understand the intricacies of architectural form. This paper explores the research question: how can architectural form become machine computable? The research objective is to develop “Deep Architectural Archiving” (DAA), a new method devised to address this question. DAA consists of the combination of four distinct steps: (1) Data mining, (2) 3D Point cloud extraction, (3) Deep form learning, as well as (4) Form mapping and clustering. The paper discusses the DAA method using an extensive dataset of architecture competitions in Switzerland (with over 360+ architectural projects) as a case study resource. Machines learn the particularities of forms using ‘architectural’ point clouds as an opportune machine-learnable format. The result of this procedure is a multidimensional, spatialized, and machine-enabled clustering of forms that allows for the visualization of comparative relationships among form-correlated datasets that exceeds what the human eye can generally perceive. Such work is necessary to create a dedicated digital archive for enhancing the formal knowledge of architecture and enabling a better understanding of innovation, both of which provide architects a basis for developing effective architectural form in a post-carbon world.

**Keywords.** Artificial Intelligence; Deep Learning; Architectural Form; Architectural Competitions; Architectural Archive; 3D Dataset; SDG 11.

## 1. Introduction

The roles attributed to form and the attendant positions within the discipline of architecture have undergone notable transitions that, in effect, reframe the very status of form itself (Eisenman, 1999). The focus on the genesis of form shifted over time as it was overshadowed by various social, cultural, ecological, or political influences on architecture. However, with recent developments of deep learning via neural networks, machines now have the capacity to learn from unlimited volumes of data, thereby lending novel significance to the issue of form-making in design disciplines, while also

enabling architects to continue their disciplinary focus on form, albeit in the still-emergent interface between human and machine (Huang et al. 2021). This research explores the following overarching research question: how can architectural form become machine computable?

The transition to the machine age allows form to operate transversally in the environment where it embodies "all forms of ecology together, whether environmental, mental, or social" (Moussavi & Lopez, 2009). However, even with such innovative thinking, there are still limitations in developing applications of 3D machine learning (ML) models because of the complexity that arises from higher dimension data and the flexibility that decreases with different representations of 3D data such as point clouds, wireframes, voxels, and meshes (Zhang & Huang, 2021). Furthermore, the inconsistencies in methods, details, and resolution of data produced by different individuals also hinder the gathering of consistent 3D data required to compile a cohesive dataset (Stoter et al., 2020). Finally, compared to readily available 3D objects with distinct shapes, such as chairs or sofas, that have been widely used to develop the 3D ML techniques buildings have considerably more formal variables that must be quantified and categorized for archiving purposes. Moreover, such variables demand even greater efforts in compiling the respective architectural datasets.

To address such limitations, this paper introduces a new technique for 3D deep learning of architectural forms: The Deep Architectural Archiving or DAA. DAA employs four steps: (1) Data mining, (2) 3D Point cloud extraction, (3) Deep form learning, and (4) Form mapping and clustering. The paper discusses the procedure in detail and applies the technique to a competition dataset as a case study resource. The application of DAA to a competition dataset can contribute to the structuring of culture and building knowledge in a materialistic and intellectual way for the edification cultural heritage (Chupin et al., 2015). In addition, the dataset for this research is an archive of architectural competitions in Switzerland from 2009 to 2021 that consists of unexploited knowledge of formal innovation, much of which is often undervalued, if not ignored outright. This repertoire of architectural propositions provides the necessary and extensive data for closely examining understudied relationships, distinguishing morphological patterns between architectural forms across different competitions using deep learning, and creating an archive of architectural competition comprising a 3D dataset using ML as a formal indexing strategy.

## 2. Background

Due to the inherent complexity of 3D information and the computational power needed to learn 3D datasets, 3D machine learning techniques require the development of creative methods capable of achieving an appropriate format for the machine readability of 3D architecture. Several methods have been developed for machines to learn from different data formats such as multiple 2D images, voxels, point clouds, or wireframes.

Using a dataset of 2D images, Steinfeld et al. proposed a framework for how 2D ML can be synthesized with 3D generative architectural design (Steinfeld et al. 2019). The volumes of gabled bay houses are translated into multi-view, tiled projection

images to feed into Generative Adversarial Networks (GAN), which generate a composition of elevations that can be constructed into a new form. Similarly, Zhang and Huang also created new forms from different building styles by using 2D images, a sequence of section cuts of 3D buildings, and the multiple layers of training networks such as StyleGAN, Waifu2X, and Pix2Pix (Zhang and Huang, 2021). Newton experimented with GANs using 3D voxels of massing models in downtown NYC by showcasing the results of 3-D IWGAN for encoding and generating new building forms (Newton 2019). In addition, Rodríguez et al. proposed using a 3D grid of wireframes of buildings and variational autoencoder (VAE) (Rodríguez et al. 2020).

Inasmuch as the previous approaches have demonstrated the efficacy of creative methods utilizing different formats of data with 3D ML algorithms for the field of architecture, the current research will deploy an unsupervised machine learning model called 3D Adversarial Autoencoder (3dAAE) using point clouds (Zamorski et al. 2020). The autoencoders are used for a discriminative task, where inputs are data and outputs are both latent points and reconstructed data, which allows for classification and representation learning for indexing and clustering. The model accepts direct input and generation of 3D point clouds composed of 2048 points for learning architectural form. Finally, the current research will be using the discriminative tasks of 3dAAE by using 3D architectural data and different point cloud strategies to develop a method for creating a machine-learned archive of architectural forms.

### 3. The DAA Method

The DAA method explores a framework of unsupervised machine learning of architectural forms using optimized point cloud representations to better understand the formal relationships among the 3D datasets. Figure 1 demonstrates a diagram of the proposed methodology, which includes (1) collecting data to create a 3D dataset of architectural forms, (2) translating the representation of data into machine learnable formats, (3) experimenting with ML to get the latent representation for organizing data, and (4) visualizing the training results using ML that allows humans to observe organizational capacity.

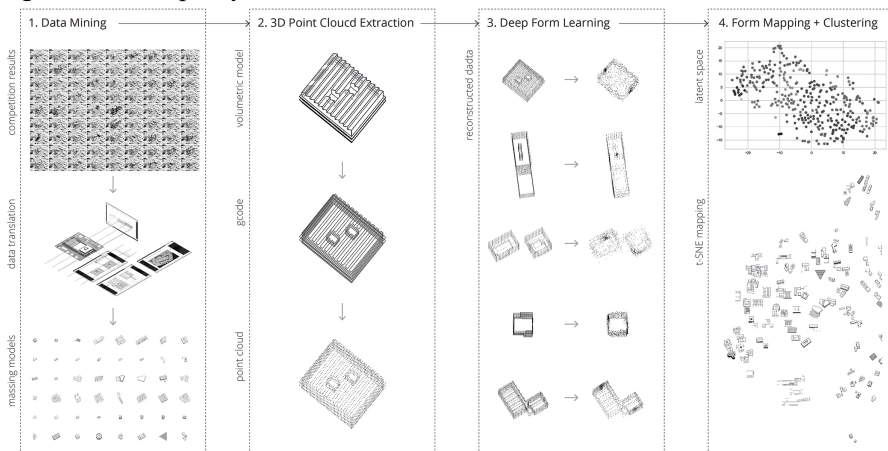


Figure 1. DAA, the proposed method for machine learning of architectural forms.

### 3.1. DATA MINING

The research utilizes the results of the well-organized Swiss architecture competition system as the source of the data. Specifically, the distinctively abstract white plaster massing models, required within the detailed procedural guidelines by the Swiss Society of Engineers and Architects (SIA), provide a key reference point and the primary source of data for the current research on form as they display essential and unadorned qualities of an architectural form in a given context, while also underscoring the importance attributed to form as an autonomous parameter during the evaluation process. Jury reports containing site plans, floor plans, sections, elevation drawings, and model photos from the web-based competition databases of Konkurado and Espazium are collected to reconstruct the white massing models in digital format. The 3D volumes are manually modelled to maintain accuracy and achieve structural consistency of the 3D models. Using Rhino 3D, each 3D volume was generated via an algorithmic approach. The first step was to scale the plans, sections, and elevations to 1:1 and position them in 3D space. The second step was to operate a series of extrusion and Boolean operations for constructing the volumes. Finally, in the third step, photographs of white massing models were used to verify any missing information from the drawings or complete the 3D modelling. Initially modelled using NURBS geometry, the models can be subsequently transformed into different data formats that are required by various ML algorithms.

The 3D models are developed with respect to their sizes in real-world dimensions and orientation in context. The results from 63 competitions of schools that ranged from 3 to 10 in project ranking are translated into 366 3D volumes from the past fifteen years. The 3D volumes are then used to generate a cohesive dataset for exploring architectural forms. Figure 2 illustrates the dataset in order of ranking on the vertical axis and chronological order on the horizontal axis.

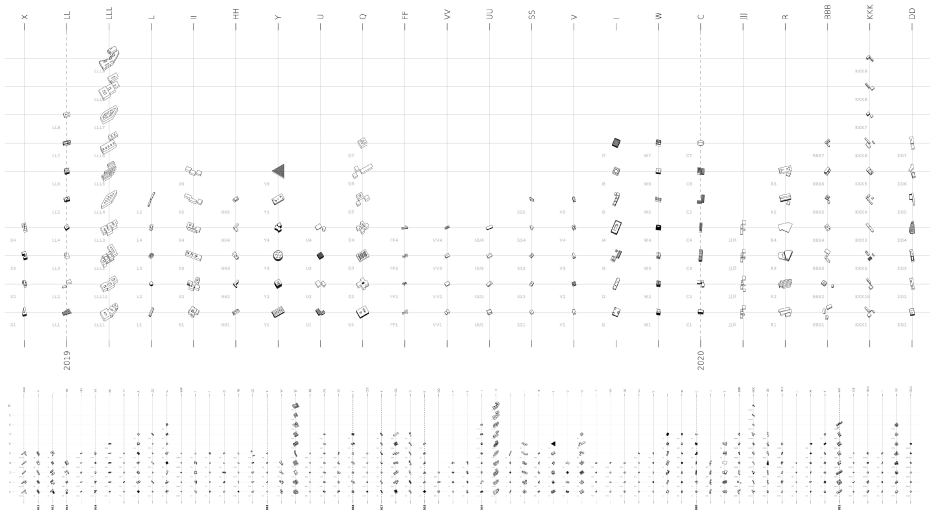
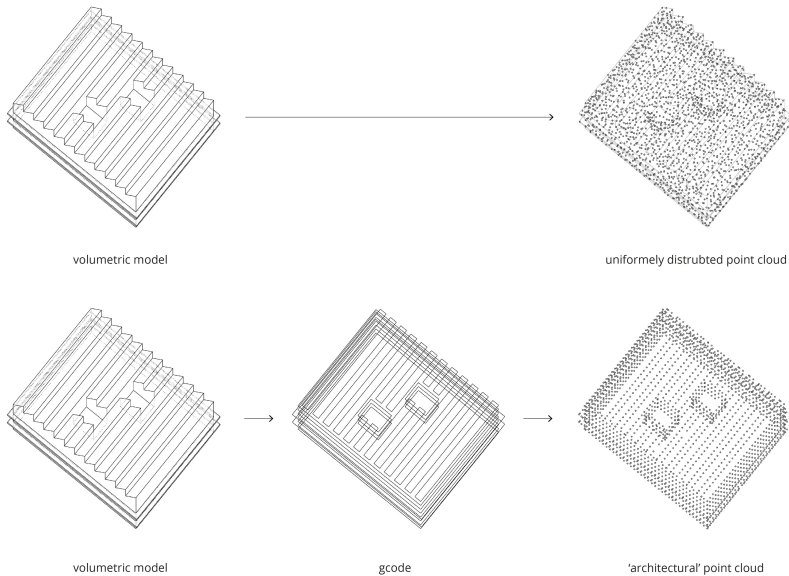


Figure 2. 3D dataset of 366 models from 63 school competitions in Switzerland arranged in chronological order on the x-axis and rankings on the y-axis.

### 3.2. 3D POINT CLOUD EXTRACTION

This step entails the preparation of a dataset, with 3D volumes transformed into an optimal representation of the architectural form using ‘point clouds’ as a machine learnable data format. Rather than a typical uniformly distributed point cloud, an ‘architecturalized’ point cloud is considered by optimally distributing 2048 points to represent the 3D architectural information for ML training. ‘Architectural’ point clouds attempt to structure the points in a meaningful architectural representation that both machines and humans can understand. As seen in Figure 3, inspired by the idea of the ‘G-codes’ (or geometric codes), a tool path produced by slicing a model for 3D printing is used to distribute the points. The points are evenly spaced out along the lines of the slicing paths to create consistency in the structure of point clouds. Using the G-code to extract the points allows different data formats of 3D volumes to be translated into a consistent dataset. The technique also normalizes the location of the points along the z-axis for each point cloud and reduces a degree of freedom for machine learning.



*Figure 3. Data translation from volumetric representation to uniformly distributed point cloud on the top. Data translation through G-code representation to achieve architectural point cloud on the bottom.*

Since there is no architectural scale in ML, the question of how 3D models are generalized from different sizes of buildings must be addressed. To develop the dataset, each model was scaled relative to the largest model to fit into a bounding box of  $1 \times 1 \times 1$  unit. In addition, a custom Grasshopper definition was developed using Rhino 3D to automate the process of transforming the input geometry into ‘architectural’ point clouds or uniformly distributed points. For the ML experiment, two-point cloud generation methods were used for developing two datasets: standard uniformly distributed point clouds and ‘architectural’ point clouds.

### 3.3. DEEP FORM LEARNING

Using the two datasets, the 3dAAE model, an unsupervised machine learning algorithm (Zamorski et al. 2020), is trained with 15,000 epochs, a learning rate of 0.0001, a batch size of 5, and the latent space dimension of 128. The training results of encoding and decoding 3D data are evaluated qualitatively by comparing the reconstruction of the point cloud to its sampled point cloud from the original architectural form. The success of the results is demonstrated in the regeneration of the point clouds that are remarkably similar to the original architectural form. Figure 4 shows the progression of the reconstruction of the point clouds at different iterations. In this example, the network shows successful learning of data composed of two rectangular volumes. As seen at 1,000 epochs, the point clouds split in two. After 10,000 epochs, the reconstructed data show stability in their point-cloud structure.

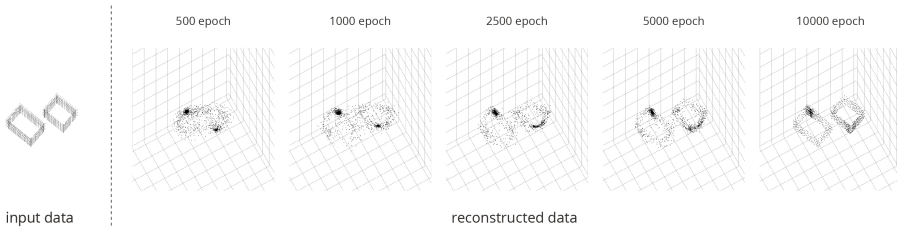


Figure 4. Sample of training progress at different epochs.

Fig. 5 shows notable differences between using uniformly distributed point clouds and architectural point clouds in ML training. Reconstructed data using uniform point clouds show randomly distributed points with a single region of a dense cluster of points within each volume. Although overall shapes are captured from the input data, there is still a lack of architectural details. As for the architectural point clouds, reconstructed data captures the general shapes and formal characteristics of architecture, such as the details of the surfaces along the volume. Compared to the single region of concentrated points when using uniformly distributed point clouds, reconstructed architectural point clouds resemble a magnetic field with the interplay of two concentrated regions of points. Figure 5 illustrates the results of comparing reconstructed and input data of uniformly distributed point clouds on the left and architectural point clouds on right at 10,000 epochs. As seen in the input data of architectural point clouds, the horizontal planes do not contain the points due to their z-direction on the slicing of volumes. This leads to a higher concentration of points around certain elements in the buildings for a clear representation of the reconstructed point clouds. For both point cloud structures, the smaller models captured greater levels of details than the larger ones due to their compactness of the density of points caused by the limited number of points for representing different sizes of 3D models.



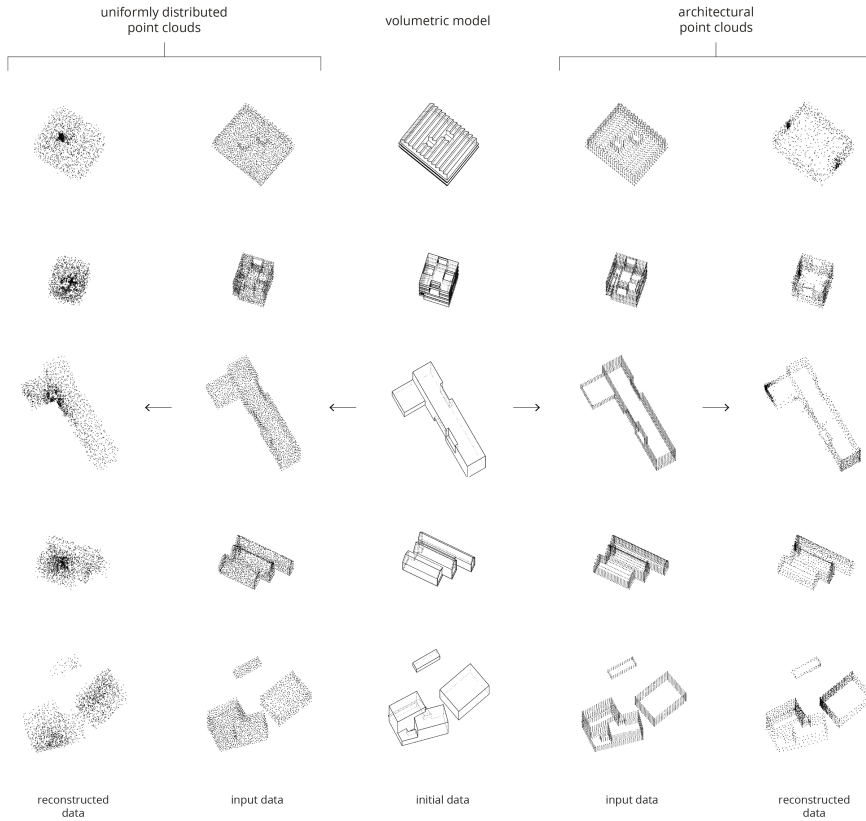


Figure 5. Comparison of reconstructed and input data of uniformly distributed point clouds on the left and architectural point cloud on right at 10,000 epochs.

### 3.4. FORM MAPPING AND CLUSTERING

During the machine learning training, the data is compressed to 128-dimensional encoded representations and visualized into 2D space using t-Distributed Stochastic Neighbor Embedding (t-SNE) (Maaten and Hinton 2008). The translation allows for the interpretation of the training results, insofar as the original 3D models are projected according to their encodings in the latent space. The spatialized arrangement of the machine-read relationships offers a more fundamental understanding of correlations across different formal groups expressed by the affinities among the forms. In this process, the t-SNE technique is used to map out the volumetric models at the encoded data points at every 1,000 epochs to evaluate the performance of ML using a qualitative evaluation of human interpretation by observing the placements of architectural forms. Rather than the reconstructed point clouds, the original volumetric models are used as the visual representations of the encoded data points. For example, Figure 6 shows the results of t-SNE plots at 9,000 epochs of previously mentioned experiments using two different point cloud datasets for ML training.

As for comparing the results between two different point cloud structures, the example shows subtle differences in the overall mapping of the 3D forms. The performance of an 'architectural' point cloud is compatible with the performance of standard point clouds. Furthermore, the positioning of different clusters of formal groups in the architecturalized point clouds even shows smooth transition in the arrangement of the volumes. For example, the clusters of linear bar buildings in different directions appear to be closer to each other in the mapping of volumes. For further experimentation, one can also imagine orienting the 3D models in the same direction to exclusively examine shapes as the main comparative feature. However, t-SNE showed a pitfall after 9,000 epochs. It showed a collapsing behaviour with merged clusters, whereby it was no longer possible to find a correlation between the encoded models (Wattenberg *et al.* 2016).



Figure 6. t-SNE plots of 3D models showing training results from two datasets at 9,000 epochs.

#### 4. Results and Discussion

DAA is proposed as a viable approach for making architectural form machine-computable and for enabling humans to analyse and interpret the machine understanding and learning of 3D architectural forms. Since it is not quantitatively evident how machines understand forms, the visualization of latent space for the qualitative measure is essential for examining the performance of encoding and decoding processes of networks at different stages of training. As shown in Figure 7, it is clear that the network understands the formal properties of the 3D models at multiple scales. In the overall view of the latent space, the models are positioned in a gradient from large to small buildings. The orientation of the buildings is reflected in the mapping, which creates a smooth change in direction between the models like a vector field. In addition, different shapes are clustered together. For example, linear buildings are placed in the perimeter of the latent space, and cubic-shaped buildings are placed towards the central area. Zoom-ins on the right side of Figure 7 reveal details such as voids, courtyards, the complexity of roof surfaces. Finally, facades are reflected in the proximities between the 3D models.

While the proposed method is promising, there are limitations with the DAA method worth noting: (1) time-intensive manual labour in translating 2D architectural drawings into 3D data. (2) a lack of data that makes it difficult to identify the subtle differences of the mapping results of learned forms. To address these limitations, further experimentation will be undertaken within each step of the DAA method. For example, an increase in the number of 3D datasets from the Swiss competition archives is currently in progress. Optimization of the number of points and structure of the point clouds will also be investigated. In addition, different hyperparameters, such as the number of dimensions, learning rates, and batch sizes, can be further explored to optimize the training performance. Finally, clustering algorithms, such as K-means could be used to identify machine categorizations of encoded models.



*Figure 7. t-SNE plot of machine-learned forms using 'architectural' point clouds at 9,000 epochs with close-ups of clustering of different formal groups.*

## 5. Conclusion and Future Work

This research develops a 3D dataset drawn from the architecture competitions in Switzerland to examine how machines can understand the particularities of architectural forms and evaluate comparative relationships among form-correlated datasets that far exceed human comprehension in their scale and complexity. Using DAA, the findings show promising results in reconstructing architectural forms by optimizing data structure for the training of networks using a relatively small dataset composed of uniform and 'architectural' point clouds. The visualizations of the results with t-SNE mapping demonstrate how formal properties such as geometrical features, scale, and orientation can be identified and how hidden connections can be discovered.

The trans-scalar view of the forms from different competitions enables a more precise examination of understudied relationships among architectural forms.

In view of future applications, DDA aspires to create a novel multi-dimensional and dynamic digital archive that facilitates a machine-enabled reading of architectural forms. It aims to advance our understanding of the inherent morphological patterns within the collection of 3D data and offer insights to architects in their formal exploration throughout the design process. Such an archive could index not only the formal properties but also preserve building data such as program, area calculations, cost estimates, energy efficiency, time factor, and jury rankings from design phases contributing towards sustainable building culture. It is also foreseeable that such knowledge could enable a detailed understanding of the distinct contextual, ecological, and cultural relevance of specific architectural forms in the post-carbon world.

## References

- Chupin, J.-P., Cucuzzella, C., & Helal, B. (2015). *Architecture Competitions and the Production of Culture, Quality and Knowledge: An International Inquiry*. Potential Architecture Books Inc.
- Espazium. *Competitions*. Retrieved December 2, 2021, from <https://competitions.espazium.ch>.
- Huang, J., Johanes, M., Kim, F. C., Doumpiotti, C., & Holz, G.-C. (2021). On GANs, NLP and Architecture: Combining Human and Machine Intelligences for the Generation and Evaluation of Meaningful Designs. *Technology|Architecture + Design*, 5(2), 207–224. <https://doi.org/10.1080/24751448.2021.1967060>
- Konkurado. *Web of Design Competitions*. Retrieved December 2, 2021, from <http://konkurado.ch>.
- Maaten, L. van der, & Hinton, G. (2008). Visualizing Data using t-SNE. *Journal of Machine Learning Research*, 9(86), 2579–2605.
- Moussavi, F., & Lopez, D. (2009). *The Function of Form*. Actar.
- Newton, D. (2019). Generative Deep Learning in Architectural Design. *Technology | Architecture + Design*, 3(2), 176–189. <https://doi.org/10.1080/24751448.2019.1640536>
- Rodríguez, J. de M., Villafañe, M. E., Piškorec, L., & Caparrini, F. S. (2020). Generation of geometric interpolations of building types with deep variational autoencoders. *Design Science*, 6. <https://doi.org/10.1017/dsj.2020.31>
- Steinfeld, K., Park, K. S., Menges, A., & Walker, S. (2019). Fresh Eyes: A Framework for the Application of Machine Learning to Generative Architectural Design, and a Report of Activities. *Smartgeometry 2018*. [https://doi.org/10.1007/978-981-13-8410-3\\_3](https://doi.org/10.1007/978-981-13-8410-3_3)
- Stoter, J. E., Arroyo Ohori, G. a. K., Dukai, B., Labetski, A., Kavisha, K., Vitalis, S., & Ledoux, H. (2020). State of the Art in 3D City Modelling: Six Challenges Facing 3D Data as a Platform. *GIM International: The Worldwide Magazine for Geomatics*, 34.
- Wattenberg, M., Viégas, F., & Johnson, I. (2016). How to Use t-SNE Effectively. *Distill*, 1(10), e2. <https://doi.org/10.23915/distill.00002>
- Zamorski, M., Zięba, M., Klukowski, P., Nowak, R., Kurach, K., Stokowiec, W., & Trzcinski, T. (2020). Adversarial autoencoders for compact representations of 3D point clouds. *Computer Vision and Image Understanding*, 193, 102921. <https://doi.org/10.1016/j.cviu.2020.102921>
- Zhang, H., & Huang, Y. (2021). Machine Learning Aided 2D-3D Architectural Form Finding at High Resolution. In P. F. Yuan, J. Yao, C. Yan, X. Wang, & N. Leach (Eds.), *Proceedings of the 2020 DigitalFUTURES* (pp. 159–168). Springer. [https://doi.org/10.1007/978-981-33-4400-6\\_15](https://doi.org/10.1007/978-981-33-4400-6_15)

# WEB-BASED THREE-DIMENSIONAL AUGMENTED REALITY OF DIGITAL HERITAGE FOR NIGHTTIME EXPERIENCE

TOMOHIRO FUKUDA<sup>1</sup>, SHIHO NAGAMACHI<sup>2</sup>, HOKI NAKA-MURA<sup>3</sup>, YUJI YAMAUCHI<sup>4</sup>, NAO ITO<sup>5</sup> and SHUNTA SHIMIZU<sup>6</sup>

<sup>1</sup>*Osaka University*

<sup>2,3</sup>*LEM Design Studio Co., Ltd.*

<sup>4,5</sup>*ALAKI Co., Ltd.*

<sup>6</sup>*FORUM8 Co., Ltd.*

<sup>1</sup>*fukuda.tomohiro.see.eng@osaka-u.ac.jp, 0000-0002-4271-4445*

<sup>2,3</sup>*{shihonagamachi|hokinakamura}@gmail.com*

<sup>4,5</sup>*{yamauchi|itoh}@alaki.jp, 0000-0002-8068-4611, 0000-0002-6258-5214*

<sup>6</sup>*simizu@forum8.co.jp*

**Abstract.** Digital heritage is a sustainable medium that allows people to understand the historical shape and context of cities and architecture, leading to visions for the future. Opportunities for the public to experience life-size representations of digital heritage in three-dimensional augmented reality (3D-AR) at outdoor sites are still limited, especially at night. Therefore, the objective of this study is to develop a web-based 3D-AR method to digitally reconstruct a heritage site. A prototype system was developed using the five-storey pagoda of Tango Kokubunji Temple, which was built around 1330 AD and later destroyed, as a digital heritage reconstruction. An interactive initial positioning method was developed to display the five-storey pagoda on real historical foundation stones by tapping a crosshair button, under the condition that the artificial lighting is insufficient at night and the distance between the viewpoint and the 3D model of the pagoda is large. Combining the lighting effects of the real and virtual worlds at night was also demonstrated. We held an event where the general public could experience 3D-AR on their own mobile devices, and conducted a usability evaluation to verify the system.

**Keywords.** Digital Heritage; Digital Restoration; Augmented Reality (AR); Web System; Lighting Design; Virtual and Real Worlds; SDG 4; SDG 8.

## 1. Introduction

Harnessing the power of heritage will accelerate the achievement of Sustainable Development Goals (SDGs) to achieve the well-being of people (SDGs 1, 2, 3, 4, 5, 6 and 11), the well-being of the planet (SDGs 6, 7, 11, 13, 14 and 15), the prosperity of communities (SDGs 5, 8, 9, 11, 12 and 14), peace within and among societies (SDGs

10, 11 and 16), and the creation of partnerships (SDGs 11 and 17) (Labadi et al., 2021). Digital heritage is a sustainable medium that allows people to understand the shape and context of cities and architecture, leading to visions for the future (UNESCO, 2009). Historical buildings that no longer exist or are severely weathered can be restored as three-dimensional (3D) digital models and brought back to life by superimposing models onto the real world using augmented reality (AR) technology as “3D-AR” experiences (Challenor and Ma, 2019). In addition to experiencing digital heritage representations in a museum (White et al., 2007), AR allows users to better imagine buildings that existed previously at outdoor sites (Lu et al., 2010). While the use of virtual reality (VR) has been reported, which includes a complete 3D computer graphics representation of a non-existent past nightscape (Fukuda et al., 2015), an AR application that can provide an outdoor, full-scale experience of a nightscape has not yet been reported.

AR systems have long required specialised hardware and software, which prevents many individuals from using their own mobile devices to experience AR. Recently, frameworks such as ARCore and ARKit have been developed to handle AR on mobile devices. However, because of model dependency, ARCore and ARKit cannot be used on some devices and further development for both ARCore and ARKit is required, which increases costs.

Recently, these problems have been alleviated by using an AR environment that is independent of the operating system and runs in a web browser via standard scripts (hereafter, web-based AR). Although the use of web-based AR to digitally restore heritage sites has been reported (Abergel et al., 2019), we are not aware of any reports of web-based AR that can be used outdoors at night.

The objective of this study is to propose a method for experiencing digital heritage in web-based 3D-AR at night. A prototype system was developed using the five-storey pagoda of Tango Kokubunji Temple (hereafter, the five-storey pagoda), which was built around 1330 AD and later destroyed, as a digital heritage reconstruction. An interactive initial positioning method is proposed to display the five-storey pagoda on real historical foundation stones by tapping a crosshair button, under the condition that the artificial lighting is insufficient at night and the distance between the viewpoint and the 3D model of the pagoda is large. Combining the lighting effects of the real and virtual worlds at night is also demonstrated. The system was verified through an event during which the public could experience 3D-AR using their own mobile devices.

## 2. Project Challenges

The five-storey pagoda digital heritage reconstruction project had to overcome the following challenges:

- Users need to experience web-based 3D-AR at night. As the digital heritage site is located in a rural area with inadequate artificial lighting at night, the 3D-AR application needs to have a tracking process that keeps the 3D virtual model in the right position and orientation in the real world (live image), rather than simply displaying it. The conventional and simple method of tracking geometric markers while detecting them with a web camera is impractical owing to insufficient illumination.

- The distance from the 3D-AR viewpoint to the digital heritage display (the digitally reconstructed five-storey pagoda) is large. Mainly for safety reasons, users will experience the 3D-AR application on a raised platform about 50 m away from the pagoda site while facing the site and keeping their smartphone in front of them.
- The 3D-AR application needs to be artistic. The AR content must be attractive and interactive for the users visiting the site so they can experience what the site looked like when the five-storey pagoda was built.

Figure 1 shows a plan of the project, the viewpoint and a photograph of the foundation stones of the five-storey pagoda.



Figure 1. Tango Kokubunji five-storey pagoda AR Revival Project: (left) plan view, (top right) view of the pagoda remains from the AR viewpoint, (bottom right) closeup view of the foundation stones

### 3. Development of a Web-based 3D-AR Digital Heritage for Nighttime Experience

#### 3.1. 3D MODELLING OF THE FIVE-STOREY PAGODA

There are no accurate drawings of the five-storey pagoda to be brought back to life virtually. The five-storey pagoda depicted in the ink painting "View of Ama-no-Hashidate" by Sesshu, a painter active around 1330 AD, provides helpful clues, but it contains insufficient information to create a complete 3D model.

Therefore, after discussions with the museum curator, we referred to other five-storey pagodas that were built in the same period. The design and the successive diminution with height of the five-storey pagoda differ according to the date of construction. Here, "diminution" refers to the width and eaves of a pagoda becoming smaller with each successive storey, and "successive diminution" refers to the ratio of the width of the top (fifth) storey to the first storey. The Myooin five-storey pagoda built in 1348, which was built in the same period as the Tango Kokubunji five-storey pagoda, has a successive diminution of 0.714.

Next, the distance between the two end pillars of the ground floor of the five-storey pagoda of Tango Kokubunji was 6.54 m, which is about 1.5 times longer than that of the five-storey pagoda of Myooin (4.36 m). If the successive diminution is the same as that of the Myooin five-storey pagoda (0.714), the total height of the Myooin five-

storey pagoda is 29.1 m, and the total height of the Tango Kokubunji five-storey pagoda is estimated to have been 43.7 m.

This information was then used as the basis for 3D modelling. It was necessary to set the upper limit of the number of polygons and the texture image to execute real-time rendering as a web-based 3D-AR. However, when the lighting effect is rendered for the 3D model of the five-storey pagoda, the architectural components that appear on the exterior affect the expression of shading. Therefore, the architectural elements characteristic of the exterior were modelled in 3D, taking into account the upper limit of the number of polygons (Table 1 and Figure 2). The file format used was the GL Transmission Format (.gltf or .glb).

Table 1 Upper limit and used value of the number of polygons and texture

Characteristic	Upper limit	Used value
Number of polygons	100,000	99,552
Texture (pixels)	2048 × 2048	1024 × 1024



Figure 2. Created 3D model of the five-storey pagoda

### 3.2. DEVELOPMENT OF THE WEB-BASED 3D-AR SYSTEM

A web-based 3D-AR system was developed by adapting the A-frame framework. Figure 3 shows the directory structure of the developed web application. The main development items are described below.

- Permission to acquire mobile device information: After loading the URL to access the AR application, permission to acquire the camera image and gyro sensor/accelerometer attached to the mobile device is required.
- Initial positioning: The 3D model of the five-storey pagoda is displayed. To anchor the display to the foundation stones in the real world, first, crosshairs are displayed in the centre of the screen. The user places the crosshairs on the foundation stones and interactively taps the crosshairs to display the 3D model of the pagoda (initial positioning). After the initial positioning is complete, the 3D model is re-rendered to correctly draw it on the foundation stones when the mobile device is rotated



around the X, Y, and Z axes.

- **Lighting installation in the virtual space:** After the initial positioning, spotlighting effects are performed on the 3D model of the five-storey pagoda. Next, the 3D model is illuminated from the foot of the pagoda. Then, a lighting effect is supplied from the roof of the pagoda. Lighting is then supplied from the inside of the pagoda. Finally, an entertaining lighting performance is performed with the 3D model. The lighting animation lasts 3 minutes and 10 seconds (Figure 4). These lighting effects are implemented by adapting a shader program.
- **Background music (BGM):** After the initial positioning is complete, BGM is played.

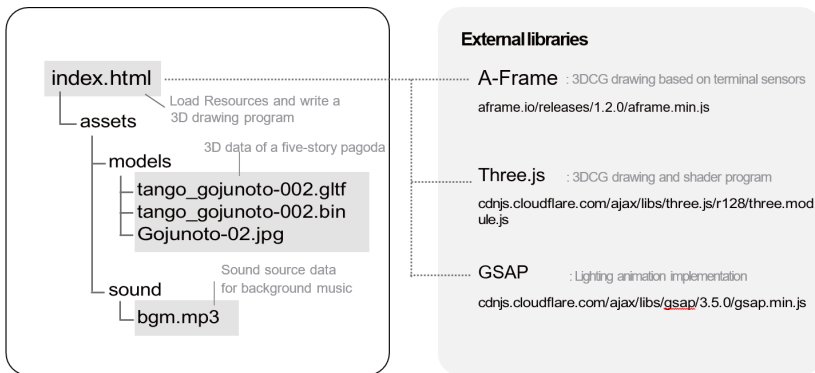


Figure 3. Directory structure of the developed web application

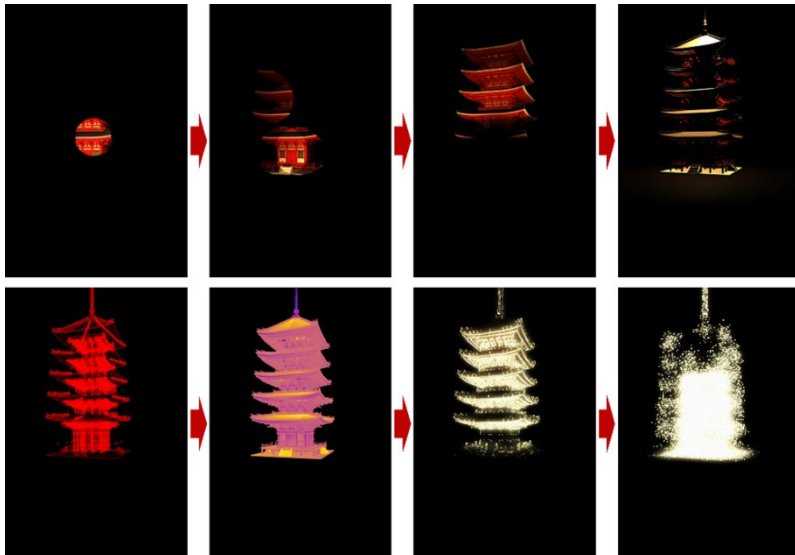


Figure 4. Screenshots of animation of lighting effects in the virtual space

### 3.3. LIGHTING DESIGN IN THE REAL WORLD

Because it is dark at night in the target area (almost 0 lx) and it is difficult to discern the foundation stones from the viewpoint, a narrow-angle light fixture was set up in the middle of the foundation stones. In the initial positioning described in section 3.2, this light is used as a target and users overlay the crosshairs displayed on the screen on this target to display the 3D model of the five-storey pagoda.

To create the illusion that the lighting environment of the real world and the virtual world are merged on the AR screen, the grassland and the foundation stones of the five-storey pagoda were illuminated from near the viewpoint (Figure 5). Lighting was also provided for users to approach the viewpoint.

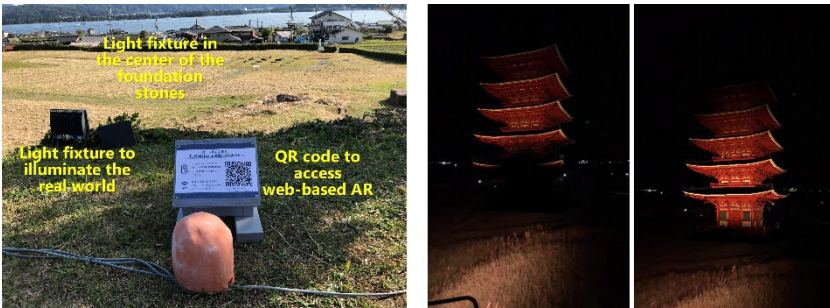


Figure 5. (left) QR code to access AR application and light fixtures, (right) AR screenshots showing the fusion of real-world and virtual world lighting

## 4. Verification Experiment and Results

### 4.1. OVERVIEW OF THE EXPERIMENT

The verification experiment took place for 29 days, from 31 October to 28 November 2020 (Figure 6). During this period, QR codes placed at the site contained the URL to access the web-based AR application described in section 3. Users visited the Tango Kokubunji five-storey pagoda site and scanned the QR code with the camera of their mobile device (smartphone or tablet) to access the AR application. The user then experienced an AR presentation lasting 3 minutes and 10 seconds by interactively tapping the crosshair button displayed in the middle of the screen, after overlaying the



Figure 6. Photos of the experiment

crosshairs on the light placed in the middle of the foundation stones. During the AR application experience, users were able to move and rotate their mobile devices freely.

At the end of the AR application, a URL was displayed, which directed users to a survey form that users could voluntarily access and complete usability evaluation.

#### 4.2. RESULTS

The attributes of the 73 participants who responded to the questionnaire were as follows:

- The age ranges were as follows: 1 (1.4%) younger than 20, 12 (16.4%) in their 20s, 12 (16.1%) in their 30s, 21 (28.8%) in their 40s, 23 (31.5%) in their 50s, and 4 (5.5%) in their 60s. Individuals aged in their 50s and 40s together accounted for more than 60% of the respondents, but a wide range of ages was included.
- The gender was almost equally divided: 37 (50.7%) females and 36 (49.3%) males.
- 21 (28.8%) lived in Miyazu City (where the five-storey pagoda site is located), 27 (37%) lived in Kyoto Prefecture (excluding Miyazu City), and 25 (34.2%) lived in the Kinki region (excluding Kyoto Prefecture). Thus, all Respondents were from relatively nearby areas, likely because it was difficult to travel long distances in Japan owing to the ongoing pandemic at the time of the experiment.
- Only smartphone mobile devices were used, including 63 (86.3%) iPhones and 10 (13.7%) Android phones. No tablet devices were used by any of the respondents.

Figure 7 shows the results of the questionnaire. No significant differences were observed between days.

- In terms of the overall evaluation of the AR application, 68.5% rated it as very excellent and 86.3% rated it highly (Figure 7 (a)).
- As for the length of the AR application, 76.7% answered that it was just right. However, 12.3% rated it as very long (Figure 7 (b)).
- Regarding whether the five-storey pagoda model could be displayed on the foundation stones in the real world (indicated by a light in the centre), 60.6% answered that it could be displayed. On the other hand, 21.1% answered that they could not display it well (Figure 7 (c)).
- 60.9% stated that the lighting in the real world "blended in" with the lighting in the virtual world. In contrast, a small number of respondents answered that the displays were not integrated (Figure 7 (d)).

#### 5. Discussion

We developed a 3D-AR method to display a digital heritage reconstruction of the five-storey pagoda on its foundation stones at night, when artificial lighting was not sufficient, and under outdoor conditions, when the distance between the viewpoint and the 3D model of the five-storey pagoda in Tango Kokubunji Temple was 50 m. To achieve this challenge, we set up a narrow-angle light fixture at the centre of the actual foundation stones and added a crosshair button on the screen of the mobile device,

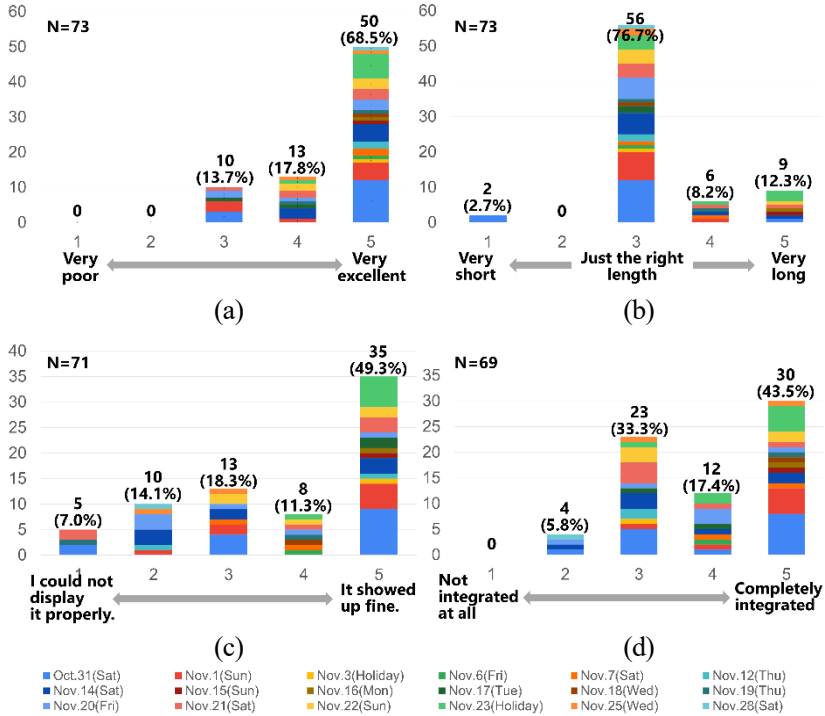


Figure 7. Questionnaire daily results: (a) Overall rating, (b) Length of the AR presentation, (c) Alignment of the five-storey pagoda model display with respect to the foundation stones and (d) Integration of the lighting in the real and virtual worlds

which the user taps after superimposing the crosshairs on the light source to display the five-storey pagoda 3D model interactively. After the initial positioning, the 3D model was re-rendered in real time by using the smartphone's gyro sensor/accelerometer to obtain the position information.

From the results of the usability evaluation (Figure 7 (c)), more than 60% of the respondents were able to display the five-storey pagoda model on the foundation stones in the real world. A possible reason why relatively many respondents reported that they could not display the five-storey pagoda model well on the foundation stones was that other activities were also occurring during the experiment period, and there were more lights in the surrounding environment and from night cruises. This may have made the target light in the centre of the foundation stone relatively less clear.

While users visiting the site virtually experienced the scene at the time when the five-storey pagoda was built, we tried to make users feel the attraction of the 3D digital heritage reconstruction of the pagoda. To achieve this, we created a 3D-AR real-time animation 3 minutes and 10-second long, during which the lighting effects changed successively based on various illumination scenarios, which included not only realistic lighting but also artistic expressions. To create the illusion that the lighting environment of the real world and the virtual world are fused, we illuminated the grassland and the foundation stone of the five-storey pagoda from the viewpoint.

In the results of the usability evaluation (Figure 7 (b)), 76.7% of the respondents answered that the length of the AR animation was just right, which suggests that the application performed well. In addition, as shown in Figure 7 (d), 60.9% of the respondents answered that the lighting of the real world and the virtual world were "integrated", which also suggests satisfactory performance.

Descriptions of the issues we found are as follows:

- There are limits to the number of polygons and texture sizes of 3D models that can be handled by current web-based 3D-AR applications. These limits are below the data-handling capabilities of most commercial architectural and urban design programs. Approaches to solving this problem include reducing the number of polygons while maintaining a realistic representation of the model, relying on the continued improvement of mobile device performance, or rethinking the role of the server and the mobile device (client). One technique for the third approach is to render a large number of 3D models at high speed on the server side and send the result of the rendering to the mobile device as a video.
- In the real world, lighting effects are caused by shading due to the position of light fixtures and the eaves of buildings. Lighting designers and computer graphics creators have had many discussions on how to express the same lighting effects in the virtual world to match those in the real world, but the construction of a shader lighting program is very challenging. The differences between virtual shader lighting effects and real-world lighting effects are still significant, and it is still difficult to express the same architectural lighting effects in the virtual world as in the real world. This is an issue that we discovered once again by merging the lighting effects from the virtual and real worlds.
- Initially, we had planned to load the content of the virtual world's lighting effects into the real world's lighting equipment in real-time, so that the light emitted from the lighting equipment would be completely synchronised with the virtual world's lighting effects, but this was not possible. Therefore, in this experiment, we prepared the content of the real-world light fixtures and the virtual-world lighting effects in advance and represented them on a device screen as if they were synchronised. It is necessary to develop a method for synchronising the lighting environment of the real world and the virtual world in real time.

## 6. Conclusion

Opportunities for the public to experience life-size representations of digital heritage in 3D-AR at outdoor sites are still limited, especially at night. Therefore, the objective of this study was to develop a web-based 3D-AR method to construct a digital heritage for nighttime experience. A prototype system was implemented using the five-storey pagoda of Tango Kokubunji Temple, which was built around 1330 AD and later destroyed, as a digital heritage reconstruction. The system was verified through an event during which the public could experience 3D-AR on a web browser using their own mobile devices.

The contributions of this research are as follows:

- A method was developed to combine 3D-AR and a night-time viewing environment

with inadequate artificial lighting and a distance of 50 m between the AR viewpoint and the imaged object, namely, the five-storey pagoda of the Tango Kokubunji Temple. The method displayed the digital heritage reconstruction of the pagoda interactively on its actual foundation stones in the real world by allowing the user to align the virtual and real scenes by tapping on a crosshair button on the screen.

- The 3D-AR real-time presentation, which was 3 minutes and 10 seconds long, displayed a 3D model of the five-storey pagoda and included artistic lighting effects. A user survey confirmed that the lighting effects in the virtual world and the lights from the light fixtures placed in the real world were combined on the 3D-AR screen.

Future challenges include the exclusion of restrictions on the number of polygons and texture size of 3D models in the web-based 3D-AR application and the synchronisation of lighting effects in the real and virtual worlds.

### Acknowledgements

We would like to thank all the participants for their support in the project, especially the Kyoto by the Sea Destination Management/Marketing Organization and the Kyoto Prefectural Tango Reginal Museum.

### References

- Abergel, V., Jacquot, K., De Luca, L., & Veron, P. (2019). Towards a SLAM-based augmented reality application for the 3D annotation of rock art. *International Journal of Virtual Reality*, 19(2), 17–26. <https://doi.org/10.20870/IJVR.2019.19.2.2910>.
- Challenor, J., & Ma, M. (2019). A review of augmented reality applications for history education and heritage visualisation. *Multimodal Technologies and Interaction*, 3(2). <https://doi.org/10.3390/mti3020039>.
- Fukuda, T., Ban, H., Yagi, K., & Nishiie, J. (2015). Development of high-definition virtual reality for historical architectural and urban digital reconstruction: A case study of Azuchi castle and old castle town in 1581. *Communications in Computer and Information Science*, 527. Springer. [https://doi.org/10.1007/978-3-662-47386-3\\_5](https://doi.org/10.1007/978-3-662-47386-3_5)
- Labadi, S., Gilberto, F., Rosetti, I., Shetabi, L., Tildirim, E. (2021). *Heritage and the sustainable development goals: Policy guidance for heritage and development actors*. ICOMOS. Retrieved November 25, 2021, from [https://www.icomos.org/images/DOCUMENTS/Secretariat/2021/SDG/ICOMOS\\_SDGs\\_Policy\\_Guidance\\_2021.pdf](https://www.icomos.org/images/DOCUMENTS/Secretariat/2021/SDG/ICOMOS_SDGs_Policy_Guidance_2021.pdf).
- Lu, B. V., Kakuta, T., Kawakami, R., Oishi, T., & Ikeuchi, K. (2010). Foreground and shadow occlusion handling for outdoor augmented reality. In *9th IEEE International Symposium on Mixed and Augmented Reality 2010: Science and Technology, ISMAR 2010* (pp. 109–118). <https://doi.org/10.1109/ISMAR.2010.5643558>.
- UNESCO. (2009). *Charter on the Preservation of the Digital Heritage*. CL/3865. Retrieved November 25, 2021, from <https://unesdoc.unesco.org/ark:/48223/pf0000179529.page=2>.
- White, M., Petridis, P., Liarokapis, F., & Plecinckx, D. (2007). Multimodal Mixed Reality Interfaces for Visualizing Digital Heritage. *International Journal of Architectural Computing*, 5(2), 321–337. <https://doi.org/10.1260/1478-0771.5.2.322>.

# ROBOTIC CARVING CRAFT

*Research on the application of robotic carving technology in the inheritance of traditional carving craft*

ZHE GUO<sup>1</sup>, ZIHUAN ZHANG<sup>2</sup> and CE LI<sup>3</sup>

<sup>1,3</sup>*Hefei University of Technology.*

<sup>2</sup>*Wuhan University*

<sup>1</sup>*guogal@hotmail.com, 0000-0002-7660-0622*

<sup>2</sup>*540530969@qq.com, 0000-0002-8300-6525*

<sup>3</sup>*26650299@qq.com, 0000-0002-6290-4297*

**Abstract.** In order to realize the inheritance of handicraft skills via digital fabrication technique, so as to preserve the traditional construction culture, this paper discusses a method of control industrial robot (six-axis KUKA kr-60 robotic arm) simulate carving craftsmen working process and explores the relationship between carving posture and different clay states. This paper starts with discussion with cultural heritage in the background of digital tools application. Next, a method to determine the pose of robotic arm by giving the angle value of the six axis is applied in the subsequent carving experimental research, which can make the robotic arm have a smoother and reasonable motion performance by disable the redundant axis movement of the robotic arm when adjusting those poses. Then, a series of carving experiments has been carried out to explore the connection between robotic movement and carved detail, together with a carving path arrangement method that allow for specific carved lines caused by given axis value. This research shows the possibility to create complex form through defining robot movement, which could fundamentally make robot manufacturing a new formal meaning.

**Keywords.** Clay Carving; Robotic Arm Control; Crafts Inheritance; Form Algorithm; SDG 8.

## 1. Introduction

### 1.1. BACKGROUND

### *1.1.1. Cultural heritage with digital technique*

In recent years, digital technology application has gradually become a popular topic in the field of architecture design (Bechthold, 2010). In such a direction of academic development, the performance potential of traditional materials and crafts under the support of digital technology has gradually become the focus of attention. In the context of world integration, maintaining traditional handicraft skills is conducive to carry forward the national cultural self-confidence, and show national characteristics under the impact of various cultures in the world. Chinese traditional sculptures arts is the one of the most ancient crafts (Wei, 2006), which usually present on the wooden beams and enclosure structures of traditional buildings. The inheritance of traditional intangible heritage needs to be realized from generation to generation, this is why handicrafts are easily lost during this process.

Besides, based on the new proposal of "*Revitalization plan of Chinese traditional crafts in digital age*" posted by central official cultural organization, this study attempts to establish a "digital craft" workflow, in order to to develop a suitable digital technique in the preservation of intangible cultural heritage. The method to solve this problem should start with analysis and digitization of motion characteristics of the craftsman's manufacturing process, which is exactly what the industrial robot be created for by simulating human real arm.

### *1.1.2. Adaptive technology in robotic crafts*

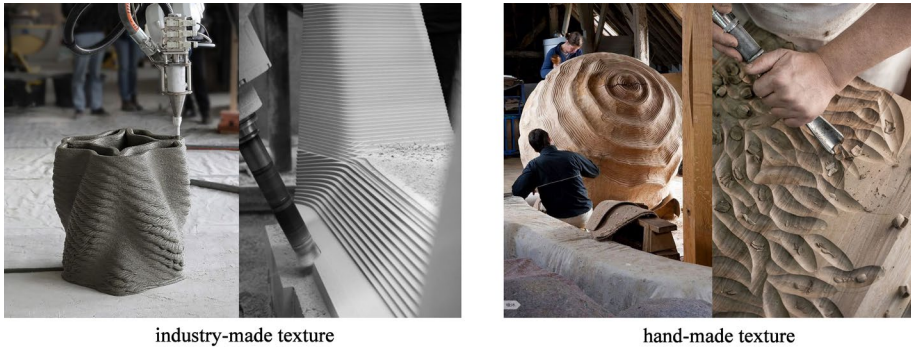
Compared with the method of digital reconstruction for handicraft products under common processes through 3D scanning technology and converting them into virtual digital geometric data for storage (Cheng, 2009), this study is committed to developing a design and production workflow for carving process via industrial robot.

In the manufacturing process of dealing with nonlinear shape or complex surface 3D texture, additive manufacturing and CNC (such as milling and laser cutting which shapes by reducing materials) have been proved to be two suitable methods for complicated fabrication among various scales with robotic arm. Since 3D printing, or additive manufacturing, was pioneered in the 1980s, it has been widely expected to revolutionise the manufacturing method of complex components covering multiple scales, which is also specially applied in digital construction field with robotic arm using in recent years such as concrete component 3D-printing (Aghaei, 2020) and plastic shell formwork 3D-printing (Zhang, 2021). Besides, in terms of reduced material manufacturing process such as robotic milling (Brell-Cokcan, 2010) and saw cutting (Chai, 2021), industrial robotic arm, as an equipment that can be programmed and controlled automatically, has played its spatial advantages of high precision and great flexibility in many construction processes.

However, no matter which of the above manufacturing methods, the behavior trajectory data of the robotic arm is still given by reverse solving the given formal simulation results. Such as the robot milling research led by Sigrid Brell-Cokcan (Brell-Cokcan, 2010). In this project, the geometric surface is designed from the beginning and applied as the milling target for resolving robot control code. Compared with the above methods, this research attempts to induce the internal correlation between form and craft from the rules of mechanical movement itself. Besides, the



creature manufactured by robotic additive and milling technique surfaces shows a trace of industry-made, which runs counter to the texture beauty of hand-made artwork (see Figure 1).



*Figure 1. Texture contrast*

## 1.2. RESEARCH AIM

Based on the above two research backgrounds, the significance of this research is supposed to explore a method deducing the results from the manufacturing process, and then link the form results with robot trajectory control code. Exploration is mainly carried out in three aspects, (shows in Figure 2)

- To analyze the relationship between formal results and movement mode from the perspective of robotic arm kinematics, in order to make robot simulate the action of real human hand.
- To create the robot movement coding mode of the manipulator that defines the axis value, and discuss the possible carved formal results of different robot motion postures.
- To make carving test under the influence of different trajectories that is carried out by disable some axis, from which summarize and accumulate the relationship between carving action and the carved result, in order to establish the direct link between axis value and form.

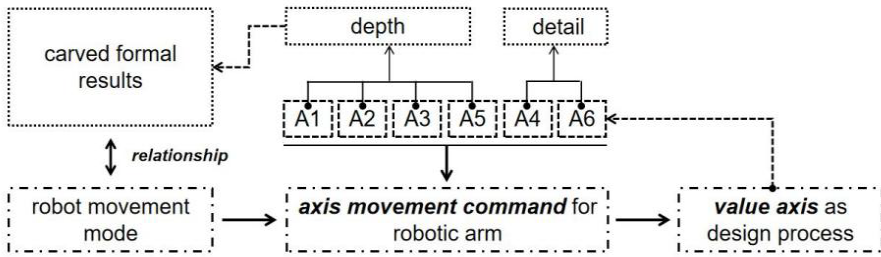


Figure 2. Research workflow

## 2. Methods

In the experiment-based study process, three key parts are carried out in turn, namely the establishment with mathematical model of 6-axis controlled robotic arm movement, formal relationship between material properties of clay and the influence of different movement modes of robotic arm with multi rigid carving tools on form in the process of engraving.

### 2.1. AXIS MOVEMENT COMMAND FOR ROBOTIC ARM

Within the scope of this study, kinematics and path planning of KUKA six axis robotic arm with off-line programming that will be applied in our design system. The movement commands of robotic arm are analyzed thoroughly and abstracted as a simple geometric mathematical model for the subsequent definition of 6-axis more intuitively.

As a general rules, industrial robot off-line programming is similar to the traditional way of using CNC (computer numerical control) machines which is also known as G codes (Mülhe et al. 2010). Different from that, the KUKA Robot Language (KRL) is a proprietary programming language with similarities to Pascal and is used to control KUKA robots (Braumann, 2011). Unlike G-Code, it does not contain just tool and machine movement commands, but can also declare variables and work with conditional clauses. There are some ways of defining movement commands to a robotic arm, the most common ones are the following two (Takase et al. 1981):

- The joint coordinate programming includes an absolute axis rotation command that instructs the robot to move each of its six axes to a defined rotation value in the format like [A1 0, A2 10, A3 90, A4 30, A5 80, A6 30].
- Cartesian coordinate programming defines the position and orientation of the end effector in the previously defined Cartesian coordinate system. The command format code like [X 10, Y 30, Z 60, A 45, B 75, C 15].

The different between these two command method is that the Cartesian way can be a point-to-point (PTP) commands for the end-effector to move from one position to the next with the least amount of axis rotation, or a linear (LIN) commands for the end-effector to move from one position to the next along a straight line. Meanwhile, the

Axis way only can control end-effector to move between positions in the most efficient way.

Obviously, the joint control method is more in line with the command giving method of simulating the movement of human arm. When the end-effector already reach to one certain position, no matter changing which value of any one of the 6 axis, and the angles of the other axis will be solved and transformed accordingly. The posture calculation can be simplified to the geometric relationship shows in the Figure 3. In this geometrical model, the 6 axes are expressed as  $A1(+/-185^\circ)$ ,  $A2(-35^\circ /-135^\circ)$ ,  $A3(+158^\circ /-120^\circ)$ ,  $A4(+/-350^\circ)$ ,  $A5(+/-119^\circ)$ ,  $A6(+/-350^\circ)$  and each axis end are expressed as  $p0, p1-2, p3-4, p5, p6$  and  $T$  is tool center point (TCP).

When  $A4$  is locked in  $0^\circ$  value, all axis end point will be in one space plane called “robot plane”; When  $A4$  is set free,  $p3-4, p5$  and  $p6$  will constitute new space plane which always forms a certain angle with robot plan ( $p0, p1-2, p3-4, p5$ ). In the second condition,  $p5$  is the key factor to make the axis angle solution to be set up, which is defined by vector constructed by  $p6$  and  $p5$ . By this meaning, as long as the coordinates of  $P6$  (When the coordinate value of  $T$  is determined, the coordinate of  $P6$  can be deduced from the geometric relationship of the tool) point in space are uniquely determined, the angle values of the axes  $A1$  to  $A5$  can be calculated according to the fixed length of each axis. In addition, the  $A6$  value defines the interactive relationship between tools posture and working normal plane.

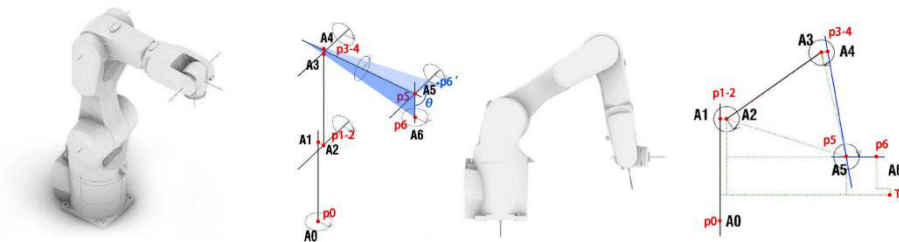


Figure 3. Robot axis value calculation

## 2.2. CARVING PROPERTIES WITH CLAY MATERIAL AND VARIOUS TOOLS

Clay as a common environmental-friendly material is selected as the initial traditional matter of carving experiment, which has good performed density and plasticity, could be the safest and cheapest way for study start. Meanwhile, clay will show some hardness under different drying stages and the moisture content of the material will also affect the edge details of the lines. On the corresponding carving tool selection, it is will know that different carving tools with varied sections shape will leave different carving lines on the clay material. Several carving tools are used for manual engraving research showed in Figure 4, through that the most suitable tools could be selected to do further robotic carve.

One complete carving action includes lowering the knife from one point, moving and lifting after material reduction. According to the length of the moving position, the

carving state will present solid surface textures with different detail characteristics (shows in Figure 4). Conclusion can be drew from one single carving movement that the shape of carved grain is determined by depth and angle of tools, which are defined by arm, wrist and fingers postures. Furthermore, the movement direction of the tool is always at an acute angle with the normal plane of the carved object.

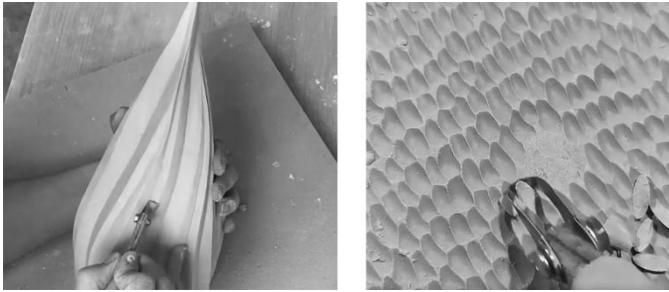


Figure 4. long line carve (left) and short line carve (right)

In order to explore the deep relationship between motion paths and form features, four cross-section shape tools were used for manual experiments which simulates human craftsman. As can see from Figure 5, through the change of certain depth, angle and motion direction, the carved texture will produce small form diversity, and even very rich and delicate details will appear under the change of particular tools combined movement (Figure 6). Similar with hand carving, robotic arm that controlled by axis value would be able to perform like hand movement and cause these complex carved texture features. The action decomposition characteristics of craftsmen in the process of carving complex texture focus on two motion method, the overall displacement of forearm (used to control the general direction of carving lines) and the depth caused by the turnover action of wrist (which determines the local depth change of carving lines).

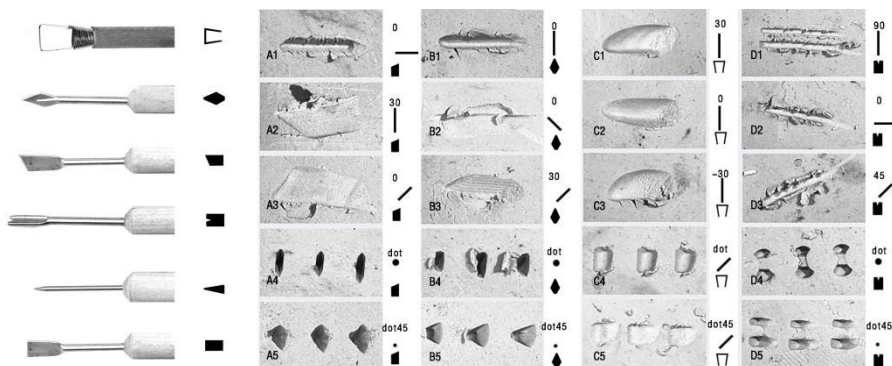


Figure 5. diversity carving movement and their carved shape texture results

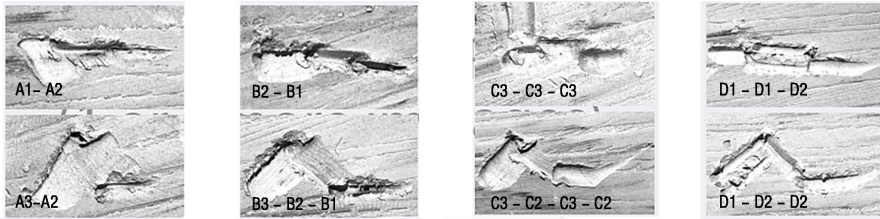


Figure 6. combined movement cause various results

### 2.3. DESIGN STRATEGIES OF CARVING MOVEMENT COMMAND

Taking the research process of robotic carving on a plane as an example (this method could be extended to any type of surfaces with arbitrary spatial curvature), the single value change and regular combination of the six axis of robotic arm would have an impact on carved shape result, such as the depth of carving and the complexity of carved grain. A complete plane-based carving process is carried out in the following order.

- Firstly, the overall trajectory of the carving tools moving on the tangent plane of target clay surface is determined, which is controlled by the A1, A2, A3, A5 and A6 axis.
- Then, the depth of carving lines is determined, which is mainly controlled by the A2 and A3 axis. Based on position points come from the first, the angel  $\theta$  by axis 5 and working plane x direction is a constant value while the angel  $\delta$  equal to 0. The P5 determine the tool vector  $n$  which is also a constant value defined by  $\theta$  and  $\delta$ .
- Finally, the value of A4 and A6 that causes the section shape of carving grain could be defined if the complex carving detail on clay surface is necessary in the design expectation. In this process vector  $n$  will valued by a function consists of two variables  $\theta$  and  $\delta$  (also can be deconstructed into three value x, y, z).

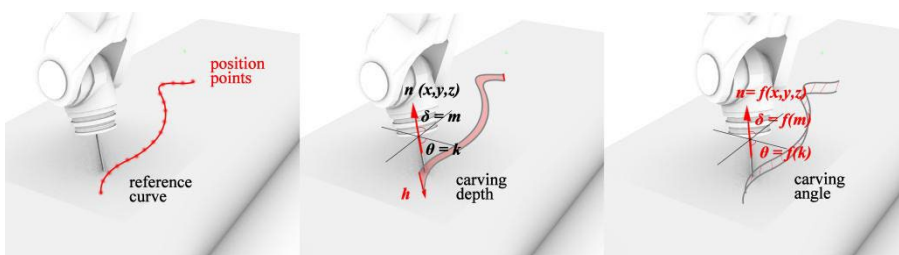


Figure 7. An image with a caption

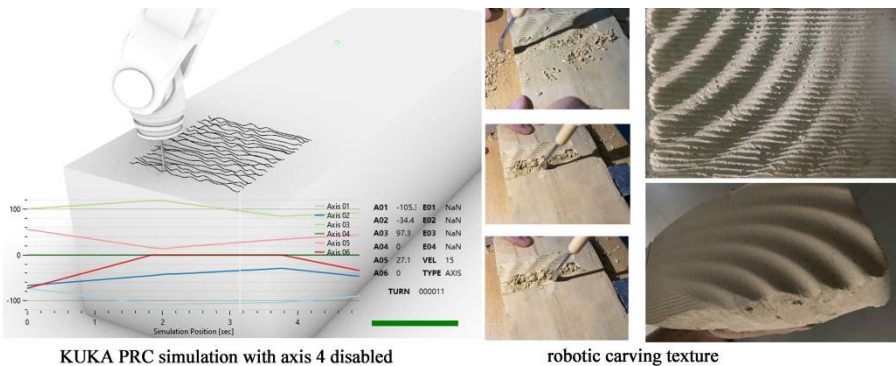
### 3. Robotic clay carving experiments

Applying with the axis movement command, a series robotic carving experiments carried out to testify method discussed above. As an important factor controlling the

carved form, the axis 4 will be discussed in two states of disable and free.

### 3.1. DISABLE 4 AXIS

In the first carving experiment, the axis 4 of the robotic arm is disabled, and the given motion track is the depth fluctuation action perpendicular to the working plane. According to the section width of carving tools and the arrangement density of reference curves, the carving result presents the texture detail state of different resolutions (Figure 8). By giving axis 3 a relatively violent periodic changes showed in simulation graph, the tool end will produce deep and shallow fluctuations on clay surface.



	A1	A2	A3	A4	A5	A5
COMMAND	[0]	[0]	[0]	[0]	[0]	[0]
A01	0 -108.140223	0 -44.950177	0 119.000672	0 0	0 15.949505	0 108.140223
A02	1 -108.140223	1 -42.369802	1 118.426191	1 0	1 13.943611	1 108.140223
A03	2 -107.859068	2 -41.51451	2 116.696636	2 0	2 14.817873	2 107.859068
A04	3 -107.586816	3 -41.496081	3 115.168095	3 0	3 16.327986	3 107.586816
A05	4 -107.302371	4 -40.804506	4 113.330588	4 0	4 17.473919	4 107.302371
VEL	5 -107.02677	5 -40.061967	5 111.458469	5 0	5 18.603498	5 107.02677
C_PTP	6 -106.764234	6 -39.602581	6 109.666701	6 0	6 19.93598	6 106.764234
x	7 -106.525904	7 -38.357387	7 107.757265	7 0	7 20.600123	7 106.525904
axis1	8 -106.280829	8 -37.926176	8 105.948089	8 0	8 21.978088	8 106.280829
y	9 -106.04034	9 -36.960242	9 103.951196	9 0	9 23.009046	9 106.04034
axis2	10 -105.80392	10 -35.992293	10 101.902339	10 0	10 24.089954	10 105.80392
axis3	11 -105.570099	11 -35.108314	11 99.812434	11 0	11 25.29588	11 105.570099
axis4	12 -105.363449	12 -34.771107	12 98.018083	12 0	12 26.753024	12 105.363449
axis5	13 -105.154422	13 -33.566961	13 95.877879	13 0	13 27.689082	13 105.154422
n	14 -104.952038	14 -33.107787	14 93.930309	14 0	14 29.177478	14 104.952038
axis6	15 -104.743975	15 -32.208096	15 91.712983	15 0	15 30.495113	15 104.743975
pT	16 -104.541473	16 -31.149588	16 89.403329	16 0	16 31.746258	16 104.541473
plane45T	17 -104.348113	17 -30.133392	17 87.096845	17 0	17 33.036547	17 104.348113
planeal	18 -104.169687	18 -28.949658	18 84.795737	18 0	18 34.153921	18 104.169687

Figure 8. Disable axis 4 motion data and carved detail

### 3.2. SET ALL AXIS FREE

Next step, all axis free method is applied to carve the same humidity clay board. In the design of motion, the rotation of axis 6 can rotate the tool section, and then make the carving results produce rich shape changes (see figure 9). By constructing the vectors value discussed in the previous chapter, the tool end can move freely on the work plane to interact with clay surface in a more flexible way.



KUKA PRC simulation with all axis free

Figure 9. All free axis motion data and carved detail



Figure 10. Comparison between the results details of robotic engraving and traditional carving craft

After realizing the experiment of mechanical engraving based on plane, this method can also be applied to spatial surfaces with arbitrary curvature to realize multi-dimensional complex carving which will explore in our further research.

#### 4. Conclusion

The study of handicrafts in traditional culture is a reference for designers to excavate history and create new forms. In this process, preserving Chinese ancient skills especially traditional fabrication crafts is the main motivation for this research to discuss with digital manufacture tools and parametric design system.

From these research background, proposing the process-oriented design method and realizing the way of manipulating robot imitating real human craftsmen are two main goals that payed more attention to in the study. The experiments show the

possibility to create complex carving details on the material surface that can be easily carved such as clay and soft wood through defining robot movement way, which could fundamentally make robot manufacturing a new formal meaning. In this way, when facing the mechanical manufacturing tools, the designer's attention will change from backward resolving the posture of the robotic arm according to the results of the 3D model, to manipulation the movement of machine itself for “process led design” deduction. Making robot thinking and moving like a real human, which shows a great possibility and charming challenge in further research.

## References

- Aghaei Meibodi, M., Voltl, C., Craney, R. Additive Thermoplastic Formwork for Freeform Concrete Columns. *ACADIA 2020: Distributed Proximities / Volume I: Technical Papers* (pp. 516-525), ISBN 978-0-578-95213-0.
- Bärtschi, R., M. Knauss, T. Bonwetsch, F. Gramazio, and M. Kohler. 2010. Wiggled Brick Bond, *Advances in Architectural Geometry 2010*: 137-148, Vienna: Springer.
- Bechthold, M. 2010. The Return of the Future- A Second Go at Robotic Construction *Architectural Design* 80(4):116-121. Hoboken: Wiley and Sons.
- Braumann, Johannes; Brell-Cokcan, Sigrid (2011). Parametric Robot Control: Integrated CAD/CAM for Architectural Design. *ACADIA 11: Integration through Computation Proceedings of the 31st Annual Conference of the Association for Computer Aided Design in Architecture (ACADIA)* (pp. 242-251), Banff (Alberta) 13-16 October, 2011.
- Brell-Cokcan, Sigrid; Braumann, Johannes. A New Parametric Design Tool for Robot Milling, *ACADIA 10: LIFE in:formation, On Responsive Information and Variations in Architecture [Proceedings of the 30th Annual Conference of the Association for Computer Aided Design in Architecture (ACADIA)* (pp. 357-363), New York 21-24 October, 2010).
- Chai, H., So, C., & Yuan, P. F. (2021). Manufacturing double-curved glulam with robotic band saw cutting technique. *Automation in Construction*, 124, 103571.
- Cheng, Hung-Ming; Ya-ning Yen and Wun-bin Yang (2009). Digital Archiving in Cultural Heritage Preservation. *Proceedings of the 14th International Conference on Computer Aided Architectural Design Research in Asia / Yunlin (Taiwan) 22-25 April 2009*, pp. 93
- Mühe, H., A. Angerer, A. Hoffmann, and W. Reif. 2010. On reverse-engineering the KUKA Robot Language. *Proceedings of the DSLRob ' 10, IEEE/RSJ International Conference on Intelligent Robots and Systems 2010, Taipeh*.
- Takase, K., R. Paul, and E. Berg. 1981. A Structured Approach to Robot Programming and Teaching, *IEEE Transactions on Systems, Man, and Cybernetics* 1(4). New York: IEEE.
- Wei, W. U. (2006). On the Relationship between Artistic Design and Crafty Arts in China. *Journal of Zhuzhou Institute of Technology*, 1.
- Zhang, Xiao, Yuan, Chao, Yang, Liu, Yu, Peiran, Ma, Yiwen, Qiu, Song, Guo, Zhe and Yuan, Philip F. 2021. Design and Fabrication of Formwork for Shell Structures Based on 3D-printing Technology. Stojakovic, V and Tepavcevic, B (eds.), *Towards a new, configurable architecture - Proceedings of the 39th eCAADe Conference* (pp. 487-496), Volume 1, University of Novi Sad, Novi Sad, Serbia, 8-10 September 2021.



# REMEMBERING URBAN VILLAGE

*Using CloudXR Technology as an Enhanced Alternative to Better Disseminate Heritage*

XINYI ZHOU<sup>1</sup>, YAO CHEN<sup>2</sup>, FUKAI CHEN<sup>3</sup>, KAN LI<sup>4</sup>,  
TIANTIAN LO<sup>5</sup>, RUFENG XIANG<sup>6</sup> and LIQUAN LIU<sup>7</sup>

<sup>2,3,5</sup>*Harbin Institute of Technology (Shenzhen)*

<sup>1</sup>*xinyiz.uk@gmail.com, 0000-0002-8781-209X*

<sup>2</sup>*chenaoyaoyao0809@gmail.com*

<sup>3</sup>*stevenchen1703@gmail.com*

<sup>4</sup>*kan.li@aschool.ac.uk*

<sup>5</sup>*skyduo@gmail.com*

<sup>6</sup>*xrf7000@163.com*

<sup>7</sup>*leo@configreality.com*

**Abstract.** Urban villages are strictly related to urban growth. It reflects the era characteristics and memory in urban growth, which has significant value in heritage and sustainable cities (SDG 11). Due to the continuous development of urbanization and the shortage of urban land, many urban villages will be replaced by more valuable functions. Therefore, better preserving the digitalization of urban villages and making more people understand the value of urban villages is particularly important. However, the existing technology still has shortcomings in disseminating digital heritage. For urban villages, usually a large-scale and complex environment, the hardware requirements will be very high for high-precision visualization. Most existing solutions use large hardware devices, such as the virtual sand table. Unlike hand-held devices, such devices are expensive and not portable, limiting better dissemination of such heritage. Due to the hardware limitation of hand-held devices, neither the display resolution nor the interaction effect is satisfying. Therefore, this paper proposes a new workflow by NVIDIA CloudXR streaming technology to achieve high-precision visualization and a rich interactive experience on hand-held devices. Such heritage can be promoted and cities can be more sustainable.

**Keywords.** CloudXR Technology; Urban Village; Digital Heritage; Preservation; Dissemination; Portable Devices; SDG 11.

## 1. Introduction

Urban villages appear in the central parts of major Chinese cities, including Shenzhen

and Guangzhou. They are surrounded by skyscrapers and the infrastructure of modern cities, a unique phenomenon in China's urbanization process. Urban villages have become a powerful booster for the urbanization process, providing low-income immigrants with affordable housing and daily needs in the past few decades. It reflects the characteristics and memories of the current era in the city's development, and it is of great value to be recorded and widely disseminated as a historical heritage.

However, many urban villages in China are endangered. Due to the continuous urbanization process and the shortage of urban land, many urban villages will be replaced by new spaces with more valuable functions. Therefore, utilizing digital technologies to preserve urban villages and let more people realize the value of urban villages is particularly essential.

Since the emergence of digital technology, digital heritage has become an essential part of the whole world's cultural heritage. With the digital evolution, the data accumulation and the application of information processing systems are becoming universal, enabling updates of heritages such as urban villages. Nowadays, data sources are highly accurate, have timeliness, and ample space span. It enables cultural heritage (such as the spatial appearance of a village in a city) to be stored in digital formats (such as CDs and disks) to achieve persistent data Storage (Tobiasz et al., 2019). At the same time, digital storage data provides valuable resources for cultural heritage preservation and research in this field (Hoon et al., 2019).

However, the explosive growth of data poses specific challenges to data analysis and processing capabilities. There are still bottlenecks in transforming data into more general information or knowledge to serve management departments, research institutions, museum industry, education, and the public related to cultural heritage (Hou et al., 2018). The application of digital technology in urban villages still has problems such as low information integration and weak information. In addition, the limitations of the current equipment's communication capabilities and user sharing capabilities have severely restricted the promotion and integration of the cultural heritage of urban villages and society. The village in the city is usually a large-scale and complex environment, and the requirements for high-precision visualization of hardware will be very high. Most existing solutions use large hardware devices, such as virtual sandboxes. Unlike hand-held devices, such devices are expensive and not easy to carry, which dramatically limits the spread of large-scale urban village heritage. Moreover, due to the hardware limitation of the hand-held device, its display resolution and interactive effect are not ideal.

Therefore, it is urgent to organize and correlate multi-source data both accurately and efficiently. Potential solutions include extracting essential information based on the characteristics of different data types, such as drone scanning, and visualizing it in various data formats. In this way, various scalable and customizable solutions can be further proposed to meet the requirements of different applications. With the evolution of computer technology, cloud computing platform, as the core technology of digital urban villages to promote resource sharing and iteration, is an inevitable choice to respond to the current situation.

This article proposes a new workflow based on NVIDIA CloudXR streaming media technology to achieve a high-precision visualization effect and a rich interactive experience on hand-held devices. Thus, digital heritages can be popularized more

effectively.

## **2. Related Technologies**

For the digitization of large-scale and complex heritage types such as urban villages, the main applications can be divided into two types: focusing on specific applications and application systems and workflows.

### **2.1. THE APPLICATION OF DIGITAL HERITAGES IN URBAN VILLAGES**

The purpose of protecting the digital cultural heritage of urban villages is conducive to the persistence of historical archives and to use these digital resources to provide more beneficial help for the sustainable development of cultural heritage. Researchers in this field have made great efforts. Digital technology has attempted to visualize three-dimensional urban villages, but it is challenging for most researchers to understand and interact with portable devices. For example, the "4D-GIS" model in Longwu Village, Shenzhen, synthesizes material flow and stock analysis (MFSA) with geographic information system (GIS) to illustrate time and spatial attributes of buildings and the evolution of materiality metabolism (Wang et al., 2019).

Despite the continuous iterations in technological development and dissemination efficiency, the preservation and dissemination of digital heritages in urban villages is non-user-oriented, designed, and developed in a non-descriptive way. They only focus on the "process" (data authentication, on-site investigation, and display) or "product" (re-interpretation and the representation of technical art which are more similar to reality), but seldom consider the "end-user" (how end-users are disseminated) (Rahaman and Tan, 2011).

### **2.2. APPLICATION SYSTEM AND WORKFLOW**

With the advancement of computing skills, mobile devices, display technology, wide-ranging high-speed networks, and the increasing research and development of immersive digital media and artificial intelligence (AI), more and more attention is paid to the dissemination and preservation of urban villages by end-users. In terms of equipment of extended reality (XR), including augmented reality (AR), virtual reality (VR), and mixed reality (MR), including having provided consumers with convenient and affordable experiences (such as mobile phones or iPads). It offers more opportunities for immersive interaction. XR on mobile devices provides a new experience for us to interact and interpret the world, mobile games such as Pokemon Go, Ingress Prime, and Knightfall: AR, educational experiences such as BBC Civilisations AR, and practical everyday tools such as Google Lens, Translate and Maps AR. Regarding data processing for hand-held devices, improving 5G communications has shown great potential in real-time rendering and rapid delivery of the high-quality immersive experience. There are many technologies, including cloud-based services and machine learning, create location-based personalized experiences, such as NVIDIA's CloudXR, streaming dense graphics for XR and GridRaster, support artificial intelligence-based immersive MR platforms, and assist in creating a more meaningful and more engaging experience, such as PopStic VR and Epic's Apollo 11: Mission AR on Microsoft HoloLens. Microsoft Mesh also interrupts our interaction

with the real and virtual worlds, introducing digital intelligence into the real world to disrupt the way we share experiences, collaborate, communicate and recognize existence.

To promote sustainable development of the digital cultural heritage of urban villages and disseminate it effectively, researchers should cooperate to achieve a higher level of information visualization sharing skills and data interaction with certain restrictions on copyright. Therefore, it is urgent to design a more flexible systematic structure to satisfy end-users various demands and device restrictions to create a general CloudXR workflow for the digital cultural heritage in urban villages.

### 3. Method

#### 3.1. METHODOLOGY

There are many methodologies for architectural heritage preservation, such as the eCAADe 2016 paper on 3D Digital Reconstruction of Lost Buildings that has inspired a further workflow in this paper (Mascio et al., 2016). This paper proposes a methodology with documentation-computation-dissemination (DCD) workflow, mainly utilizing NVIDIA CloudXR technology to promote high accessibility, fidelity, and interactivity in disseminating digital heritage. This methodology proposes a new perspective of the preservation of digital architectural heritage. In figure 1, The application of the method is shown in the three sections of "What," "Why," and "How." The workflow content such as Documentation, Digital Reconstruction, Computational Process, and Dissemination is listed in "How."

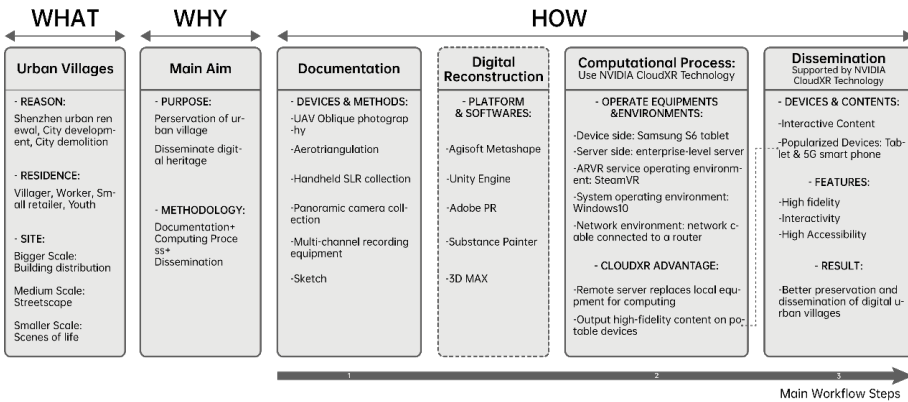


Figure 1. DCD Workflow that highlights technology to promote the spread of digital heritage

#### 3.2. DOCUMENTATION

In this documentation part, various technical methods are being used to collect data on Baishizhou from three different scales to collect and preserve the appearance of urban villages more comprehensively. In terms of large-scale: The general form, overall layout, and regional distribution of urban villages are preserved. The specific method is to use a drone to carry a 20-megapixel lens to perform oblique photography and then operate aerial triangulation. In terms of medium-scale: Preserve the daily street scenes

and the spatial experience in human perspectives, which use four different ways to synthesize. Small-scale: Through field measurement and digital modeling to document interior living spaces.

### 3.3. COMPUTATIONAL METHOD

#### 3.3.1. Introduction to NVIDIA CloudXR Technology

Compared with using traditional localized high computing power equipment, NVIDIA CloudXR technology can remotely control server equipment for data processing. It is innovative digital heritage preservation and can replace localized high computing power equipment.

CloudXR breaks the traditional limitations of VR and AR, streaming XR content to untethered devices, and its fidelity is indistinguishable from the native bondage configuration. Figure 2 visualizes the equipment and operating environment of NVIDIA CloudXR Technology from three parts: Server, Network, and Client.

#### 3.3.2. Operate Equipment Used in NVIDIA CloudXR Technology

When using NVIDIA CloudXR technology, it needs to use devices at both sides: the device and the server, respectively. Device: In the AR (augmented reality) environment, the Samsung S6 tablet is used. If in the VR environment, the Meta Quest 2 is used. Server: The server is configured with two Intel Zhiqiang E5 processors, 96GB ECC memory, SDD disk matrix with a total capacity of 2T, NVIDIA RTX 2060S four graphics cards are connected in a pass-through mode. ARVR service operating environment: The SteamVR environment. System operating environment: Windows10.

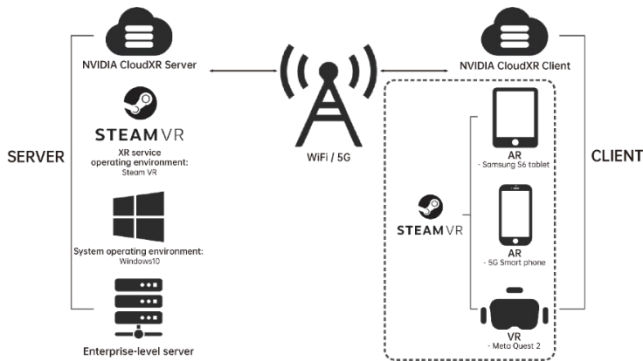


Figure 2. NVIDIA CloudXR Architecture

#### 3.3.3. Practical results of using NVIDIA CloudXR technology workflow

To practice the workflow using NVIDIA CloudXR technology, we use the tablet as the output terminal of AR and Meta Quest 2 as the VR output terminal to visualize. As shown in figure 3, this table highlights the advantages of CloudXR technology and the results of CloudXR technology testing using different devices such as AR and VR.

Computational Process: Use NVIDIA CloudXR Technology					
OPERATE EQUIPMENTS & ENVIRONMENTS:	CLOUDXR ADVANTAGES:	TEST RESULT:			
-Device side: Samsung S6 tablet -Server side: enterprise-level server -ARVR service operating environment: SteamVR -System operating environment: Windows10 -Network environment: network cable connected to a router	-1. Remote server replaces local equipment for computing -2. Output high-fidelity content on portable devices	The server is connected to the router-ASUS AiB8u through a physical network cable, and it is maintained at the 60MHz frequency band width			
		AR - Samsung S6 tablet	VR - Meta Quest 2		
		Single clean channel	Video stream compression to 50Mbps	Single clean channel	Video stream compression to 50Mbps
		Number of devices present at the same time:			
10	30	4	12		

Figure 3. Computational process

### 3.4. DISSEMINATION

#### 3.4.1. Interactive design in the dissemination

When disseminating digital heritage, NVIDIA ClouxXR technology significantly promotes the accessibility and fidelity of the content. Meanwhile, the Unity engine enables visualized information, amusing animation, sound effects, and UIUX design to be linked to the digital heritage to enhance the fun and interactivity of the content. Therefore, it enhances the experienter's curiosity and acceptance of digital heritage.

#### 3.4.2. Hand-held devices in the dissemination

In figure 4, this picture shows the three advantages of CloudXR technology in disseminating digital heritage. The following are the three aspects that why CloudXR technology is disseminated efficiently: a. Based on the application of CloudXR technology, complex data is transferred by the network to the edge and remote high computing power devices for calculation, maintaining the high fidelity of content b. Combined with the popularization of 5G communication networks, CloudXR technology is better used in hand-held devices with higher popularity. c. Better interactive experience.

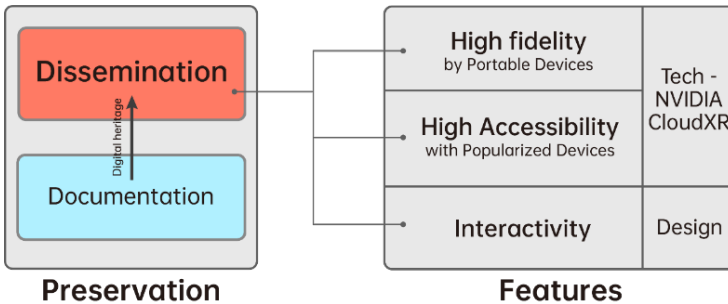


Figure 4. Portable Devices in Dissemination

#### 4. Experimental System and Case Study

This study takes Baishizhou, Shenzhen, Guangdong Province, China as an example to test the feasibility of a new workflow based on NVIDIA CloudXR streaming media technology in digital heritage preservation. As a newly emerged city, Shenzhen has a limited history that can be traced back, and the urban villages in the city are witness to this collective memory. Baishizhou is a place for the fringe people to integrate into city life as much as possible. Low-income labour and rural migrations live temporarily in an urban village like Baishizhou, different from most developed urban spaces. However, it still could be possible to identify their urbanites identities. The preservation of the memory space of Baishizhou is not only related to the materiality and construction history of its architectural space but also to how its spatial characteristics affect our experience, thus emerging the place property of Baishizhou. (Junyi, 2017, p.24) The protection of heritage in Baishizhou also means the protection of collective memory of people related to the space.



*Figure 5. The Location and Street View of Baishizhou*

Based on the current situation of Baishizhou, we tried to make the Baishizhou AR interactive sandbox open to the public for the experience. The purpose is to allow the people of Baishizhou to experience and communicate in the IDG industry incubation centre space. Figure 6 shows the experimental process.

First, we extract the data collection of Baishizhou; use Agisoft Metashape and other software to generate 3D architectural models from 3D scanning data. After that, use 3DMax to model and make animations. Then, import digital assets such as the generated models and animations into the Unity game engine; carry out model reduction and resource integration. Finally, use NVIDIA CloudXR workflow to transfer the content that needs to be presented to the remote tablet for display. The interactive user experience is as follows: a.Activates the device camera's Ray casting algorithm. When the user aligns the device camera at a specific area of the digital model (Baishizhou consists of six regions), the area will be highlighted, and the device screen will pop out of it. b.After the highlighted area is displayed, slide the screen upwards: the specified animation and sound effects emerge, and then the detailed information interface of the village pops up. c.Swipe the screen up again: the designated animation and sound effects appear, the iconic digital building in the corresponding area with animation is shown. d.Based on the technical support of bi-directional audio in CloudXR, users who experience the product remotely can simultaneously have real-time communication and discussion.



Figure 6. AR representation and Multiplayer experience in experiment

We invited 27 people to participate in this experiment: Currently living in Baishizhou: 11 people and previously living in Baishizhou: 16 people. In the experiment, 27 volunteers were divided into two groups to conduct experiments respectively: The first group consists of 11 people currently living in Baishizhou; two to three people experience at the same time as a group; hold Samsung tablet devices; the experience time is controlled within 5-10 minutes. The second group consists of 16 people who have lived in Baishizhou; the settings of factors such as the number of experienced persons, experience equipment, and experience duration are the same as those of the first group of experiments. Figure 8 shows the experimental process of the multiplayer experience.

We evaluate the experience process into three parts: model presentation content, interactive experience, and overall experience. Meanwhile, we refined the specific content of each part based on the Baishizhou AR sandbox experience, made these nine items into a table (experience level is divided into ten Levels), and invited a simple questionnaire survey after the people have experienced it. In the end, the questionnaire data filled out by the 27 people we interviewed in the future was calculated into the following table:

## 5. Results & Discussion

This section begins with a summary of the seminar. We choose to use the main subject categories that appear in the data to present the results (Figure 9):

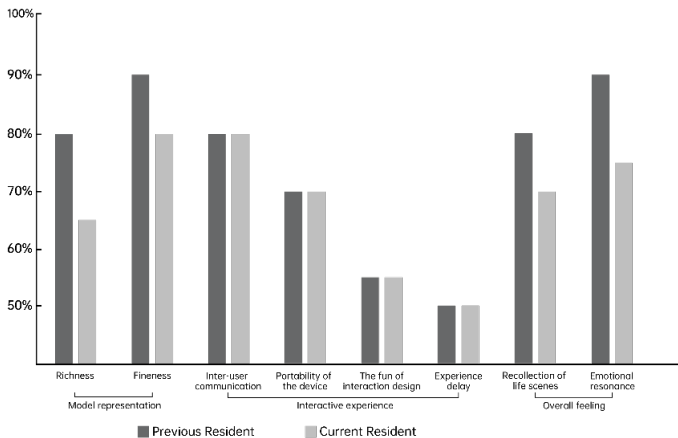


Figure 7. Data of 3 aspects (model representation, interactive experience, and overall feeling)



According to the comprehensive table, based on the high-precision Baishizhou data collection and NVIDIA CloudXR technology for the transmission of high-precision content, the Baishizhou AR Sandbox shows excellent performance in the presentation model content, regardless of the richness or the accuracy of the model content. It exceeds the imagination of the viewer significantly. From the high-precision data collection of Baishizhou and the presentation of high-precision content, NVIDIA CloudXR technology has played a vital role in the workflow. From the perspective of interactive experience, the portable tablet archives the multiplayer AR experience. It creates real-time communication opportunities for the viewers while watching the sandbox content. It is noticeable that the amuse and interactivity of UI design, animation, and sound effects in the display process need to be improved. The content delay caused by NVIDIA CloudXR technology needs to be continuously improved technically. Finally, from the viewer's overall experience, the high-precision AR sandbox brings a fantastic visualization experience to the experienter, breaking through the limitation of time and space, creating a "space-time journey" and a natural, localized experience.

Essential aspects of Cloud XR optimization are finding the right balance between the calculations performed and the calculations performed by the cloud and optimizing the cloud's computation efficiency (R. Yadav et al., 2020). Many similar pieces of research have proposed different strategies to optimize the XR scene rendering, such as caching the actively generated video frames and their transmission by reserving network resources (Y. Sun et al., 2019; M. S. Elbamby et al., 2018; T. Xu et al., 2019). However, the application of CloudXR technology in architectural heritage and sustainable cities is still very limited. Based on this attempt, it is believed that new ideas will make urban villages such communities preserved and documented better.

## 6. Conclusion

The research uses this workflow based on CloudXR technology to present the large-scale and complex architectural heritage of urban villages on mobile-device. For the large-scale and complex architectural heritage of urban villages, limited by the computing power of hardware devices, the existing presentation methods have certain limitations. That is, relatively large-scale hardware devices are required. Therefore, we hope to explore a mobile-device method to disseminate urban villages better. CloudXR provides technical support for this demand.

To summarize, in the era when the wave of urbanization in China is still raging, one after another. Urban villages are about to disappear, and cities' sustainability gradually becomes an essential thesis. Using scientific technology and innovative design to retain and preserve them has become an urgent research content.

## Acknowledgments

We very much appreciate the technical production of "Baishizhou Renewal" provided by Configuration Studio. The specific thanks are as follows: Derek Peng provides information on technical details; Rufeng Xiang's technical contribution to Baishizhou documentation and implementation of NVIDIA CloudXR Technology; Zhiwei Deng and Haida Sun's contribution to the realization of NVIDIA CloudXR Technology;

Mengqi Tu's contribution to interactive animation content production; Leo Liu provides technical and innovative advice; Derek Peng and Shenmeizi Cao's contribution to commercial publicity and promotion of the project.

## References

- Elbambay, M. S. , Perfecto, C. , Bennis, M. , & Doppler, K. . (2018). Edge Computing Meets Millimeter-wave Enabled VR: Paving the Way to Cutting the Cord. *IEEE WCNC'2018*. IEEE.
- Hafizur, Rahaman, Beng-Kiang, & Tan. (2011). Interpreting digital heritage:a conceptual model with end-users' perspective. *International Journal of Architectural Computing Ijac*.
- Hao, P., Sliuzas, R & Geertman, S. (2011). The development and redevelopment of urban villages in Shenzhen. *Habitat International*, 35.
- Hou, M., Yang, S., Hu, Y., Wu, Y., Jiang, L., Zhao, S., & Wei, P. (2018). Novel method for virtual restoration of cultural relics with complex geometric structure based on multiscale spatial geometry. *ISPRS International Journal of Geo-Information*, 7(9), 353.
- Jo, Y. , & Hong, S. . (2019). Three-dimensional digital documentation of cultural heritage site based on the convergence of terrestrial laser scanning and unmanned aerial vehicle photogrammetry. *International Journal of Geo-Information*, 8(2).
- Leng, J. (2017). The Value and Utilization of the Memory Space of Urban Villages in Urban Renewal from the Perspective of Symbiosis Thought: A Case Study of the North District of Baishizhou, Shenzhen. *China Academic Journal Electronic Publishing House*. TU-10590-31.
- Sun, Y. , Chen, Z. , Tao, M. , & Liu, H. . (2019). Communications, caching and computing for mobile virtual reality: modeling and tradeoff. *IEEE Transactions on Communications*, 67(99), 7573-7586.
- Tobiasz, A. , Markiewicz, J. S. , Sawomir apiński, Nickel, J. , & Muradov, M. . (2019). Review of methods for documentation, management, and sustainability of cultural heritage. *Case study: museum of king jan iii's palace at wilanów*. *Sustainability*, 11(24), 7046.
- Wang, H., Chen, D., Duan, H., Yin, F., & Niu, Y. (2019). Characterizing urban building metabolism with a 4D-GIS model: A case study in China. *Journal of Cleaner Production*, 228, 1446-1454.
- Xu, T. , Sun, Y. , Xia, S. , Li, H. , & Chen, Z. . (2019). Optimal Bandwidth Allocation with Edge Computing for Wireless VR Delivery. *2019 IEEE/CIC International Conference on Communications in China (ICCC)*. IEEE.
- Yadav, R. , Zhang, W. , Huang, C. , & Tao, G. . (2017). MuMs: Energy-Aware VM Selection Scheme for Cloud Data Center. *International Workshop on Database & Expert Systems Applications*. IEEE.

## INDEX OF AUTHORS

### A

Akbar, Zuardin 1-393  
 Alani, Mostafa 1-151  
 Althoff, Klaus-Dieter 1-323,  
 1-403, 2-435  
 Alva, Pradeep 1-525  
 Aman, Jayedi 1-595  
 Amtsberg, Felix 2-141  
 Anifowose, Hassan 2-547  
 Antonio, Rishan 1-435  
 An, Yudi 1-605  
 Araujo, Goncalo 2-689  
 Ataman, Cem 1-383

### B

Bae, Jiyoona 2-121  
 Balaban, Özgün 1-69  
 Banihashemi, Farzan 2-629, 2-679  
 Bao, Ding Wen 1-121  
 Bao, Dingwen 2-71  
 Barath, Shany 1-171  
 Barton, Jack 1-687  
 Bedarf, Patrick 2-61  
 Beebe, Aaron G. 1-39  
 Belek Fialho Teixeira, Müge 2-21  
 Bidgoli, Ardavan 1-373  
 Bielski, Jessica 1-323,  
 1-403, 2-435  
 Boim, Anna 1-223  
 Bolojan, Daniel 1-353, 1-363  
 Brown, Nathan 1-111  
 Bui, Do Phuong Tung 2-495  
 Burden, Alan 2-21  
 Byrne, Daragh 1-333

### C

Cai, Chengzhi 2-233  
 Caldwell, Glenda 2-21  
 Cao, Chufan 1-565  
 Cao, Qingning 2-233  
 Cao, Xiaoyu 2-233  
 Carta, Silvio 1-615

Cattaneo, Tiziano 1-717  
 Chadzynski, Arkadiusz 1-555  
 Chai, Hua 2-41  
 Chang, Teng-Wen 2-201  
 Chavan, Tejas 1-313, 1-435  
 Cheddadi, Aqil 1-585  
 Chen, Chun-Yen 2-201  
 Chen, Fukai 1-757  
 Cheng, Cesar 1-313, 1-435  
 Cheng, Chung-Chieh 2-213  
 Cheng, Nancy 2-345  
 Cheng, Sifan 1-495  
 Chen, Guoyi 1-161  
 Chen, Hao 1-89  
 Chen, Jielin 2-385  
 Chen, Kian Wee 1-415  
 Chen, Lei 2-517  
 Chen, Pei 1-181  
 Chen, Qinchuan 1-19  
 Chen, Shi Yu 1-505  
 Chen, Ting-Chia 2-243  
 Chen, Yao 1-757  
 Chen, Yulong 1-625  
 Cheung, Ling Kit 1-181  
 Choi, Seungcheol 1-161  
 Choo, Seungyeon 2-597  
 Chowdhury, Shuva 2-465  
 Chronis, Angelos 1-273,  
 1-545, 2-577

Cole, Laura 2-455  
 Cui, Qiang 2-101, 2-171  
 Cui, Qinyu 1-515  
 Cui, Weiwen 2-567

### D

Dai, Sida 1-151  
 Danahy, Patrick 1-131  
 Davies, Peter 2-749  
 Davis, Felecia 1-111  
 De Wolf, Catherine 2-577  
 Dengel, Andreas 1-323,  
 1-403, 2-435  
 Deshpande, Rutvik 1-313, 1-435

- Dias Guimaraes, Gabriela 2-587  
 Dillenburger, Benjamin 2-61  
 Ding, Xinyue 2-425  
 Dixit, Manish 2-547  
 Dong, Zhiyong 1-213, 1-263  
 Donovan, Jared 2-21  
 Doria, David 2-759  
 Dorteimer, Jonathan 1-223  
 Dritsas, Stylianos 2-263  
 Duering, Serjoscha 1-545  
 Dunn, Kate 1-687  
 Duran, Ayca 2-669  
 Düring, Serjoscha 1-273
- E**  
 Eisenstadt, Viktor 1-323,  
 1-403, 2-435  
 Eshaghi, Sarvin 1-69
- F**  
 Fang, Yu-Cyuan 2-201  
 Farr, Marcus 2-293  
 Feng, Jiajia 1-425  
 Feng, Xiqiao 2-101  
 Fernandez, Alberto 2-759  
 Fernandez, Javier 2-263  
 Fingrut, Adam 2-11  
 Fink, Theresa 1-545  
 Förster, Nick 1-635  
 Fujii, Haruyuki 2-475  
 Fuji, Takaaki 1-303  
 Fukuda, Tomohiro 1-89, 1-737,  
 2-395, 2-607
- G**  
 Gamberro, Julien 2-325  
 García del Castillo y  
 López, Jose Luis 1-505,  
 2-283, 2-485  
 Gardner, Nicole 1-1, 1-253,  
 1-687, 2-1  
 Gero, John S. 2-315  
 Globa, Anastasia 2-649, 2-749  
 Goel, Abhimanyu 2-365  
 Gomes da Silva, Vanessa 2-587  
 Gomes, Ricardo 2-689  
 Gong, Lei 2-273  
 Gong, Pixin 1-455, 1-645  
 Gough, Phillip 2-649
- Gramazio, Fabio 2-111  
 Grisiute, Ayda 1-555  
 Grobman, Jacob 1-283  
 Grugni, Francesco 1-717  
 Gu, Hyeongmo 2-597  
 Guida, George 2-485  
 Gu, Ning 2-527, 2-587  
 Guo, Xiangmin 2-223, 2-425  
 Guo, Yefei 2-233  
 Guo, Yiyao 2-91  
 Guo, Yuchen 2-233  
 Guo, Zhe 1-747, 2-233  
 Guo, Zhixian 2-41  
 Gursel Dino, Ipek 2-669
- H**  
 Haeusler, Matthias Hank 1-253,  
 1-273,  
 1-687, 2-739  
 Hanegraaf, Johan 2-465  
 Hao, Qi 1-425  
 Helmreich, Matthias 2-111  
 He, Mengxi 2-171  
 Herthogs, Pieter 1-555  
 Heusi, Alex 2-61  
 Hirano, Yuji 2-131  
 Holloway, Leona 2-709  
 Hong, Soonmin 2-597  
 Hoo, Jian Li 2-263  
 Hotta, Akito 1-303  
 Hotta, Kensuke 1-303, 2-131  
 Hou, June-Hao 1-99, 1-667,  
 2-445  
 Hsiao, Chi-Fu 2-201  
 Huang, Chenyu 1-141,  
 1-233,  
 1-455, 1-645  
 Huang, Ching-Wen 2-243  
 Huang, Jeffrey 1-727  
 Huang, Jie 2-567  
 Huang, Shuyi 1-475  
 Huang, Tracy 1-687  
 Huang, Xiaoran 1-455,  
 1-645, 1-707  
 Huang, Yiting 1-515  
 Hu, Huiyao 2-495  
 Hyun, Kyung Hoon 1-1, 1-59,

	2-1	Kuroki, Mitsuhiro	2-111
<b>I</b>			
Ikeda, Yasushi	1-585, 2-31, 2-131		
Inaba, Tooru	2-131		
Irger, Matthias	2-739		
Iseri, Orcun Koral	2-669		
Ito, Koki	2-131		
Ito, Nao	1-737		
<b>J</b>			
Jahn, Gwyllim	2-191		
Janssen, Patrick	1-415, 2-495, 2-505		
Jaramillo Pazmino, Pablo	2-181		
Isaac			
Jeong, Joowon	1-19		
Ji, Guohua	2-699		
Johansson, Mikael	1-29		
<b>K</b>			
Kahlon, Yuval	2-475		
Kaljevic, Sofija	1-677		
Kalkan, Sinan	2-669		
Karastathi, Nikoletta	2-759		
Kawakami, Takuma	2-131		
Keane, Adrienne	2-749		
Khajehee, Arastoo	2-31, 2-131		
Khean, Nariddh	1-273		
Kiesewetter, Laura	1-393		
Kikuchi, Naoki	2-607		
Kikuzato, Naoto	1-303		
Kim, Dongyun	2-485		
Kim, Frederick Chando	1-727		
Kim, Hwan	1-59		
Kim, Jong Bum	1-595, 2-455		
Kim, Ki	2-587		
Kim, Nayeon	1-19		
Kim, Taehoon	2-597		
Kleiss, Michael	1-151		
Kocaturk, Tuba	1-677		
Koch, Ethan	1-655		
Koh, Immanuel	1-465		
Kohler, Matthias	2-111		
König, Reinhard	1-273, 1-545		
Krezlik, Adrian	2-619		
Krishnamurti, Ramesh	1-333		
<b>L</b>			
Ladron de Guevara, Manuel	1-333, 1-373		
Langenhan, Christoph	1-323, 1-403, 2-435		
Lang, Werner	2-629, 2-679		
Larkin, Nicole	2-709		
Latteur, Pierre	2-151		
Lee, Alric	2-131		
Lee, Ching-Han	2-201		
Lee, Hyunsoo	1-19		
Leitão, António	2-689		
Leong, Siew Leng	2-505		
Leung, Carson Ka Shut	1-293, 1-495, 2-11		
Li, Andre	2-567		
Liang, Yuebing	1-425		
Li, Biao	1-625		
Li, Ce	1-747, 2-233		
Li, Cong	2-233		
Li, Haimiao	1-565		
Li, Kan	1-757		
Lim, Chor-Kheng	1-445		
Lin, Alexander	2-365		
Lin, Han-Ting	1-667		
Lin, Jinru	1-213		
Lin, Yinshan	1-79		
Lin, Yuxin	1-343		
Li, Shuyang	1-79		
Liu, Jia	2-31		
Liu, Liquan	1-757		
Liu, Nuozhi	1-465		
Liu, Sijie	2-405		
Liu, Xiao	1-213		
Liu, Xinyu	1-485		
Li, Wenjing	1-243, 1-565		
Li, Yuanyuan	1-141		
Li, Yuke	1-435		
Li, Zhixian	1-707		
Li, Ziyang	2-71		
Lo, Chun-Yu	1-99		
Lok, Leslie	2-121		
Lopez Rodriguez, Alvaro	2-181		
Lo, Tian Tian	1-757, 2-223, 2-425		

- Luo, Dan 1-1, 1-243,  
1-565,  
1-595, 2-1,  
2-161
- Luo, Yang 2-91
- Lu, Xuanyu 2-31
- Lu, Yen-Cheng 2-243
- Lv, Xueyuan 2-517
- M**
- Makki, Mohammed 1-161,  
1-181, 1-655
- Marengo, Mathilde 2-577
- Markopoulou, Areti 2-577
- Marsatyasti, Naya 2-475
- Marschall, Max 2-639
- Mathers, Jordan 1-161
- Matisziw, Timothy C 1-595
- Mayer, Hannes 2-111
- Mayer, Wolfgang 2-587
- May, Kieran 2-527
- Mazza, Domenico 1-677
- McMeel, Dermott 2-415
- Menges, Achim 1-393, 2-41,  
2-81, 2-141
- Meng, Leo Lin 1-253
- Meral Akgul, Cagla 2-669
- Mete, Burak 1-323
- Miao, Junyi 2-233
- Miller, Clayton 1-525
- Milovanovic, Julie 2-315
- Min, Deedee 1-201
- Miyaguchi, Mikita 2-131
- Mondal, Tushar 2-253
- Moscovitz, Or 1-171
- Mosteiro-Romero, Martin 1-525
- Mo, Yichen 1-625
- Murata, Ryo 2-475
- N**
- Naboni, Roberto 2-375
- Nagakura, Takehiko 1-69
- Nagamachi, Shiho 1-737
- Nakajima, Tadahiro 2-111
- Nakamura, Hoki 1-737
- Nanasca, James 1-39
- Napier, Ilaena Mariam 2-303
- Narahara, Taro 1-11
- Neri, Iacopo 2-577
- Newnham, Cameron 2-191
- Ng, Provides 2-759
- Nguyen-Tran, Khang 2-475
- Nguyen, Tommy Bao 1-243
- Nghi
- Ni, Eryu 2-161
- Nisztuk, Maciej 1-313, 1-435
- O**
- Ochoa Paniagua, Jorge 2-587
- Odaibat, Baha 2-759
- Oghazian, Farzaneh 1-111
- Oki, Takuya 2-475
- Onishi, Ryo 2-395
- Oprean, Danielle 2-455
- Ortner, F. Peter 1-191
- P**
- Pacher, Matteo 2-111
- Pantic, Igor 2-181
- Panya, David Stephen 2-597
- Pan, Yongjie 2-729
- Park, Hyejin 2-597
- Park, Hyoung-June 1-697
- Patel, Sayjel Vijay 1-313, 1-435
- Pawar, Siddharth Suhas 2-101, 2-171
- Pebryani, Nyoman 1-151
- Perrault, Simon 1-383
- Perry, Gabriella 2-283
- Petrovic, Emina K. 2-415
- Petzold, Frank 1-635,  
2-435, 2-557
- Pintacuda, Luigi 1-615
- Poranne, Roi 1-49
- Q**
- Qiu, Song 2-101
- Qiu, Waishan 1-243,  
1-425, 1-565
- R**
- Raanan, Noam 1-283
- Raghu, Deepika 2-577
- Raja, Twisha 2-51
- Rameezdeen, Rameez 2-587
- Reinhardt, Dagmar 2-649,  
2-709, 2-749
- Reitberger, Roland 2-629, 2-679

- Rezaei Rad, Aryan 2-151  
 Rhee, Jinmo 1-373  
 Ricafort, Kim 1-655  
 Riggio, Mariapaola 2-345  
 Robinson, Richard 1-697  
 Rogeau, Nicolas 2-151  
 Roupé, Mattias 1-29  
 Rust, Romana 1-49  
 Ruszczewski, Szymon 1-707
- S**  
 Sakai, Yasushi 1-585  
 Santos, Luis 2-689  
 Sateei, Shahin 1-29  
 Schmidiger, Robin 1-49  
 Schneidman, Alexander 1-333  
 Schubert, Gerhard 1-635  
 Schumann, Kyle 2-355  
 Schwinn, Tobias 2-81  
 Sepulveda, Pablo 2-639  
 Settimi, Andrea 2-325  
 Shao, Tong 2-699  
 Sheng, Yu-Ting 2-213  
 Sheth, Urvi 1-1  
 Shibuya, Masako 2-131  
 Shimizu, Shunta 1-737  
 Shi, Xing 2-537  
 Shi, Zhongming 1-555  
 Silveira, Sue 2-709  
 Silvennoinen, Heidi 1-555  
 Singh, Mayank 2-659  
 Skoury, Lior 2-141  
 Smith, Ross 2-527  
 Someya, Syunsuke 2-131  
 Sopher, Hadas 2-315  
 Sprecher, Aaron 1-223  
 Stark, Tim 2-41  
 Stieler, David 2-81  
 Stouffs, Rudi 1-525, 2-385  
 Stuart-Smith, Robert 1-131  
 Subramanian, Ramanathan 1-313, 1-435  
 Sukegawa, Chika 2-131  
 Sun, Chengyu 1-79  
 Sun, Ke Nan 2-223  
 Szabo, Anna 2-61
- T**  
 Tabi, Salma 1-585  
 Taima, Masahiro 1-585  
 Tang, Chongyi 2-161  
 Tang, Ming 1-575  
 Tang, Peng 1-625  
 Tang, Sheng Kai 2-445  
 Tan, Ying Yi 2-91  
 Tay, Jing Zhi 1-191  
 Thomas, Bruce 2-527  
 Tian, Jing 1-575  
 Tohidi, Alex 2-739  
 Tong, Ziyu 2-719  
 Toohey, Gabrielle 1-243  
 Tracy, Kenneth 2-91  
 Tsai, Tsung-Han 2-243  
 Tsubata, Shinya 2-111  
 Tunçer, Bige 1-383  
 Tung, Nguyen 1-585  
 Turchi, Tommaso 1-615
- V**  
 Vaez Afshar, Sepehr 1-69  
 Vaidhyanathan, Vishal 2-51  
 van Ameijde, Jeroen 1-1, 1-293, 1-485, 1-495, 2-1  
 van den Berg, Nick 2-191  
 Varinlioglu, Guzden 1-69  
 Veloso, Pedro 1-373  
 Vermisso, Emmanouil 1-353, 2-335  
 Vestartas, Petras 2-151, 2-325  
 Vilppola, Ritva 1-243  
 Voltolina, Marco 1-717  
 von Richthofen, Aurel 1-555
- W**  
 Wagner, Hans Jakob 2-41, 2-141  
 Walsh, James 2-527  
 Wang, Hanmo 2-365  
 Wang, Julian 1-575  
 Wang, Ke 2-425  
 Wang, Likai 1-415, 2-699  
 Wang, Qiang 2-131  
 Wang, Shih-Yuan 2-213, 2-243  
 Wang, Sihan 2-91  
 Wang, Sining 2-405  
 Wang, Siqi 1-213

- |                    |                       |                        |                             |
|--------------------|-----------------------|------------------------|-----------------------------|
| Wang, Tsung-Hsien  | 1-535                 | Yoffe, Hatzav          | 1-283                       |
| Wang, Xiang        | 2-517                 | Yousif, Shermeen       | 1-353,<br>1-363, 2-335      |
| Watanabe, Yoshiaki | 2-131                 | Yuan, Philip F.        | 2-41, 2-71,<br>2-273, 2-517 |
| Weijenberg, Camiel | 1-313, 1-435          | Yu, Chuan              | 2-101, 2-171                |
| Weinand, Yves      | 2-151, 2-325          | Yu, Daniel             | 1-253,<br>1-687, 2-739      |
| Wei, Ziru          | 2-405                 | Yu, Junah              | 1-201                       |
| White, Marcus      | 1-455, 1-645          | <b>Z</b>               |                             |
| Won, Junghye       | 2-597                 | Zahedi, Ata            | 2-557                       |
| Wood, Dylan        | 1-393                 | Zangori, Laura         | 2-455                       |
| Wortmann, Thomas   | 1-393                 | Zanini, Michele        | 2-61                        |
| Wu, Hao            | 2-71                  | Zhang, Garry Hangge    | 1-253                       |
| Wu, Jinxuan        | 2-223                 | Zhang, Gengjia         | 1-141, 1-233                |
| Wu, Wanling        | 2-233                 | Zhang, Hong            | 2-567                       |
| Wu, Xinyu          | 2-71                  | Zhang, Huikai          | 2-101                       |
| Wu, Zihao          | 2-719                 | Zhang, Qiyan           | 1-625                       |
| <b>X</b>           |                       | Zhang, Shuyu           | 1-515                       |
| Xiang, Rufeng      | 1-757                 | Zhang, Tong            | 2-729                       |
| Xiao, Yahan        | 1-303                 | Zhang, Yunsong         | 2-719                       |
| Xin, Zhuoyang      | 2-161                 | Zhang, Zihuan          | 1-747                       |
| Xu, Hang           | 1-535                 | Zhao, Yucheng          | 1-121                       |
| Xu, Jiaqi          | 1-213                 | Zheng, Hao             | 1-475                       |
| Xu, Ke             | 1-425                 | Zheng, Lang            | 2-273                       |
| Xu, Yijia          | 1-213                 | Zheng, Zifei           | 2-233                       |
| Xu, Zhitao         | 1-181                 | Zhou, Margaret Z.      | 1-505                       |
| Xu, Zhiyan         | 2-233                 | Zhou, Xinjie           | 2-71                        |
| <b>Y</b>           |                       | Zhou, Xinyan           | 2-233                       |
| Yabe, Taisei       | 2-31                  | Zhou, Xinyi            | 1-757                       |
| Yabuki, Nobuyoshi  | 1-89, 2-395,<br>2-607 | Zhou, Yifan            | 2-273                       |
| Yamauchi, Yuji     | 1-737                 | Zhou, Ziqi             | 2-517                       |
| Yang, Qing         | 1-565                 | Zhuang, Dian           | 2-537                       |
| Yang, Xiliu        | 2-141                 | Zhu, Guanqi            | 2-161                       |
| Yang, Xuyou        | 1-121                 | Ziegler, Christoph     | 1-403, 2-435                |
| Yan, Wei           | 2-547                 | Zomparelli, Alessandro | 2-375                       |
| Yan, Xin           | 1-121                 |                        |                             |
| Yao, Jiawei        | 1-141, 1-233          |                        |                             |
| Yin, Minggang      | 1-233                 |                        |                             |