Cities 156 (2025) 105473

Contents lists available at ScienceDirect

Cities

journal homepage: www.elsevier.com/locate/cities

A perception-powered urban digital twin to support human-centered urban planning and sustainable city development

Junjie Luo [a,b], Pengyuan Liu [c], Wenhui Xu [b], Tianhong Zhao [d], Filip Biljecki [e,f,*]

[a] Zhejiang Province Key Think Tank: Institute of Ecological Civilization, Zhejiang A&F University, Hangzhou, China
[b] College of Landscape Architecture, Zhejiang A&F University, Hangzhou, China
[c] Future Cities Lab Global, Singapore-ETH Centre, Singapore
[d] College of Big Data and Internet, Shenzhen Technology University, Shenzhen, China
[e] Department of Architecture, National University of Singapore, Singapore
[f] Department of Real Estate, National University of Singapore, Singapore

ARTICLE INFO

Keywords:
Urban digital twin
Semantic segmentation
Virtual reality
Visual perception
Random forest

ABSTRACT

Urban Digital Twins (UDTs) offer a promising avenue for advancing sustainable urban development by mirroring physical environments and complex urban dynamics. Such technology enables urban planners to predict and analyze the impacts of various urban scenarios, addressing a global priority for sustainable urban environments. However, their potential in public engagement for environmental perception remains unfulfilled, with existing research lacking the capability to analyze urbanscapes' visual features and predict public perceptions based on photo-realistic renderings. To fill the gap, our study developed and implemented a UDT platform designed for the dual purposes of objective feature evaluation and subjective visual perception, alongside the prediction of perceptions in simulated scenarios. We incorporated DeepLabV3, a deep learning model for imagery semantic segmentation, to quantify a series of visual features within the built environment, such as the proportion of vegetation and architectural elements. Subjective visual perceptions (e.g. safety and lively) are captured using immersive virtual reality to gather public perceptions of different scenarios and learn patterns. Further, utilizing a photo-realistic rendering engine, high-quality renderings of textures and materials for UDT were achieved, and we proved their veracity based on a perception experiment. Afterwards, we employ the random forest algorithm for automated perception predictions of rendering scenarios. The implementation was demonstrated with a case study on an urban greenway in the central area of Singapore. We compared both the objective evaluation and subjective perception results, followed by a demonstration of automated visual perception prediction through photo-realistic scenario simulations, such as modifying vegetation density or introducing new architectural elements to the skyline, to predict the perception of scenarios before they are built, leading to more efficient and automated urban planning.

1. Introduction

Sustainable urban development is increasingly recognized as a global imperative, calling for innovative solutions to navigate the complex dynamics of contemporary urban environments (Son et al., 2023). In this evolving landscape, Urban Digital Twins (UDTs) have emerged as transformative tools, simulating and analyzing urban scenarios to offer nuanced insights into city functioning (Haraguchi et al., 2024; Ivanov et al., 2020; Lei, Janssen, et al., 2023). Originally rooted in aerospace, the concept of Digital Twins has broadened its reach, proving instrumental across various domains, including urban management, thereby setting the stage for a new paradigm in urban planning and development (Deng et al., 2021; Elkefi & Asan, 2022; Francisco et al., 2020). A digital twin includes the detailed replication and virtual representation of objective physical entities, thereby facilitating the establishment of a mirrored mapping relationship between the virtual and the physical (Jafari et al. Wang & Vu, 2023; Weil et al., 2023; Li et al., 2023; Tao & Qi, 2019). Ideally, this relationship is bidirectional, leading to decision-making and interventions in the real world (Luo, Liu, & Cao, 2022; Weil et al., 2023). Currently, UDTs can facilitate scientific analysis, virtual simulation, predictive inferences, iterative optimization of physical entities, and promise automation for testing urban scenarios and

* Correspondence to: F. Biljecki, Department of Architecture, National University of Singapore, Singapore.
E-mail addresses: luojunjie@zafu.edu.cn (J. Luo), filip@nus.edu.sg (F. Biljecki).

https://doi.org/10.1016/j.cities.2024.105473
Received 23 May 2024; Received in revised form 20 July 2024; Accepted 30 September 2024
Available online 21 October 2024

translating insights into real-world actions (Bauer et al., 2021; Botín-Sanabria et al., 2022; Jiang, Ma, et al., 2022; Liu et al., 2021; Tzachor et al., 2022). UDTs and related technologies can provide a deeper understanding of various issues in the built environment (Bolívar et al., 2023; Hamada et al., 2024; Piras et al., 2024), with visual perception being a particularly significant aspect (Liang et al., 2023). The visual perception and experience of residents regarding urban environments not only directly impact their satisfaction with urban spaces but also profoundly influence their sense of happiness and willingness to participate in community activities (Wang et al., 2024; Yosifof & Fisher-Gewirtzman, 2024). Therefore, it is crucial to gain a deep understanding of residents' visual perceptions and integrate them comprehensively in urban planning, which can help to create a more responsive, human-centered, and sustainable urban environment (Li, Wang, et al., 2022; Liu et al., 2023; Wang & Vu, 2023).

However, in the expanding realm of UDTs, a significant aspect that has yet to be thoroughly explored is the complex relationship between urban landscape characteristics and the subjective perceptions of their inhabitants through the lens of UDTs (Weil et al. Wang & Vu, 2023; Al-Sehrawy et al., 2023). To date, the potential of UDTs to evaluate and predict human subjective perceptions in response to urban environmental changes has been tapped scarcely. Current literature, while rich in exploring the objective evaluation of built environments, seldom intersects with the subjective visual perceptions that fundamentally influence residents' experience of urban spaces (Adade & de Vries, 2023; Kikuchi et al., 2022). With the integration of human perception into the UDTs, simulations can account for factors like field of view area, height-to-width ratio of street, green view ratio and so on, providing a more holistic view of how residents experience and interact with their urban environment. Recognizing this, our study seeks to address the existing research gap by focusing specifically on human visual perception within the context of UDTs. The integration involves the use of advanced algorithms such as computer vision and parametric spatial analysis methods, which facilitate the efficient evaluation of spatial features of the built environment, thereby supporting urban scientific management and aiding decision-making processes (Liu et al., 2023; Mazzoli et al., 2023; Ramu et al., 2022). In addition to integrating spatial features evaluation, UDT-related technologies also allow for simulation modelling, optimization of the built environment (Deng et al., 2021; Lu et al., 2023), and have the potential to automatically predict human visual perceptions with the support of artificial intelligence and photo-realistic renderings (Xu et al., 2022). The integration of human perception into UDTs paves the way for simulations that reflect real-world experiences more closely, enabling urban planners to predict and mitigate potential issues before they manifest in physical form. Through this innovative methodology, we can provide a comprehensive understanding of how urban environments are perceived and experienced by individuals.

Heading in this direction, the precise 3D models (role as a 'digital base') of urban environments that contain multiple landscape elements (e.g. buildings, vegetation, and fences) are essential (Lei, Stouffs, & Biljecki, 2023; Li, Xue, et al., 2022; Najafi et al., 2023). In addition, panoramic imagery that can be used for virtual perception is also one of the important data sources for achieving urban visual perception (Kang et al., 2020; Liu et al., 2023; Luo, Zhao, et al., 2022b). Panoramas from crowdsourcing platforms and map enterprises can also serve as the input data for UDTs (Luo, Zhao, et al., 2022c). Therefore, the practice of continually updating these panoramas, whether in real-time or at regular intervals, serves to maintain the relevance and accuracy of the UDTs, ensuring they reflect the current state of the physical environment accurately. Simultaneously, there has been a growing emphasis on public-friendly participation in visual perception research, especially within the academic lifecycle of DTs (Dembski et al., 2020; Luo, Shi, et al., 2022; Schrotter & Hürzeler, 2020). This focus not only promotes DT technology but also enhances the accessibility to the general public (Lei, Janssen, et al., 2023). Moreover, developing user-friendly modes of visual perception in the cyber counterpart is widely acknowledged as a prerequisite for achieving human-(virtual) environment interaction and fostering public engagement (Luo, Liu, & Cao, 2022). To achieve such a goal, existing research has demonstrated how the integration of immersive virtual perception devices, such as VR headsets, into UDT platforms can significantly enhance public participation (Kuru, 2023). By providing an immersive experience that closely mimics the physical environment, these technologies allow users to engage deeply with potential urban planning scenarios, offering valuable insights into public preferences and perceptions (Dang et al., 2023; Xu et al., 2022). This holistic approach, which combines detailed 3D modelling with immersive user experiences, exemplifies the evolving landscape of urban planning. Digital innovation now facilitates a more inclusive and interactive process, aligning technological advancements with the nuanced needs of urban residents.

Additionally, a theoretical understanding of how individuals process and evaluate environmental information is paramount, as it provides a structured basis to cohesively guide relevant research (Wu et al., 2024). Mindsponge Theory, proposed by Vuong (2023), offers a comprehensive framework elucidating the dynamic processes of information absorption, filtering, and integration. This theory conceptualizes the human mind as a sponge, selectively absorbing information from the urban environment based on its perceived relevance and value to the individual. Adopting Mindsponge Theory enables us to establish a robust approach for analyzing and evaluating urban landscape characteristics through UDTs with the help of virtual reality technology. Such a theoretical foundation enhances our comprehension of how processed environmental data translates into subjective evaluations, such as perceived safety, beauty, and liveliness, informing and shaping sustainable urban planning initiatives.

Therefore, in this study, we introduce a human-centric UDT concept, emphasizing the crucial role of incorporating public visual perceptions into the digital representation of urban environments. We adopt the Mindsponge Theory as the foundation for understanding how individuals process and evaluate environmental information. The main objective of this study is to weave human visual perceptions into the fabric of UDTs to analyze and evaluate the characteristics of urban landscapes. By combining UDT with human-computer interaction technology (e.g. virtual reality), this study achieves the subjective visual perception of cityscapes and complements the ongoing developments in digital twins that are focused primarily on physical and non-human aspects, such as wind flow simulations. Furthermore, we utilize DeepLabV3, a deep learning architecture, for imagery semantic segmentation and Grasshopper parametric spatial evaluation methods in the UDT platform for objective feature analysis and employ the random forest algorithm for automated perception prediction, providing data support and a decision-making basis for environmental analysis. Building upon this, we illustrate the viability of replacing photographs with photo-realistic renderings, enabling the prediction and testing of diverse scenarios prior to actual implementation. The primary application of the UDT is manifested in simulating various urban scenarios, such as modifying vegetation density or introducing new architectural elements to the skyline. This study, therefore, contributes a unique perspective to the discourse on digital twins, emphasizing the indispensable role of human perception in creating cities that are not only sustainable but profoundly human-centric and attractive to their residents and visitors.

2. Background and related work

2.1. The role of digital twins in sustainable urban development

Sustainable urban development has emerged as a critical concept in addressing the environmental and social challenges posed by rapid urbanization (Foroozesh et al., 2022; Zhang, Han, et al., 2022). In this context, DTs are widely regarded as the primary driver of the sustainable transition within human settlements, offering unprecedented capabilities to model, analyze, and optimize the multifaceted dynamics of urban environments (Tzachor et al., 2022; Weil et al., 2023). However, the concept of DT lacks a universally accepted definition or a blueprint for its construction (Lehtola et al., 2022; Xia et al., 2022), leading to the emergence of a diverse array of interpretations that vary in focus and scope according to their specific application contexts. It is generally acknowledged that UDTs do not constitute perfect replicas - the idea of an exact counterpart remains an aspirational ideal rather than a tangible goal. The endeavor to incrementally refine digital models to more closely approximate their real-world counterparts fundamentally underpins the development of computerized simulations. The application of UDTs in urban development transcends traditional planning methods by emphasizing the importance of environmental sustainability. Through the detailed analysis of environmental data, such as air quality, water management, and green spaces, UDTs support the design of cities that not only meet the current needs of their inhabitants but also safeguard the environment for future generations. This aligns with the principles of sustainable urban development by promoting a balance between economic development, social inclusion, and environmental protection.

Digital twin models serve as the foundation for simulating different urban scenarios, ranging from infrastructure development to disaster response strategies (Adade & de Vries, 2023; Hu et al., 2023; Jiang, Ma, et al., 2022). By providing a holistic view of city functioning, DTs allow urban planners and decision-makers to visualize the consequences of their choices before implementing them in the real world, ensuring that urban development is both sustainable and aligned with the community's needs (Lei, Su, & Biljecki, 2023; Tancev & Toro, 2022; Trantas et al., 2023; White et al., 2021). Digital Twin models necessitate the simulation and analysis of physical environmental characteristics as well as mobile objects within urban settings on suitable software platforms. Corporations, including General Electric and Siemens, have developed Digital Twin platforms, while software platforms like Rhino-Grasshopper and ArcGIS are capable of either fully or partially supporting applications related to digital twins (Jafari et al., 2023; Kenett & Bortman, 2022; Moyano et al., 2021). The applications of DT platforms in urban contexts are manifold and transformative. For instance, Lee et al. (2022) exemplifies a DT platform designed for individual mobility objects. The DT platform diverges from previous smart city projects by focusing on the data objectification, analysis, and compression of individual pedestrians or vehicles, offering a nuanced approach to urban mobility management. It is capable of detecting individual mobility data and rendering 3D visualizations of UDT models, thereby providing an unprecedented level of detail in the monitoring and analysis of urban mobility. Similarly, Xu et al. (2022) showcases a DT platform designed for on-road emission monitoring. The proposed platform integrates a VR-based digital environment with a micro-simulation model for background traffic and utilizes the motor vehicle emission simulator for accurate emission estimations. This research not only demonstrates the applicability of DT technology in addressing complex urban challenges but also highlights its role in fostering more sustainable urban environments through improved emission monitoring and management practices. Moreover, methodologies for data collection and transmission play a pivotal role in the effectiveness of DTs (Botín-Sanabria et al., 2022). Beyond the traditional monitoring of urban physical parameters through IoT technologies, such as water temperature, soil moisture, and water clarity (Barth et al., 2023; Tancev & Toro, 2022; Trantas et al., 2023), imagery captured by an extensive network of urban surveillance cameras has emerged as an indispensable source for digital twins (Ivanov et al., 2020). The application of machine learning and deep learning algorithms underpins sophisticated multi-dimensional predictions and hazard perception analyses within urban environments (AlBalkhy et al., 2024; Li et al., 2023; Sai et al., 2024). This seamless integration from the initial stages of data collection and analysis through to the application of advanced algorithms for the predictive exploration of urban environmental dynamics is what truly transforms digital twins into entities of 'intelligence'.

The ability of DTs to simulate the long-term impacts of urban policies ensures that sustainability considerations are embedded in the planning process from the outset (Lei, Janssen, et al., 2023; Wang & Wang, 2024). By enabling precise analysis and visualization of urban scenarios, DTs facilitate informed decision-making that prioritizes environmental sustainability, economic viability, and social equity (Barth et al., 2023; Chen et al., 2024). By offering detailed insights into the complex interplay of urban systems, DTs empower cities to navigate the challenges with agility and foresight.

2.2. Integrating human visual perception in UDTs for environment optimization

The advancement of DT-related technologies has not only revolutionized urban planning but also expanded the horizons of what can be achieved in simulating and optimizing urban scenarios (Elnabawi & Raveendran, 2024; Geremicca & Bilec, 2024). Traditionally, DTs have focused on replicating physical infrastructure and operational dynamics with high precision (Boccardo et al., 2024). However, to truly transform urban living, it is imperative to incorporate human perception into these digital replicas. The integration of human visual perception becomes crucial to enhance the realism and effectiveness of UDTs in replicating real-world urban environments (Luo, Liu, & Cao, 2022; Yossef Ravid & Aharon-Gutman, 2023). By understanding and simulating how people perceive and interact with their surroundings, UDTs can provide a more holistic approach to the improvement of cities. Therefore, it is necessary to delve into the evolving landscape of UDTs applications, with a specific focus on the integration of visual perception for the optimization of urban scenarios.

Recognizing the limitations of traditional 3D models or DTs, we emphasize the crucial role of incorporating human perception, especially through the use of virtual reality (VR) or augmented reality (AR), into these platforms (Adrienne & Nora, 2024; Argota Sánchez-Vaquerizo et al., 2024; Lehtola et al., 2022). In recent years, the integration of DT with VR has emerged as a powerful approach in the field of urban planning and landscape analysis (Attaran & Celik, 2023; Bevilacqua et al., 2022; Zhang, Wen, et al., 2022). VR headsets and immersive environments allow urban planners, architects, and researchers to step into the digital spaces, exploring urban landscapes from multiple perspectives, including ground-level and bird's-eye views (Luo, Zhao, et al., 2022b). This immersion significantly enhances their ability to perceive and understand complex urban configurations and spatial relationships (Rowland et al., 2022), and can also offer dynamic and immersive experiences. Specifically, urban planners and decision-makers can visualize the impacts of proposed changes, such as new construction projects or infrastructure developments, in a realistic 3D environment (Adrienne & Nora, 2024; White et al., 2021). This dynamic simulation aids in assessing how these changes might affect landscape characteristics and the human experience within the urban environment (Shahat et al., 2021). Simultaneously, the integration of digital twins with VR or AR holds tremendous potential for engaging the public in urban planning, design, and management (Dembski et al., 2019; Künz et al., 2022). Citizens can virtually explore proposed urban designs and developments, effectively 'walking through' their city's futurescapes (Yin et al., 2023). For example, the city's planning department can employ VR to create an immersive experience for the green space simulation (Schrotter & Hürzeler, 2020). Residents can wear VR headsets to take virtual strolls through these environment, experiencing the layout, landscaping, and amenities before construction begins (White et al., 2021). These technologies also promote inclusivity by reaching a broader audience when residents access the virtual interface through web-based VR or AR tools from their smartphones and other digital devices (Ham & Kim, 2020; La Guardia & Koeva, 2023; Sepasgozar, 2020). This approach ensures participation from a wider demographic, including individuals who might have difficulty attending physical meetings, thereby empowering residents to effectively voice their preferences and concerns. These immersive technologies offer innovative ways to involve citizens in shaping their cities while contributing to more sustainable and human-centric urban development (Lv & Fridenfalk, 2023).

In addition to the visual representation of scenarios and intuitively showcasing the outcomes of planning schemes, UDTs can also engage in visual perception prediction (Liu et al., 2023). In this study, we aim to push the boundaries of traditional UDT applications by integrating cutting-edge visualization techniques and advanced predictive modelling. Utilizing machine learning algorithms (e.g. random forest), UDTs could assist planners in anticipating individuals' subjective visual experiences in different scenarios, thereby providing a scientific basis for planning and design. However, to the best of our knowledge, research that has fully achieved this objective remains scarce. Gaining inspiration from the technology fusion of VR and UDTs and to address the aforementioned gaps, the key contributions of this study are a new UDT constructed to integrate human perception, spatial evaluation and machine learning algorithms for landscape feature analysis and automatic perceptual prediction associated with photo-realistic rendering scenarios, a new use case of UDT for urbanscapes perception and betterment, and one that has the underlying capability to monitor and optimize the futurescapes, a rarity in the literature and a clear gap.

3. Methodology

3.1. Perception-guided urban digital twin platform design

According to the Mindsponge Theory, information processing involves several critical steps: information reception, filtering, interpretation, and integration (Vuong, 2023). Previous research has divided the platform of urban digital twins into multiple hierarchical layers to better support environmental analysis, information processing and visualization (Luo, Liu, & Cao, 2022; Weil et al., 2023; White et al., 2021; Xu et al., 2022). This study focuses on how urban digital twins collect sensory data (e.g., visual stimuli) that users process to form perceptions. The process begins with the reception of raw environmental data, followed by filtering based on relevance and personal significance, and ends with integrating this processed information into people's cognitive framework. Therefore, the UDT platform designed and developed in this study similarly comprises five subsystems (or layers), including the 3D model, data collection, visual characteristics analysis, immersive virtual reality, and perception analysis and prediction subsystem (Fig. 1), to mimic the steps in Mindsponge Theory. These five subsystems correspond to distinct functionalities and employ different key methodologies.

1. 3D model subsystem. Our study referenced the composition of relevant UDT models (White et al., 2021) and, in accordance with our research aims, defined six physical layers of content within this subsystem. These layers include the terrain, water, buildings, roads and squares, vegetation, and facilities. The 3D model was finalised manually, referring to the high-quality UAV oblique imagery, and own collected street view imagery. The model was built on the Rhinoceros and Grasshopper platform with parametric spatial evaluation capability, and the textures and materials for 3D models were further rendered and visualized in the Lumion software with the assistance of the high-quality rendering ability (Muhammad et al., 2019).
2. Data collection subsystem. The data collection subsystem encompasses different hardware devices, such as GoPro Max panoramic cameras and unmanned aerial vehicles. Through these devices, panoramic imagery for visual perception and evaluation can be

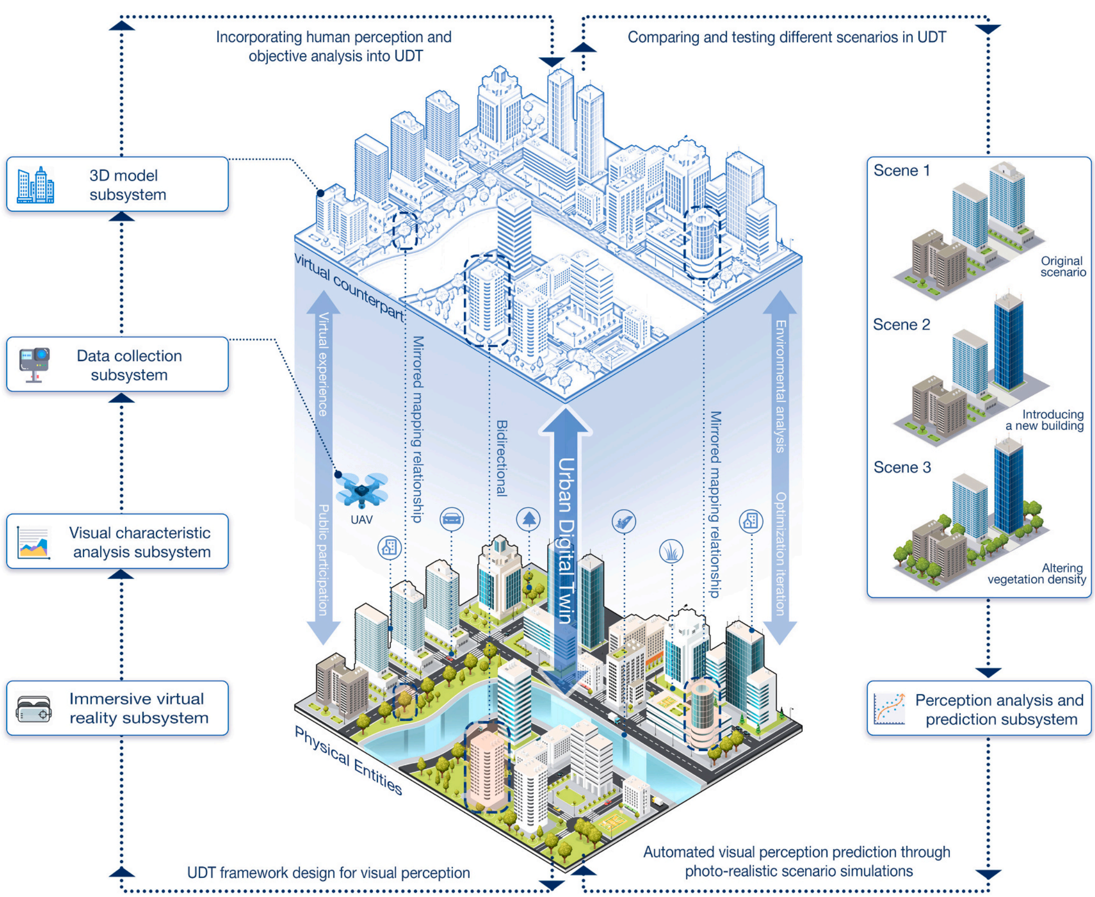


Fig. 1. The urban digital twin platform we developed for assessing and managing visual perception in the built environment.

acquired. This subsystem also includes software programs, such as Mapillary API, enabling the retrieval of publicly available imagery within the study area (Hou & Biljecki, 2022), which has global coverage (Hou et al., 2024). Additionally, the programs are equipped to read GPS information (such as latitude and longitude) in photographs. By extracting GPS information from images, it becomes feasible to pinpoint the locations of different images within the UDT, establishing a correspondence between images and the virtual model. Regularly obtaining imagery of the study area enables near-real-time updates of the UDT for the study area according to the research aim.

3. Visual characteristic analysis subsystem. The Visual characteristic analysis subsystem is an important component in our UDT platform, capable of analyzing the objective visual features of the study area. This subsystem comprises two main aspects, computer vision algorithm for imagery semantic segmentation and Grasshopper parametric algorithms for spatial feature evaluation. The former enables high-precision semantic segmentation and quantitative analysis of panoramas, indicators include the green visibility factor, building visibility factor and so on. The latter can analyze characteristics such as field of view area, field of view perimeter and others.
4. Immersive virtual reality subsystem. This subsystem can facilitate human-environment interaction and virtual perception. It enables participants to attain an immersive visual experience through virtual reality technology, offering a genuine and interactive 3D visual environment. This aids researchers in delving deeper into the visual characteristics of the study area. Head-mounted virtual reality displays (e.g. VR goggles) are the core hardware capable of presenting virtual reality environments to users. The head-mounted VR displays are typically accompanied by additional accessories such as controllers or foot pedals to facilitate user interaction with the virtual reality environment.
5. Perception analysis and prediction subsystem. This subsystem includes the Pearson correlation analysis approach, which can explore the relationship between the objective environmental characteristics (e.g. green visibility factor and field of view area) and the subjective visual perceptions, such as the lively and beauty of a place. It also incorporates the Random Forest (RF) algorithm, which investigates the mechanisms by which objective visual characteristics of urban landscapes influence subjective visual perception results. The RF can also assess the key independent variables and variable importance within the model. Furthermore, it enables the prediction of subjective visual perception in other scenarios.

3.2. Objective characteristics evaluation

Various types of artificial intelligence and spatial evaluation algorithms have been integrated into the UDTs to achieve more intelligent data processing and analysis (Fan et al., 2021; Lehtola et al., 2022; Li et al., 2023). Specifically, imagery semantic segmentation algorithms can assign each pixel in an image to its respective object category, allowing for precise segmentation of different target objects within the image (Biljecki et al., 2023; Chen & Biljecki, 2023; Liang et al., 2023). Integrating computer vision into the UDT enables semantic segmentation of the input imagery. We use DeepLabv3 to obtain the percentages of spatial features in panoramas. DeepLabv3, which can classify 30 categories of urban objects in photos, was pre-trained on the CityScape dataset using PaddleSeg (Yurtkulu et al., 2019). Thus, this imagery semantic segmentation algorithm can classify each pixel in panoramic images of the built environment into object categories accurately, such as the sky, buildings and vegetation. The performance of the pre-trained DeepLabv3 model was evaluated using Mean Intersection over Union (mIoU), which ensures accurate segmentation across different object categories. By quantitatively analyzing the proportions of objects within these image scenes, an assessment of the objective visual characteristics of the landscape can be achieved.

Furthermore, in addition to supporting parametric 3D modelling, the Grasshopper parametric platform can perform various types of spatial features evaluation (Funari et al., 2021; Gholami et al., 2022; Koenig et al., 2020; Luo, Liu, & Cao, 2022). Examples include automatically quantifying street widths, measuring building heights, and ranking them to analyze the height-to-width ratio characteristics of the study area. Additionally, on the Grasshopper platform, we quantify the size of the visible area in different locations of the study area and evaluate the field of view perimeter, which plays a crucial role in evaluating the spatial characteristics of the urban landscape environment.

We adopt 8 indexes for the objective evaluation, drawing from the previous research experience on the assessment of visual and spatial characteristics of the built environment, primarily streetscapes and open spaces (Chen & Biljecki, 2023; Gong et al., 2018; Liu et al., 2023; Ma et al., 2021; Tabrizian et al., 2020). Among these, 5 indicators come from imagery semantic segmentation outcomes, including the green view factor (GVF), hard-ground view factor (HVF), building view factor (BVF), sky view factor (SVF), and dynamic factor (DF), which are well established metrics (Biljecki et al., 2023; Wang & Vu, 2023). Each indicator is constructed from specific semantic classes in the Cityscapes dataset. GVF is composed mainly of the 'vegetation' class, quantifying the visible green area within the urban environment. HVF indicator is derived from the 'road' and 'sidewalk' classes, representing the visible hard-ground surfaces. BVF indicator utilizes the 'building' class to assess the visibility of built structures within the scene. SVF is composed of the 'sky' class, this indicator measures the visible sky area, providing an understanding of the openness of the urban space. DF indicator is calculated from the 'car,' 'truck,' 'bus,' 'motorcycle,' 'bicycle,' and 'caravan' classes, representing the dynamic elements. The other 3 novel indicators come from spatial analysis outcomes utilizing the Grasshopper parametric analysis algorithms, facilitating the quantitative analysis of spatial characteristics in the study area, including the field of view area (FVA), the field of view perimeter (FVP), and the height-to-width ratio (HWR) of building vertical facade and flat surface.

3.3. Human-in-the-loop visual perception

The human-in-the-loop experiment was conducted to obtain the subjective visual perception outcomes by utilizing the immersive virtual reality subsystem of the UDT platform. In this study, the Pico Neo 3 VR goggles were used for immersive virtual perception experiments, which can obtain the subjective visual perception results from participants (Taupin et al., 2023). By referencing the number of participants in similar studies related to visual perception using virtual reality (Birenboim et al., 2019; Luo, Shi, et al., 2022; Luo, Zhao, et al., 2022b), this study, taking into account the research aims, recruited a total of 120 participants. Among these participants, 94 had professional backgrounds in fields such as landscape architecture and urban planning. They primarily consisted of undergraduate, graduate students, and PhD candidates, as well as university faculty members. Additionally, to enhance participant diversity, our study also included a smaller group of non-professionals, totalling 26 individuals. Thus, professionals accounted for approximately 78.3 % of the study's participants. In this study, all participants were randomly divided into two groups, each consisting of an equal number of 60 individuals. The first group had an average age of 28.5 years, with 61.7 % ($n = 37$) being female, and 91.7 % ($n = 55$) having prior experience wearing VR headsets. The second group of participants had an average age of 27.3 years, with 56.7 % ($n = 34$) being female, and 85 % ($n = 51$) having prior experience wearing VR goggles. The study protocol was approved by the Institutional Review Board of the National University of Singapore (reference code: NUS-IRB-2022-191).

In developing our study's methodology for evaluating urban environmental perceptions, we drew inspiration from established research in the field. Specifically, we adopt the methodology of Place Pulse 2.0, notable for its comprehensive approach to quantifying human perceptions across six dimensions: safety, liveliness, boredom, wealth, depression, and beauty (Dubey et al., 2016). This project stands out due to its extensive dataset and its contribution to understanding urban environmental perception, making it one of the most recognized cases in urban perception research. We chose these six perceptual indicators due to their established relevance in reflecting people's perceptions of urban environments, as supported by prior studies. These indicators have been shown to effectively capture both positive and negative aspects of urban experiences, providing a balanced view of environmental perception (Liang et al., 2023; Wang et al., 2019; Zhang et al., 2018). We use immersive virtual reality devices to observe the scenarios of the imagery from all angles and assess them, using the same perceptual indications as this sort of urban scene. The two groups of participants underwent an identical visual perception process, used the same equipment, and employed the same rating method (using the 7-point Likert scale). Each participant only evaluated 20 scenarios, lasting 20 to 25 min each, due to the possibility of visual fatigue brought on by wearing VR glasses for extended periods.

Following each session of viewing panoramic images, participants were required to provide perceptual perceptions of various aspects of the scenes, such as scenic beauty and perceived safety. Upon completing the visual perception tasks, we conducted a correlation analysis of the rating results from the two groups of participants. This analysis was performed to assess the consistency and reliability of the experimental outcomes. If a high level of correlation was observed between the results of the two groups, it could be inferred that the experimental results exhibited good consistency and test-retest reliability. At the same time, participants also rated the feasibility of substituting photos with photo-realistic renderings based on the Likert 7 scale form (strongly disagree (1) to strongly agree (7)), results could demonstrate their acceptance level of the photo-realistic renderings' realism regarding the UDT visual simulations.

3.4. Prediction model construction

The most common means to gauge a linear correlation is through the Pearson correlation approach (Zhang & Zeng, 2024). For processing data that defy the hypothesis when employing four or more Likert scale categories, Pearson correlation is relatively stable, making it appropriate for the correlation's verification (Norman, 2010). Therefore, we assessed the Pearson correlation coefficient between objective characteristics and subjective human perception results.

The Random Forest model is an integrated learning method based on bagging, which enhances predictive accuracy by employing multiple decision trees (Biau, 2012). The fundamental idea behind the RF algorithm is to perform repeated random sampling of training samples and construct decision trees using random features in each sampling iteration. Subsequently, the predictions from multiple decision trees are combined, and this approach effectively mitigates the risk of overfitting. Currently, the RF model has found widespread applications in the field of built environment planning and design (Li, Wang, et al., 2022). Our research used the RF regression algorithm through the Scikit-learn Python library (Pedregosa et al., 2011) for evaluation and prediction. We put the objective visual and spatial characteristics of the research area into the model as the independent variable and human subjective visual perception results (scoring of beauty, safety, lively and so on) as the dependent variable. We utilized the GridSearchCV module to identify the most appropriate parameter set to guarantee the model's accuracy and generalizability. To further reduce the minimal number of sample divisions and leaf node samples, the N-fold approach was also applied. In order to guarantee the model's generalizability, additional parameters, including the maximum feature value, were included. In this study, the 'feature_importances_' attribute was utilized to assess the contribution of each feature to the model's predictive outcomes. Evaluation of the model's performance was accomplished by calculating metrics such as R2, MSE (Mean Squared Error), and MAE (Mean Absolute Error).

To enhance the transparency and interpretability of the RF model, SHAP (SHapley Additive exPlanations) values were employed using the TreeExplainer package (Lundberg et al., 2020). SHAP values quantify the contribution of each feature to the model's predictions by comparing predictions with and without specific features. This approach provides insights into how each feature influences the model's output and helps reveal the interrelationships among features (Iban, 2022; Xu, Xiong, et al., 2023; Zhao et al., 2024). Incorporating SHAP analysis aligns our study with current trends in the literature and strengthens the methodological rigor of our work.

The trained RF model was used for variable importance assessment, enabling the investigation of the relationship between objective landscape features and subjective visual perception scores, as well as facilitating the model's predictive capabilities for other datasets.

4. Study area

We focus on a 3-km-long linear urban greenway located in the heart of Singapore as study area (Fig. 2). Situated adjacent to the Singapore River, this greenway serves as a vital linkage connecting the surrounding residential neighbourhoods with key tourist attractions and recreational spaces, rendering it a critical element of Singapore's urban fabric.

The selection of this specific location as the following considerations underpin the study area. Firstly, Singapore's status as a nation boasting highly advanced digital infrastructure renders it well-equipped to support research endeavours focused on digital twin cities and the pursuit of sustainable urban development. Secondly, Singapore's urban environment presents a microcosm of global urban challenges within a highly dense context. These include the need to balance diverse resident needs and expectations, effectively integrate public perceptions into landscape management, and navigate urban planning complexities in limited spaces. These challenges are not unique to Singapore but are encountered in various forms across global urban landscapes. Thus, solutions and insights derived from Singapore's context have the potential to inform urban sustainability practices worldwide. Thirdly, Singapore is one of the morphologically most diverse cities in the world (Biljecki & Chow, 2022), providing various scenarios of the urban fabric, and it is going through continuous redevelopment, making it a natural focus of this kind of work that focuses on optimizing new developments. Acknowledging the socio-cultural and political uniqueness of Singapore, our study leverages this specificity as a strength. By integrating public perceptions into the urban digital twin, we harness a broad spectrum of insights that reflect a high degree of civic engagement and digital literacy. These insights are crucial for developing inclusive, participatory urban planning processes that can be adapted and applied in different socio-cultural and political contexts. Finally, the choice of this urban greenway aligns with the broader emphasis on promoting sustainable and livable urban environments in Singapore. In essence, while this study area offers a unique case study with its specific characteristics, the core principles and findings of our research are designed to be transferable. We aim to contribute to a global dialogue on the use of digital twin technology in urban planning, demonstrating how digital innovations can enhance sustainability and livability in diverse urban environments. The choice of Singapore underscores our commitment to addressing pressing urban challenges through advanced technology, with the understanding that the methodologies and insights gained here can offer valuable contributions to sustainable urban development efforts globally.

We set evaluation points every 20–40 m in the linear area, with a total of 100 evaluation points in chronological order from east to west. Then, we collected geo-tagged panoramic pictures of these evaluation points for the following analysis utilizing the data collection subsystem in our UDT platform. We extracted GPS information from each imagery and corresponded it with the spatial location of the UDT. Further, we utilized computer vision and parametric spatial quantification analysis algorithms on the Rhino-Grasshopper platform to objectively quantify each location point. The relationship between objective features and

Fig. 2. Study area and survey points with numerical labels. Field panoramic photographs of the two selected mapping points show the surveyed landscapes. (Source of the base map: Google.)

subjective visual perceptions was explored through the Pearson correlation and random forest algorithm. With the assistance of Lumion, the high-quality rendering software, we can render and visualize future scenarios and predict subjective visual perceptions using the RF models. By examining the infusion of public perceptions into the development and management of urban green spaces within a digitally augmented framework, this research seeks to unearth valuable insights that can inform the future planning and design of similar urban greenways (Fig. 3).

5. Results

5.1. The outcome of the urban digital twin

The urban digital twin developed in this study not only includes a 3D model with texture information but also integrates panoramic images. It is capable of aligning with the 3D model, the 'digital base' for the twin landscapes, by extracting the GPS information associated with the panoramas. By incorporating approaches such as computer vision and

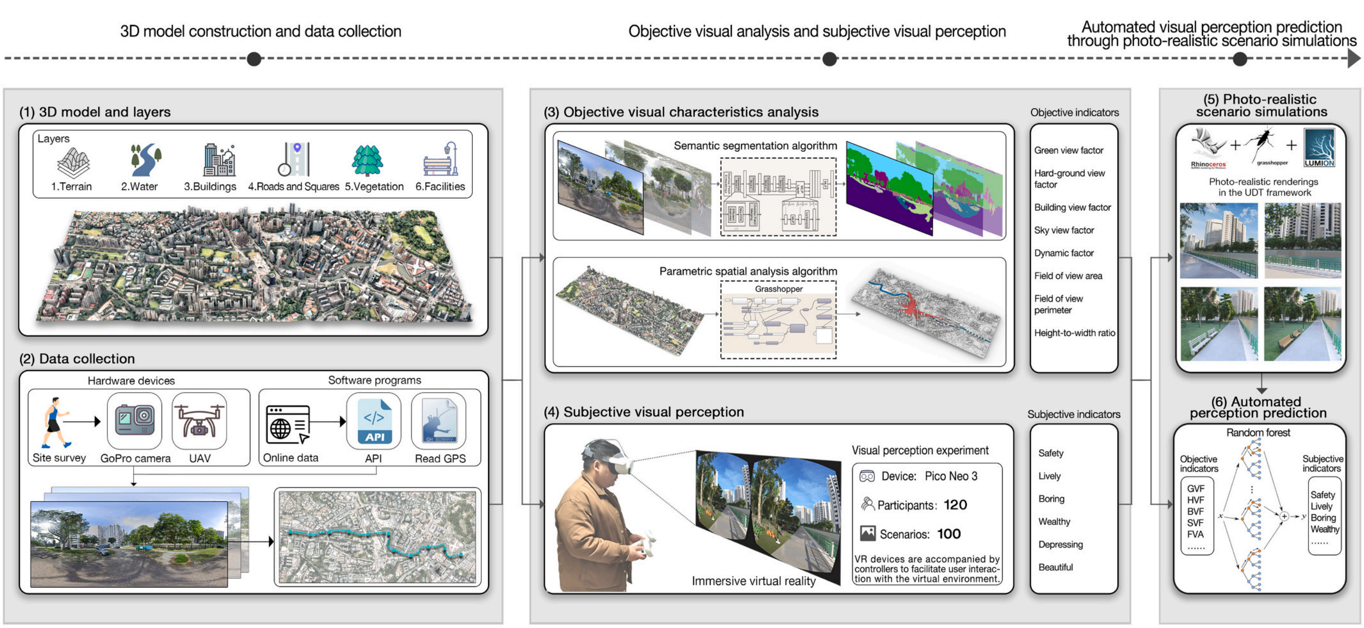


Fig. 3. Workflow of this study.

machine learning algorithms within the UDT platform, a multidimensional analysis of the visual characteristics of the urban greenway can be conducted. Additionally, with the assistance of the Lumion photo-realistic rendering engine, further high-quality renderings of textures and materials for 3D models were achieved on the Rhino Grasshopper platform. Fig. 4 illustrates the key features of the established UDT environment in Singapore. The participants showed a positive attitude towards the feasibility of substituting photos with photo-realistic renderings based on the Likert 7 scale. The evaluation yielded an average score of 5.533 with a standard deviation of 0.9. The reliability quality was high with a Cronbach's alpha of 0.727, greater than 0.7. Therefore, the accuracy of the digital counterpart is ensured by utilizing textures derived from panoramas taken in the real world.

5.2. Results of the preliminary perception study

This study demonstrated the utilization of the urban digital twin for the built environment's objective characteristics analysis and subjective visual perceptions (Fig. 5). The analysis of objective features evaluation primarily involved the application of the computer vision algorithm for pixel-level semantic segmentation of panoramic imagery, quantifying the proportions of landscape elements. Additionally, the digital twin model was employed to achieve high-precision spatial feature quantification analysis, such as the visual field area. Furthermore, immersive subjective visual perception experiments were conducted to collect participants' visual perceptions of different locations. Then, a comprehensive comparative analysis and spatial distribution characteristic examination were performed on the obtained results. Finally, the study explored the correlation between objective analysis results and subjective visual perception outcomes. The RF algorithm was employed to analyze the environmental features that influence individuals' subjective visual perceptions.

5.2.1. Objective evaluation results

This study conducted a statistical analysis and overall comparison of 8 objective indicators in the study area (Table 1). Regarding the results of imagery semantic segmentation, the DeepLabV3 model performs exceptionally well on the Cityscapes dataset, achieving a mIoU of 81.3 % (Chen et al., 2017). Even without fine-tuning, high precision can be ensured by conducting post-merge processing of our major category indicators.

In the research area, the average value of the HVF in the research area is higher than the average values of other indicators. In contrast, the average value of the DF is lower than the average value of other factors. The standard deviation of the GVF is higher than others. Additionally, the maximum value of the GVF is higher than the maximum value of other semantic segmentation indicators. The minimum value of the HVF is high, while the minimum values of the SVF, BVF, and DF are almost zero. Regarding the results of parametric spatial feature analysis, the average value of the FVA index is 0.056 km^2, and a standard deviation of 0.034. The average value of the FVP is 2.688 km, and a standard deviation of 1.462. Furthermore, the average value of the HWR of building vertical facades and flat surfaces is 2.242, and the standard deviation of the index is 1.102.

We also conducted a comprehensive spatial analysis of the evaluation indicators in the research area (Fig. 6). Notable factors such as HVF, BVF, and GVF demonstrate significant spatial variations. Additionally, SVF, DF, and FVA are explored, revealing distinct patterns across different survey points. The findings provide valuable insights into the spatial dynamics of visual and environmental characteristics in the studied region.

5.2.2. Human visual perception results

After the completion of the experiment by all participants, this study employed Pearson correlation analysis to examine the correlation between the evaluation results of the two groups of experimenters. The analysis revealed a correlation coefficient of 0.923, indicating a highly positive correlation between the results of the two groups. Therefore, the consistency of the experimental results can be considered substantial. Ultimately, the final scores for each scenario were determined as the average ratings given by all participants for each scenario.

We conducted a statistical analysis and overall comparison of the six subjective perception indicators within the research area (Table 2). It was found that the average perception score for the safety index was higher than the other indicators, while the corresponding value of depression was the lowest. The standard deviation of the boring index was high, in contrast, the value of the lively indicator was slightly lower than others. The minimum value for the boring and depressing indicators was low, while the safety index was the highest. Additionally, the maximum value for the safety perception indicator was higher than that of other indicators, and it can also be found that the maximum value of the depressing index was the lowest.

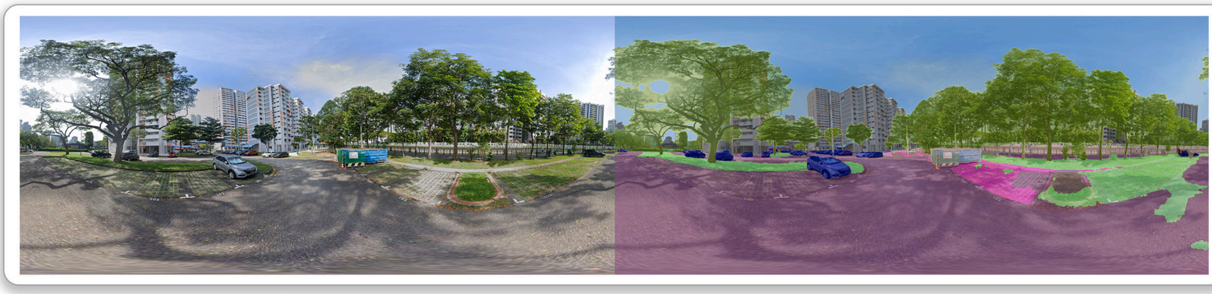
Fig. 7 shows the spatial variation characteristics of subjective visual perception results in the linear research area. The findings indicate substantial fluctuations in perception scores across the area. For example, safety perception exhibits an overall decreasing trend, while liveliness perception varies spatially. Depression and boredom indicators show significant fluctuations with specific extreme values. Overall, the study provides insights into the dynamic nature of visual perception in the built environment.

5.2.3. Prediction model accuracy and proportion of importance

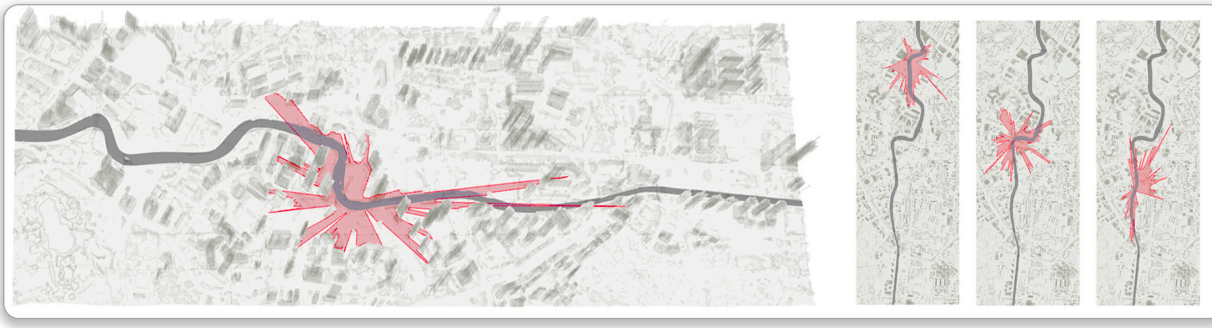
This research employs Pearson correlation analysis to investigate the strength of relationships between subjective visual perception indicators (e.g. beautiful, safety, and liveliness), and objective analytical indicators, such as green visibility factor, building visibility factor, and field of view area (Fig. 8). We found that there is a correlation between the GVF and all the subjective perception indicators. Specifically, there

(A). Photograph of an example location (B). Virtual counterpart in UDT

Fig. 4. Key features of the virtual counterpart of the study area in the UDT platform. In this research, we demonstrate that renders from 3D models give the same visual impression in surveys, confirming the value of photo-realistic renderings of scenarios that are yet to be built for participatory planning.



A. Imagery semantic segmentation



B. Parametric spatial feature analysis (field of view area)

Fig. 5. Semantic segmentation of the collected panoramic imagery and parametric spatial analysis outcomes.

Table 1
Overall comparison of objective indicators.

	HVF	BVF	GVF	SVF	DF	FVA	FVP	HWR
Average	0.301	0.081	0.234	0.224	0.012	0.056	2.688	2.242
Standard Deviation	0.060	0.052	0.120	0.085	0.011	0.034	1.462	1.102
Minimum	0.136	0.000	0.016	0.006	0.000	0.000	0.085	0.280
Maximum	0.442	0.267	0.589	0.422	0.051	0.129	6.607	5.425

is a strong positive correlation between the GVF and the beautiful perception index, a moderate positive correlation with the safety index, but negative correlations with boredom, depression, liveliness, and wealth perception. Additionally, the HVF indicator exhibits a moderate positive correlation with the boring perception index and a weak negative correlation with the beautiful indicator. The BVF index is moderately positively correlated with the wealthy and lively indexes, weakly positively correlated with the depression indicator, and weakly negatively correlated with the beautiful index. Moreover, the analysis of the correlation between the three spatial feature analysis factors (FVA, FVP, and HWR) and subjective perception indicators reveals further insights. FVA demonstrates a moderate negative correlation with depression perception, indicating that places with larger visual field areas tend to have lower depression perception. The FVP index exhibits a weak negative correlation with boredom perception and a weak positive correlation with beauty perception. The HWR indicator has weak negative correlations with safety and boredom perception, while it is not correlated with other indicators.

In this study, the results of the objective analysis of the built environment's characteristics were used as independent variables in RF models, with the subjective visual perception results serving as dependent variables for model training. It was observed that when employing the optimal model parameters as presented in Table 3, these models demonstrated the highest accuracy and generalisation ability. For instance, in the wealthy perception model, selecting parameters n_estimators as 100, max_depth as 10, and min_samples_split as 6 yielded the best model performance. In this configuration, the model achieved an R-squared value of 0.509, while the MSE and MAE values were relatively small.

After training the RF models, we computed the importance of all variables and showed the bee swarm plots to reflect the distribution and spread of SHAP values for each feature (Fig. 9). The SHAP plots provide deeper insights into the significance and influence of each feature on human perceptions. Each point on the bee swarm plot represents a sample, with its horizontal position indicating the SHAP value and its color reflecting the magnitude of the impact (red for larger impacts, blue for smaller impacts). This visualization allows us to observe both the importance and the direction of each feature's impact on the predictions.

The variable importance assessment for the lively perception model revealed that the BVF was significantly more important than other indicators. The DF, FVA, and GVF indicators followed as the second, third, and fourth most important variables, respectively, while the importance of other indicators was relatively low. For the beautiful and safety perception models, the GVF variable demonstrated substantial importance. In the wealthy perception model, the BVF was significantly more importance than other indicators, indicating its crucial role in assessing wealthy perception. Depression and boredom are two negative perception models, with the variable importance analysis of the depression perception model showing that FVA and GVF held much higher importance than other indicators. Furthermore, the variable importance ranking of the boredom perception model revealed that HVF and SVF exhibited relatively higher importance.

Building on the above statistical foundation, our study provided a nuanced understanding of which environmental features most significantly influence residents' perceptions. The RF model's variable importance assessment revealed that certain indicators, such as GVF for beauty and BVF for wealth perception, hold substantial predictive power. This suggests that green spaces and visible building structures

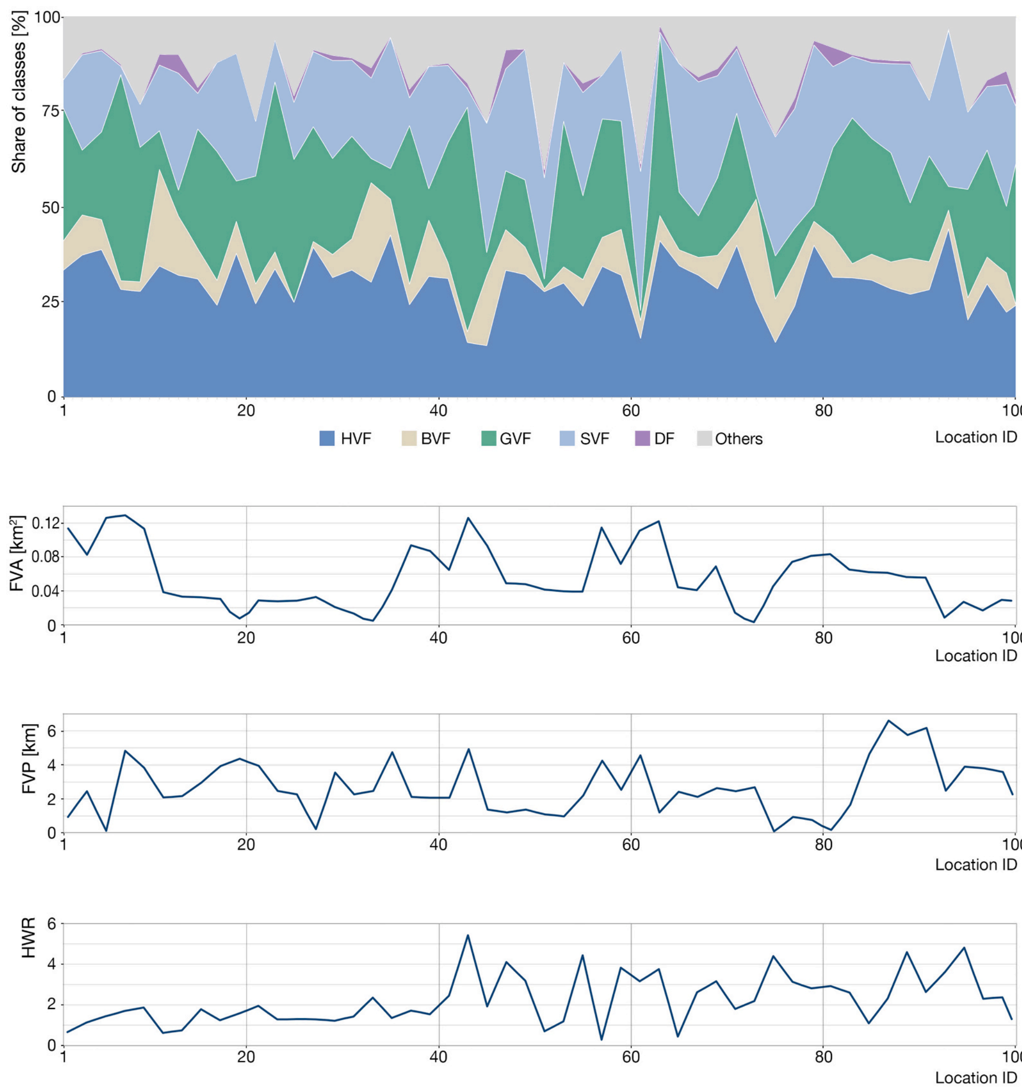


Fig. 6. Spatial variation characteristics of objective indicators.

Table 2
Overall comparison of subjective indicators.

	Wealthy	Safety	Lively	Depressing	Boring	Beautiful
Average	4.756	4.805	4.208	3.547	4.043	4.129
Standard Deviation	0.463	0.465	0.362	0.454	0.756	0.382
Minimum	3.583	3.750	3.583	2.667	2.667	3.417
Maximum	5.750	6.083	5.250	4.833	5.250	5.083

play crucial roles in shaping how people perceive their urban surroundings in terms of aesthetics and socioeconomic status. The underlying reasons for these relationships can be attributed to a variety of factors, including cultural values, economic conditions, and political frameworks. For example, the significant role of green visibility in influencing perceptions of beauty and safety may reflect a cultural appreciation for nature and green spaces as symbols of well-being and harmony within densely urbanized contexts. Economically, areas with higher visibility of buildings might be perceived as more affluent, impacting perceptions of wealth and liveliness. Politically, urban planning policies promoting green infrastructure and sustainable development could further enhance these perceptions, reflecting a government's commitment to improving urban livability and environmental sustainability. This deeper analysis of the statistical relationships and the RF model's findings could enrich the understanding of how urban environmental characteristics are intertwined with cultural, economic, and political factors to shape human perceptions (Li, Wang, et al., 2022). It highlights the potential of integrating UDT and computer vision techniques for more informed urban planning and policy-making that considers the complex determinants of public perception.

5.3. Automated visual perception prediction through photo-realistic scenario simulations

One of the primary objectives of this paper is to establish a ‘true digital twin’ within the context of sustainable urban development by demonstrating the feasibility of substituting photos with photo-realistic renderings, allowing predicting and testing different scenarios before they are built. This approach not only enhances the visual fidelity of the Urban Digital Twin (UDT) but also provides a robust platform for pre-emptive urban planning and design evaluation (Li, Xue, et al., 2022). The primary utilization of the UDT involves simulating diverse urban scenarios, such as altering vegetation density or introducing new architectural elements to the skyline. Through the incorporation of photo-realistic renderings, the UDT becomes a dynamic tool for envisioning and virtually implementing changes within the urban environment. Additionally, a distinctive feature of this work is the automated prediction of how individuals would perceive the altered urban

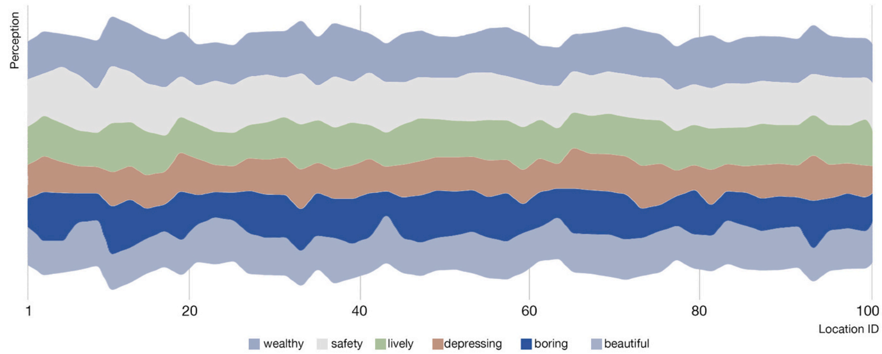


Fig. 7. Spatial variation of subjective perception results, visualized as a streamgraph.

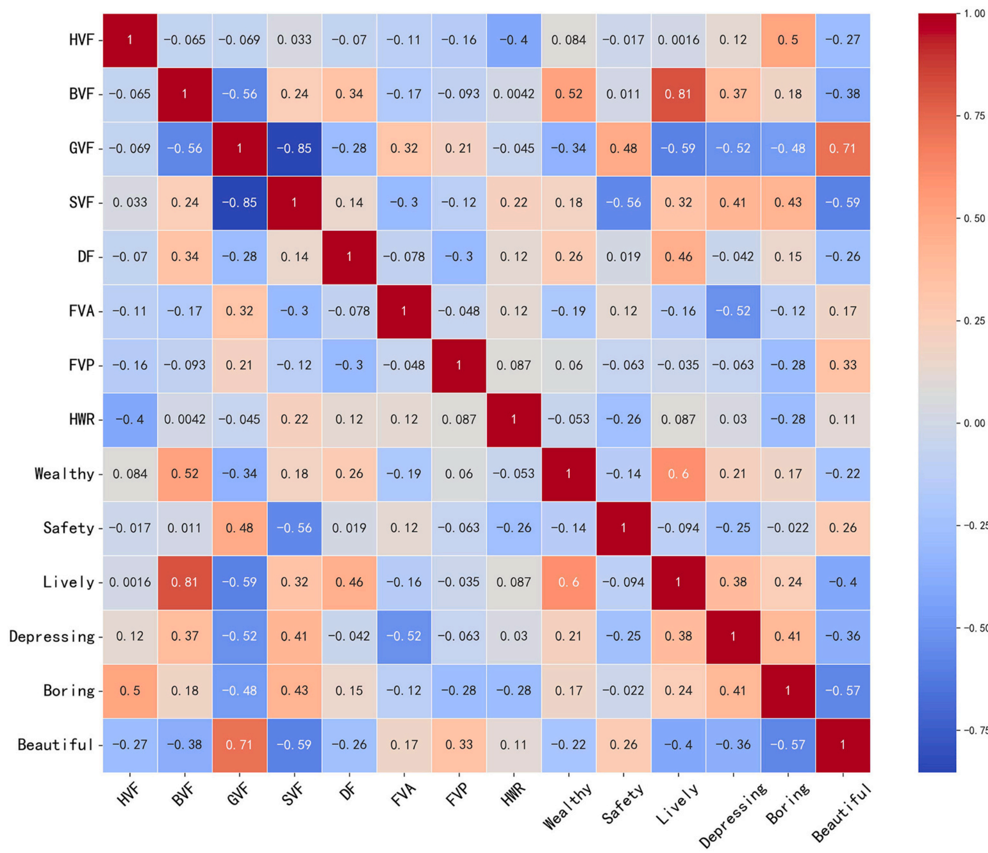


Fig. 8. Pearson correlation of data collected.

Table 3
Performance of the random forest model.

		wealthy	safety	lively	depression	boring	beautiful
Optimal parameters	max_depth	10	5	5	5	5	5
	min_samples_split	6	6	2	2	4	2
	n_estimators	100	200	100	200	50	100
R2		0.509	0.487	0.438	0.541	0.489	0.619
MSE		0.08	0.092	0.044	0.136	0.318	0.134
MAE		0.234	0.222	0.165	0.262	0.434	0.281

scenarios. Instead of relying on conventional survey methods, the RF prediction model is applied within the UDT platform. This approach dynamically assesses and anticipates the visual perceptions people would have in response to the simulated scenarios, offering a more efficient and automated approach to understanding the human experience of the urban environment.

Fig. 10 demonstrates the visual perception prediction through photo-realistic scenario simulations in our study area. Specific operations, such as ‘add, delete, modify’ can be executed on landscape elements like trees and buildings in UDT. We conducted an objective quantification analysis

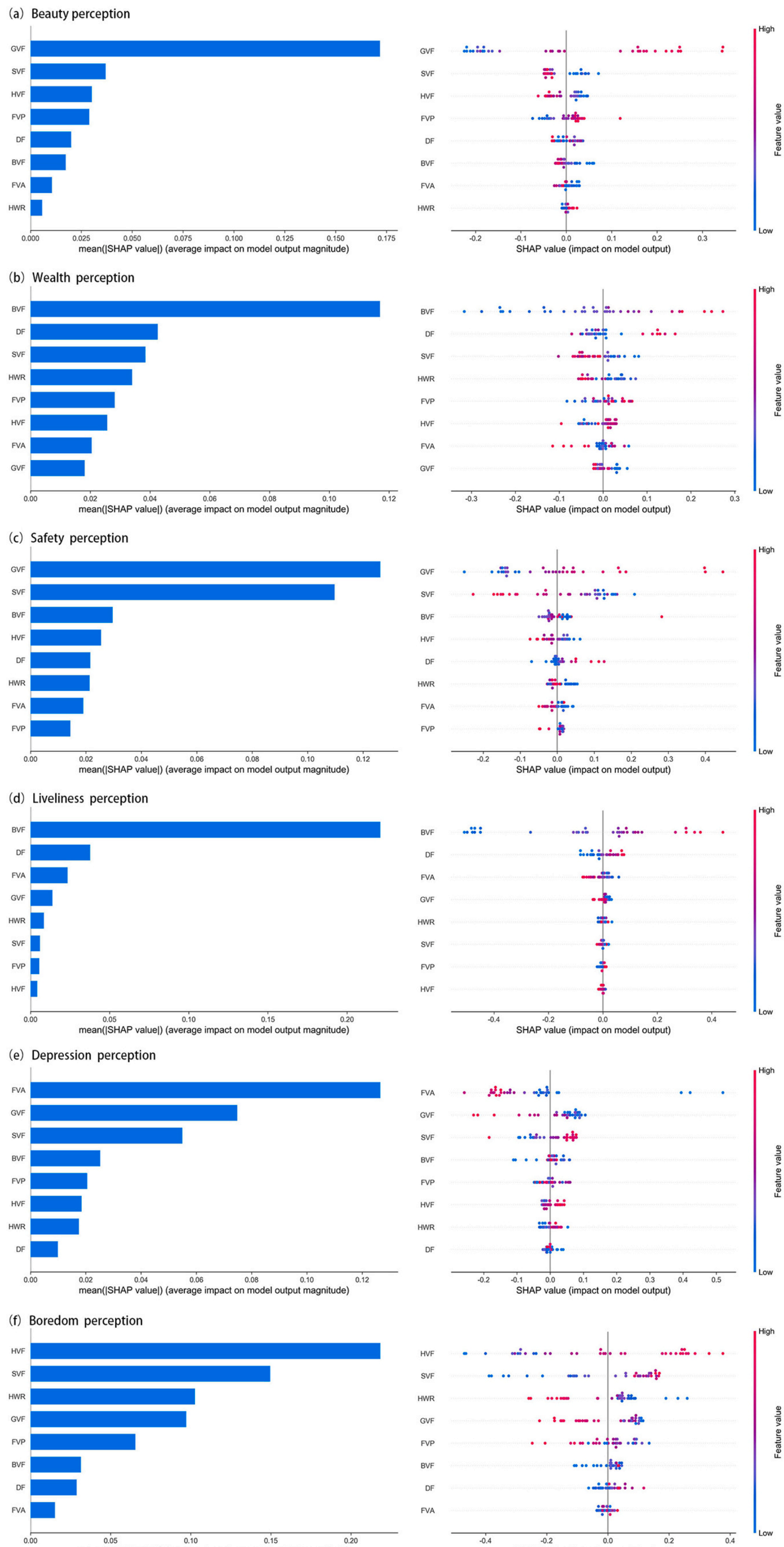


Fig. 9. Order of significance and bee swarm plots of random forest models.

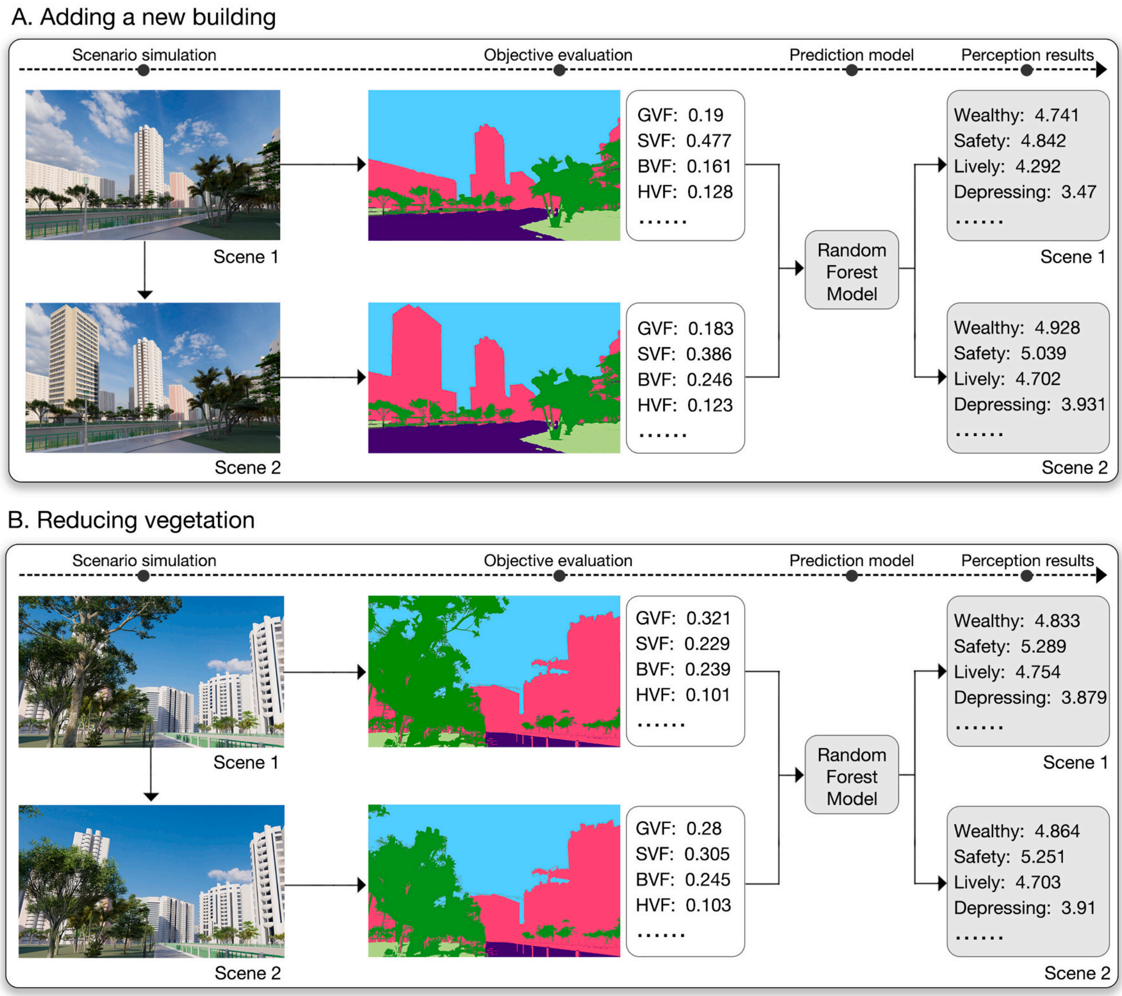


Fig. 10. Demonstrations of the UDT in testing different future scenarios for predicting visual perception results.

of visual features in the UDT, comparing scenarios before and after the addition of buildings and changes in vegetation. Several visual characteristics, such as the green view factor, were quantified to assess the visual impact of these scenarios. Subsequently, we employed the RF prediction model to automatically predict subjective visual perception outcomes (e.g. Safety and Lively).

After predicting the subjective visual perception results, the accuracy of the prediction results can be verified through additional visual perception experiments. Specifically, when obtaining the perceptual rating results through the experiments, data comparison (e.g. correlation analysis) can be conducted between the experimental results and the predicted outcomes. If the numerical values and distribution characteristics of the data are very similar, or there is a high positive correlation between the two, it can be considered that the predicted data is reliable. Moreover, by utilizing virtual reality within the UDT platform, people can personally compare urban landscapes under different scenarios, gaining a deeper understanding of each scenario's features. People can also provide feedback and suggestions in the interactive environment.

This dual approach not only establishes the credibility of the 'true digital twin' but also paves the way for advancements in predictive analytics within sustainable urban development, streamlining decision-making processes and enhancing the overall effectiveness of urban planning strategies.

6. Discussion

6.1. Digital twins for human visual perceptions and sustainable urban development

Cities, as complex mega-systems, have become a shared objective among governments worldwide to be developed into 'smarter' and more sustainable urban environments, and digital twin technology is regarded as one of the pathways to achieve this goal (Caldarelli et al., 2023; Han & Kim, 2024; Lehtola et al., 2022). Building sustainable cities and societies is a complex and enduring process that aims to improve the quality of life for their residents (Heidari et al., 2022; Xia et al., 2022), and enhancing visual perception experience is also an important part. Traditional research on visual perception often relies on different types of on-site investigations or imagery-based quantitative evaluations. The former is expensive and time-consuming, and the latter cannot evaluate the 3D characteristics (e.g. field of view area and height-to-width ratio) of the physical environment. Moreover, these methods cannot establish connections and interactions between the physical settings, virtual counterpart and human perceptions, which are limited in their ability to perform scenario simulation and planning schemes evaluation (Kruse et al., 2021; Pirker et al., 2022; Xu et al., 2022), particularly in the tasks of visual perception predictions for the environment studies. Thanks to the development of UDT-related technologies, we can digitally replicate and virtually express objective entities, reflecting the mirror mapping relationship between the virtual and the real world (Caldarelli et al., 2023; Ferré-Bigorra et al., 2022). In this context, we utilized a 3D model as the foundation of the UDT, thereby establishing a 'digital base' for the twin landscapes. In the UDT research of the built environment, the 3D model serves as both the basis for achieving all research objectives and a data type. The 3D model in the UDT platform can be periodically updated using approaches such as the UAV oblique photography modelling based on research aims to better match the current status of the landscape environment, which can achieve high-precision reproduction of real-world scenarios, and it can simulate future scenarios before they are implemented in reality, which traditionally used approaches cannot match. Furthermore, the UDT constructed in this study incorporates panoramic images as its data source, enabling different types of visual perception analysis, and it becomes possible to employ a variety of machine learning algorithms for multidimensional evaluations. This reflects the core value of the digital twin, which is to create an algorithm-driven virtual counterpart that provides quantitative analysis, and visual simulation, and supports scientific decision-making.

The relationship between information and its perceived value is crucial for understanding human responses to urban environments (Naghibi, 2024; Vuong, 2023). The human visual perception UDT, we developed in this study, enables a virtual environment where citizens can interact with and contribute to the planning and assessment of urban scenarios. Through immersive virtual reality and scenario simulations, individuals could gain firsthand experience of proposed changes, fostering informed decision-making and increasing the likelihood of public acceptance. We acknowledge that the engagement of citizens in the planning process is essential for the success and sustainability of urban development initiatives. The integration of human perceptions facilitates a dynamic feedback loop, where the public can provide insights, preferences, and concerns in time. This iterative process contributes to the creation of urban spaces that better align with the needs and expectations of the city. By incorporating the inhabitants' perspective, digital twins can be more intuitive and user-friendly, making it easier for people to explore and interact with the city.

6.2. Limitations, challenges and future directions

This study faces some challenges and limitations.

- Firstly, panoramic imagery, serving as one of the data sources for the digital twin platform in this study, has a relatively long data update cycle. Currently, this study relies mainly on panoramic imagery from the Mapillary platform and panoramas obtained by researchers for updates (Hou & Biljecki, 2022).
- Secondly, the current 3D model in the UDT platform is primarily constructed manually referencing data such as drone oblique photography and street view images, which is a laborious process. In the later stages, this study will employ fully automated drone control programs to update the UDT, thereby enhancing research efficiency regularly.
- Thirdly, while employing immersive virtual reality to perceive urban environments provides valuable insights into public perception, it may differ from on-site experiences. These differences can stem from various factors, including the absence of multi-sensory feedback, altered spatial awareness, and emotional responses that are unique to physical presence (Birenboim et al., 2019; Luo, Zhao, et al., 2022c). Recognizing this limitation, our study underscores the need for further exploration into quantifying and understanding the nuances between immersive virtual experiences and direct, physical interactions with urban spaces.

Efficiently integrating public perceptions and evaluations of landscapes into urban design and renewal is a current interdisciplinary focus (Adrienne & Nora, 2024; Jiang, Han, et al., 2022; Yıldız et al., 2020). As mentioned earlier, we witness there is potential in employing VR for comprehensive visual experience assessments. This involves simulating various visual scenarios to predict how proposed changes, such as architectural modifications, would influence the aesthetic aspects of the urban environment. In the field of urban landscape planning and design, stakeholders, including surrounding residents, urban management authorities, and researchers, can evaluate and perceive multiple planning schemes using the immersive VR within the UDT platform. The UDT platform, featuring interactive components, presents new opportunities for active public engagement in urban planning by urban management departments and has become a vital tool for public '*E*-participation' in the built environment (Orsetti et al., 2022). While the application of UDTs undeniably facilitates public understanding and participation in urban planning and design processes, thereby enhancing public engagement and decision-making transparency, it is imperative to acknowledge the nuanced implications of such digital integration. The deployment of UDTs and similar digital technologies raises legitimate concerns regarding privacy, data security, and the potential for increased surveillance (Cui et al., 2022). These concerns reflect a broader paradox inherent in the digital age, where the benefits of increased accessibility and interactive participation are weighed against the risks associated with privacy and data protection.

Moreover, the rapid advancements in technologies such as deep learning, photo-realistic rendering algorithms, Internet of Things (IoT) and edge computing, (near)-real-time UDTs are poised to play a crucial role in urban environmental research (Ghazal et al., 2023; Kikuchi et al., 2023; Puli et al., 2022; Xu, Shao, et al., 2023), which can help to simulate urban environments and human behaviours, offering novel approaches to urban planning, design, and environmental evaluation. Digital twins based on IoT and edge computing can leverage vast amounts of environmental data for (near)-real-time analysis (Dai & Zhang, 2022; Durana et al., 2022). For instance, monitoring indicators such as traffic flow, air quality, and crowd density within a city allows for a more accurate understanding of changes. Additionally, UDTs can simulate human behaviours, providing insights into human perceptions and reactions to the urban environment.

Overall, the value of UDTs in urban environmental research lies in their ability to conduct multidimensional scenario simulations and predictive analyses of landscapes, furnishing city managers with more comprehensive and accurate decision-making foundations. As technology continues to advance, the assessment and perception of built environment characteristics regarding sustainable and people-centric urban environments is evolving. This tech-driven approach fosters inclusivity, transparency, and understanding, aligning with the principles of sustainable urban development.

7. Conclusion

This research proposes a pathway for built environment characteristics evaluation and perception that combines objective feature analysis with subjective perceptions based on an urban digital twin we designed and developed. The idea of an urban digital twin is to replicate the real-world and test scenarios before they are built, driving their implementation in the real-world, which we have accomplished by integrating an automated means to human perception simulation in a digital model with high fidelity, informing design and policies.

By applying the Mindsponge Theory, we were able to better interpret the subjective visual perceptions gathered from participants and relate them to the objective environmental features quantified by the digital twin. Objective feature analysis involves using an advanced DeepLabV3 deep learning algorithm for pixel-level semantic segmentation of continuous panoramic images, which could quantify different features (e.g. green visibility and sky exposure) in the urban environment. The 3D model is employed for high-precision spatial feature quantification analysis, including 3 novel indicators in the UDT field: the field of view area, the field of view perimeter, and the height-to-width ratio of building vertical facades and flat surfaces. Additionally, subjective perception involves obtaining evaluative information from participants regarding different scenarios through immersive VR devices. A broader contribution of the work is that, with its two-way direction, it contributes towards achieving a 'true' digital twin for urban planning, and it sets the foundation for automating the optimisation of urban scenarios and enhancing urban sustainability. Further contributions include demonstrating that photo-realistic renderings based on the UDT can replace the traditionally used photos for this use case, confirming their value in participatory planning.

As the implementation of the novel approach, the paper presents a case study of an urban greenway in the central area of Singapore. Following the construction of a digital twin for this area, we conduct a comprehensive comparison and spatial distribution analysis of the objective environmental features and subjective visual perception results of the urban greenway. Subsequently, the study utilizes the RF algorithm to analyze the importance of objective environmental variables influencing subjective visual perceptions, contributing significantly to understanding the impact mechanisms of objective environmental features on subjective visual perceptions. The pivotal contribution lies in leveraging the UDT not only for simulating diverse urban scenarios but, significantly, for automating the prediction of subjective visual perceptions associated with photo-realistic renderings, and we proved their veracity based on a perception experiment. This underscores the core value of the digital twin constructed in this study, namely, creating an algorithm-driven methodology capable of providing quantitative analysis, public engagement, and assisting in scientific decision-making for urban environments.

With the accelerated pace of urbanization, increased emphasis is being placed on the quantitative analysis of the built environment and research related to public perception. Digital twins, as an innovative approach, bear substantial potential in the analysis of urban environments, serving as a pivotal method to elevate the interaction between humans and their surroundings. The integration of public perceptions into UDTs is crucial for establishing a bidirectional connection between the 'digital platform—public perception—physical environment.' This integration encapsulates a specific manifestation of the sustainable urban development concept centred around the public.

CRediT authorship contribution statement

Junjie Luo: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization, Funding acquisition, Project administration. **Pengyuan Liu:** Writing – review & editing, Methodology, Validation. **Wenhui Xu:** Investigation. **Tianhong Zhao:** Writing – review & editing, Methodology, Software. **Filip Biljecki:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization, Resources.

Declaration of competing interest

We declare that there are no competing interests and we do not have any conflict of interest we are aware of.

Acknowledgements

We gratefully acknowledge the participants of the experiment. We thank the members of the NUS Urban Analytics Lab for the discussions, in particular, Binyu Lei, Winston Yap, and Leon Gaw. The Institutional Review Board of the National University of Singapore has reviewed and approved the ethical aspects of this research (reference code NUS-IRB-2022-191). This research is part of the projects (i) Visual Perception Analysis of Urban Riverscapes Using Computer Vision, which is supported by the Zhejiang A&F University Scientific Research Development Fund (Start Up Project, grant no. 2023LFR076); (ii) Large-scale 3D Geospatial Data for Urban Analytics, which is supported by the National University of Singapore under the Start Up Grant R-295-000-171-133; and (iii) Multi-scale Digital Twins for the Urban Environment: From Heartbeats to Cities, which is supported by the Singapore Ministry of Education Academic Research Fund Tier 1.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT in order to check the grammar and improve readability. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Data availability

The authors do not have permission to share data.

References

Adade, D., & de Vries, W. T. (2023). Digital twin for active stakeholder participation in land-use planning. *Land, 12*, 538.
Adrienne, G. R., & Nora, F. (2024). Key factors to enhance efficacy of 3d digital environments for transformative landscape and urban planning. *Landscape and Urban Planning, 244*, Article 104978.
AlBalkhy, W., Karmaoui, D., Ducoulombier, L., Lafhaj, Z., & Linner, T. (2024). Digital twins in the built environment: Definition, applications, and challenges. *Automation in Construction, 162*, Article 105368.
Al-Sehrawy, R., Kumar, B., & Watson, R. (2023). The pluralism of digital twins for urban management: Bridging theory and practice. *Journal of Urban Management, 12*, 16–32.
Argota Sánchez-Vaquerizo, J., Hausladen, C. I., Mahajan, S., Matter, M., Siebenmann, M., van Eggermond, M. A., & Helbing, D. (2024). A virtual reality experiment to study pedestrian perception of future street scenarios. *Scientific Reports, 14*, 4571.
Attaran, M., & Celik, B. G. (2023). Digital twin: Benefits, use cases, challenges, and opportunities. *Decision Analytics Journal, 100165*.
Barth, L., Schweiger, L., Benedech, R., & Ehrat, M. (2023). From data to value in smart waste management: Optimizing solid waste collection with a digital twin-based decision support system. *Decision Analytics Journal, 100347*.
Bauer, P., Stevens, B., & Hazeleger, W. (2021). A digital twin of earth for the green transition. *Nature Climate Change, 11*, 80–83.
Bevilacqua, M. G., Russo, M., Giordano, A., & Spallone, R. (2022). 3d reconstruction, digital twinning, and virtual reality: Architectural heritage applications, in: 2022 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW). *IEEE. pp., 92–96*.
Biau, G. (2012). Analysis of a random forests model. *The Journal of Machine Learning Research, 13*, 1063–1095.
Biljecki, F., & Chow, Y. S. (2022). Global building morphology indicators. *Computers, Environment and Urban Systems, 95*, Article 101809. https://doi.org/10.1016/j.compenvurbsys.2022.101809
Biljecki, F., Zhao, T., Liang, X., & Hou, Y. (2023). Sensitivity of measuring the urban form and greenery using street-level imagery: A comparative study of approaches and visual perspectives. *International Journal of Applied Earth Observation and Geoinformation, 122*, Article 103385.
Birenboim, A., Dijst, M., Ettema, D., de Kruijf, J., de Leeuw, G., & Dogterom, N. (2019). The utilization of immersive virtual environments for the investigation of environmental preferences. *Landscape and Urban Planning, 189*, 129–138.
Boccardo, P., La Riccia, L., & Yadav, Y. (2024). Urban echoes: Exploring the dynamic realities of cities through digital twins. *Land, 13*, 635.
Bolívar, M. P. R., Munoz, L. A., & Munoz, C. A. (2023). *Identifying patterns in smart initiatives' planning in smart cities. an empirical analysis in spanish smart cities. Technological Forecasting and Social Change, 196*, 122781.
Botín-Sanabria, D. M., Mihaita, A. S., Peimbert-García, R. E., Ramírez-Moreno, M. A., Ramírez-Mendoza, R. A., & Lozoya-Santos, J.d. J. (2022). Digital twin technology challenges and applications: A comprehensive review. *Remote Sensing, 14*, 1335.
Caldarelli, G., Arcaute, E., Barthelemy, M., Batty, M., Gershenson, C., Helbing, D., Mancuso, S., Moreno, Y., Ramasco, J., Rozenblat, C., et al. (2023). The role of complexity for digital twins of cities. *Nature Computational Science*, 1–8.
Chen, C., Wang, H., Wang, D., & Wang, D. (2024). Towards the digital twin of urban forest: 3d modeling and parameterization of large-scale urban trees from close-range laser scanning. *International Journal of Applied Earth Observation and Geoinformation, 127*, Article 103695.
Chen, L.C., Papandreou, G., Schroff, F., Adam, H., 2017. Rethinking atrous convolution for semantic image segmentation. arXiv preprint arXiv:1706.05587.
Chen, S., & Biljecki, F. (2023). Automatic assessment of public open spaces using street view imagery. *Cities, 137*, Article 104329.
Cui, Q., Wei, R., Huang, R., Hu, X., & Wang, G. (2022). The effect of perceived risk on public participation intention in smart city development: Evidence from China. *Land, 11*, 1604.
Dai, Y., & Zhang, Y. (2022). Adaptive digital twin for vehicular edge computing and networks. *Journal of Communications and Information Networks, 7*, 48–59.
Dang, X., Liu, W., Hong, Q., Wang, Y., & Chen, X. (2023). Digital twin applications on cultural world heritage sites in China: A state-of-the-art overview. *Journal of Cultural Heritage, 64*, 228–243.

Dembski, F., Wössner, U., Letzgus, M., Ruddat, M., & Yamu, C. (2020). Urban digital twins for smart cities and citizens: The case study of Herrenberg, Germany. *Sustainability, 12*, 2307.

Dembski, F., Yamu, C., & Wössner, U. (2019). Digital twin, virtual reality and space syntax: Civic engagement and decision support for smart, sustainable cities. *Proceedings of the 12th international space syntax symposium.* pp. 316–1.

Deng, T., Zhang, K., & Shen, Z. J. M. (2021). A systematic review of a digital twin city: A new pattern of urban governance toward smart cities. *Journal of Management Science and Engineering, 6*, 125–134.

Dubey, A., Naik, N., Parikh, D., Raskar, R., & Hidalgo, C. A. (2016). Deep learning the city: Quantifying urban perception at a global scale. *European conference on computer vision, Springer.*, 196–212.

Durana, P., Krastev, V., & Buckner, K. (2022). Digital twin modeling, multi-sensor fusion technology, and data mining algorithms in cloud and edge computing-based smart city environments. *Geopolitics, History, and International Relations, 14*, 91–106.

Elkefi, S., & Asan, O. (2022). Digital twins for managing health care systems: Rapid literature review. *Journal of Medical Internet Research, 24*, Article e37641.

Elnabawi, M. H., & Raveendran, R. (2024). Meta-pragmatic investigation of passive strategies from 'uhi–climatology'nexus perspective with digital twin as assessment mechanism. *Journal of Urban Management*.

Fan, C., Zhang, C., Yahja, A., & Mostafavi, A. (2021). Disaster city digital twin: A vision for integrating artificial and human intelligence for disaster management. *International Journal of Information Management, 56*, Article 102049.

Ferré-Bigorra, J., Casals, M., & Gangolells, M. (2022). The adoption of urban digital twins. *Cities, 131*, Article 103905.

Foroozesh, F., Monavari, S. M., Salmanmahiny, A., Robati, M., & Rahimi, R. (2022). Assessment of sustainable urban development based on a hybrid decision-making approach: Group fuzzy bwm, ahp, and topsis–gis. *Sustainable Cities and Society, 76*, Article 103402.

Francisco, A., Mohammadi, N., & Taylor, J. E. (2020). Smart city digital twin–enabled energy management: Toward real-time urban building energy benchmarking. *Journal of Management in Engineering, 36*, 04019045.

Funari, M. F., Hajjat, A. E., Masciotta, M. G., Oliveira, D. V., & Lourenço, P. B. (2021). A parametric scan-to-fem framework for the digital twin generation of historic masonry structures. *Sustainability, 13*, 11088.

Geremicca, F., & Bilec, M. M. (2024). Searching for new urban metabolism techniques: A review towards future development for a city-scale urban metabolism digital twin. *Sustainable Cities and Society, 105445*.

Ghazal, T. M., Hasan, M. K., Ahmad, M., Alzoubi, H. M., & Alshurideh, M. (2023). Machine learning approaches for sustainable cities using internet of things. In *The effect of information technology on business and marketing intelligence systems* (pp. 1969–1986). Springer.

Gholami, M., Torreggiani, D., Tassinari, P., & Barbaresi, A. (2022). Developing a 3d city digital twin: Enhancing walkability through a green pedestrian network (gpn) in the city of Imola, Italy. *Land, 11*, 1917.

Gong, F. Y., Zeng, Z. C., Zhang, F., Li, X., Ng, E., & Norford, L. K. (2018). Mapping sky, tree, and building view factors of street canyons in a high-density urban environment. *Building and Environment, 134*, 155–167.

Ham, Y., & Kim, J. (2020). Participatory sensing and digital twin city: Updating virtual city models for enhanced risk-informed decision-making. *Journal of Management in Engineering, 36*, 04020005.

Hamada, T., Chabi, F., Chakir, R., Lejri, D., Dugay, F., & Waeytens, J. (2024). Smart placement of depolluting panels in urban areas as regards of the airflow using adjoint framework and district digital twin. *Building and Environment, 248*, Article 111111.

Han, M. J. N., & Kim, M. J. (2024). A systematic review of smart city research from an urban context perspective. *Cities, 150*, Article 105027.

Haraguchi, M., Funahashi, T., & Biljecki, F. (2024). Assessing governance implications of city digital twin technology: A maturity model approach. *Technological Forecasting and Social Change, 204*, Article 123409. https://doi.org/10.1016/j.techfore.2024.123409

Heidari, A., Navimipour, N. J., & Unal, M. (2022). Applications of ml/dl in the management of smart cities and societies based on new trends in information technologies: A systematic literature review. *Sustainable Cities and Society, 104089*.

Hou, Y., & Biljecki, F. (2022). A comprehensive framework for evaluating the quality of street view imagery. *International Journal of Applied Earth Observation and Geoinformation, 115*, Article 103094. https://doi.org/10.1016/j.jag.2022.103094

Hou, Y., Quintana, M., Khomiakov, M., Yap, W., Ouyang, J., Ito, K., … Biljecki, F. (2024). Global streetscapes – A comprehensive dataset of 10 million street-level images across 688 cities for urban science and analytics. *ISPRS Journal of Photogrammetry and Remote Sensing, 215*, 216–238. https://doi.org/10.1016/j.isprsjprs.2024.06.023

Hu, J., Lv, Z., Yuan, D., He, B., Chen, W., Ye, X., Li, D., & Yang, G. (2023). A spatiotemporal intelligent framework and experimental platform for urban digital twins. *Virtual Reality & Intelligent Hardware, 5*, 213–231.

Iban, M. C. (2022). An explainable model for the mass appraisal of residences: The application of tree-based machine learning algorithms and interpretation of value determinants. *Habitat International, 128*, Article 102660.

Ivanov, S., Nikolskaya, K., Radchenko, G., Sokolinsky, L., & Zymbler, M. (2020). Digital twin of city: Concept overview, in: 2020 global smart industry conference (GloSIC). *IEEE. pp., 178–186*.

Jafari, M., Kavousi-Fard, A., Chen, T., & Karimi, M. (2023). *A review on digital twin technology in smart grid, transportation system and smart city: Challenges and future.* IEEE Access.

Jiang, F., Ma, L., Broyd, T., Chen, W., & Luo, H. (2022). Digital twin enabled sustainable urban road planning. *Sustainable Cities and Society, 78*, Article 103645.

Jiang, Y., Han, Y., Liu, M., & Ye, Y. (2022). Street vitality and built environment features: A data-informed approach from fourteen chinese cities. *Sustainable Cities and Society, 79*, Article 103724.

Kang, Y., Zhang, F., Gao, S., Lin, H., & Liu, Y. (2020). A review of urban physical environment sensing using street view imagery in public health studies. *Annals of GIS, 26*, 261–275.

Kenett, R. S., & Bortman, J. (2022). The digital twin in industry 4.0: A wide-angle perspective. *Quality and Reliability Engineering International, 38*, 1357–1366.

Kikuchi, N., Fukuda, T., & Yabuki, N. (2022). Future landscape visualization using a city digital twin: Integration of augmented reality and drones with implementation of 3d model-based occlusion handling. *Journal of Computational Design and Engineering, 9*, 837–856.

Kikuchi, T., Fukuda, T., & Yabuki, N. (2023). Development of a synthetic dataset generation method for deep learning of real urban landscapes using a 3d model of a non-existing realistic city. *Advanced Engineering Informatics, 58*, Article 102154.

Koenig, R., Miao, Y., Aichinger, A., Knecht, K., & Konieva, K. (2020). Integrating urban analysis, generative design, and evolutionary optimization for solving urban design problems. *Environment and Planning B: Urban Analytics and City Science, 47*, 997–1013.

Kruse, J., Kang, Y., Liu, Y. N., Zhang, F., & Gao, S. (2021). Places for play: Understanding human perception of playability in cities using street view images and deep learning. *Computers, Environment and Urban Systems, 90*, Article 101693.

Künz, A., Rosmann, S., Loria, E., & Pirker, J. (2022). The potential of augmented reality for digital twins: A literature review, in: 2022 IEEE conference on virtual reality and 3D user interfaces (VR). *IEEE. pp., 389–398*.

Kuru, K. (2023). *Metaomnicity: Towards immersive urban metaverse cyberspaces using smart city digital twins.* IEEE Access.

La Guardia, M., & Koeva, M. (2023). Towards digital twinning on the web: Heterogeneous 3d data fusion based on open-source structure. *Remote Sensing, 15*, 721.

Lee, A., Lee, K. W., Kim, K. H., & Shin, S. W. (2022). A geospatial platform to manage large-scale individual mobility for an urban digital twin platform. *Remote Sensing, 14*, 723.

Lehtola, V. V., Koeva, M., Elberink, S. O., Raposo, P., Virtanen, J. P., Vahdatikhaki, F., & Borsci, S. (2022). Digital twin of a city: Review of technology serving city needs. *International Journal of Applied Earth Observation and Geoinformation, 102915*.

Lei, B., Janssen, P., Stoter, J., & Biljecki, F. (2023). Challenges of urban digital twins: A systematic review and a delphi expert survey. *Automation in Construction, 147*, Article 104716.

Lei, B., Stouffs, R., & Biljecki, F. (2023). Assessing and benchmarking 3D city models. *International Journal of Geographical Information Science, 37*, 788–809. https://doi.org/10.1080/13658816.2022.2140808

Lei, B., Su, Y., & Biljecki, F. (2023). Humans as sensors in urban digital twins. In *International 3D GeoInfo conference* (pp. 693–706). Springer.

Li, M., Xue, F., Wu, Y., & Yeh, A. G. (2022). A room with a view: Automatic assessment of window views for high-rise high-density areas using city information models and deep transfer learning. *Landscape and Urban Planning, 226*, Article 104505.

Li, X., Feng, M., Ran, Y., Su, Y., Liu, F., Huang, C., Shen, H., Xiao, Q., Su, J., Yuan, S., et al. (2023). Big data in earth system science and progress towards a digital twin. *Nature Reviews Earth and Environment*, 1–14.

Li, X., Wang, X., Jiang, X., Han, J., Wang, Z., Wu, D., Lin, Q., Li, L., Zhang, S., & Dong, Y. (2022). Prediction of riverside greenway landscape aesthetic quality of urban canalized rivers using environmental modeling. *Journal of Cleaner Production, 367*, Article 133066.

Liang, X., Zhao, T., & Biljecki, F. (2023). Revealing spatio-temporal evolution of urban visual environments with street view imagery. *Landscape and Urban Planning, 237*, Article 104802.

Liu, M., Fang, S., Dong, H., & Xu, C. (2021). Review of digital twin about concepts, technologies, and industrial applications. *Journal of Manufacturing Systems, 58*, 346–361.

Liu, P., Zhao, T., Luo, J., Lei, B., Frei, M., Miller, C., & Biljecki, F. (2023). Towards human-centric digital twins: Leveraging computer vision and graph models to predict outdoor comfort. *Sustainable Cities and Society, 93*, Article 104480.

Lu, W., Chen, J., Fu, Y., Pan, Y., & Ghansah, F. A. (2023). Digital twin-enabled human-robot collaborative teaming towards sustainable and healthy built environments. *Journal of Cleaner Production, 412*, Article 137412.

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., … Lee, S. I. (2020). From local explanations to global understanding with explainable ai for trees. *Nature Machine Intelligence, 2*, 56–67.

Luo, J., Liu, P., & Cao, L. (2022). Coupling a physical replica with a digital twin: A comparison of participatory decision-making methods in an urban park environment. *ISPRS International Journal of Geo-Information, 11*, 452.

Luo, J., Zhao, T., Cao, L., & Biljecki, F. (2022b). Semantic riverscapes: Perception and evaluation of linear landscapes from oblique imagery using computer vision. *Landscape and Urban Planning, 228*, Article 104569.

Luo, J., Zhao, T., Cao, L., & Biljecki, F. (2022c). Water view imagery: Perception and evaluation of urban waterscapes worldwide. *Ecological Indicators, 145*, Article 109615.

Luo, S., Shi, J., Lu, T., & Furuya, K. (2022). Sit down and rest: Use of virtual reality to evaluate preferences and mental restoration in urban park pavilions. *Landscape and Urban Planning, 220*, Article 104336.

Lv, Z., & Fridenfalk, M. (2023). Digital twins for building industrial metaverse. *Journal of Advanced Research*.

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*, Article 103086.

Mazzoli, C., Papadaki, D., & Ferrante, A. (2023). A parametric approach for optimizing design solutions in urban regeneration and reshaping: An application to a district block in bologna, Italy. *Buildings, 13*, 3123.

Moyano, J., León, J., Nieto-Julián, J. E., & Bruno, S. (2021). Semantic interpretation of architectural and archaeological geometries: Point cloud segmentation for hbim parameterisation. *Automation in Construction, 130*, Article 103856.

Muhammad, A. A., Yitmen, I., Alizadehsalehi, S., & Celik, T. (2019). Adoption of virtual reality (vr) for site layout optimization of construction projects. *Teknik Dergi, 31*, 9833–9850.

Naghibi, M. (2024). Rethinking small vacant lands in urban resilience: Decoding cognitive and emotional responses to cityscapes. *Cities, 151*, Article 105167.

Najafi, P., Mohammadi, M., van Wesemael, P., & Le Blanc, P. M. (2023). A user-centred virtual city information model for inclusive community design: State-of-art. *Cities, 134*, Article 104203.

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education, 15*, 625–632.

Orsetti, E., Tollin, N., Lehmann, M., Valderrama, V. A., & Morató, J. (2022). Building resilient cities: Climate change and health interlinkages in the planning of public spaces. *International Journal of Environmental Research and Public Health, 19*, 1355.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research, 12*, 2825–2830.

Piras, G., Agostinelli, S., & Muzi, F. (2024). Digital twin framework for built environment: A review of key enablers. *Energies, 17*, 436.

Pirker, J., Loria, E., Safikhani, S., Künz, A., & Rosmann, S. (2022). Immersive virtual reality for virtual and digital twins: A literature review to identify state of the art and perspectives, in: 2022 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW). *IEEE. pp., 114–115*.

Puli, M. S., Singarapu Swetha, E., Anuhya, A., & Reddy, J. S. (2022). Urban street cleanliness assessment using mobile edge computing and deep learning. *Journal Of Algebraic Statistics, 13*, 547–552.

Ramu, S. P., Boopalan, P., Pham, Q. V., Maddikunta, P. K. R., Huynh-The, T., Alazab, M., Nguyen, T. T., & Gadekallu, T. R. (2022). Federated learning enabled digital twins for smart cities: Concepts, recent advances, and future directions. *Sustainable Cities and Society, 79*, Article 103663.

Rowland, Z., Cug, J., & Nica, E. (2022). The geopolitics of smart city digital twins: Urban sensing and immersive virtual technologies, spatio-temporal fusion algorithms, and visualization modeling tools. *Geopolitics, History and International Relations, 14*, 56–71.

Sai, S., Gaur, A., Hassija, V., & Chamola, V. (2024). Artificial intelligence empowered digital twin and nft-based patient monitoring and assisting framework for chronic disease patients. *IEEE Internet of Things Magazine, 7*, 101–106.

Schrotter, G., & Hürzeler, C. (2020). The digital twin of the city of Zurich for urban planning. PFG–journal of photogrammetry. *Remote Sensing and Geoinformation Science, 88*, 99–112.

Sepasgozar, S. M. (2020). Digital twin and web-based virtual gaming technologies for online education: A case of construction management and engineering. *Applied Sciences, 10*, 4678.

Shahat, E., Hyun, C. T., & Yeom, C. (2021). City digital twin potentials: A review and research agenda. *Sustainability, 13*, 3386.

Son, T. H., Weedon, Z., Yigitcanlar, T., Sanchez, T., Corchado, J. M., & Mehmood, R. (2023). Algorithmic urban planning for smart and sustainable development: Systematic review of the literature. *Sustainable Cities and Society, 104562*.

Tabrizian, P., Baran, P. K., Van Berkel, D., Mitasova, H., & Meentemeyer, R. (2020). Modeling restorative potential of urban environments by coupling viewscape analysis of lidar data with experiments in immersive virtual environments. *Landscape and Urban Planning, 195*, Article 103704.

Tancev, G., & Toro, F. G. (2022). Towards a digital twin for air quality monitoring networks in smart cities, in: 2022 IEEE international smart cities conference (ISC2). *IEEE. pp., 1–4*.

Tao, F., & Qi, Q. (2019). Make more digital twins. *Nature, 573*, 490–491.

Taupin, M. L., Romeas, T., Juste, L., & Labbé, D. R. (2023). Exploring the eﬀects of 3d-360° vr and 2d viewing modes on gaze behavior, head excursion, and workload during a boxing specific anticipation task. *Frontiers in Psychology, 14*.

Trantas, A., Plug, R., Pileggi, P., & Lazovik, E. (2023). Digital twin challenges in biodiversity modelling. *Ecological Informatics, 102357*.

Tzachor, A., Sabri, S., Richards, C. E., Rajabifard, A., & Acuto, M. (2022). Potential and limitations of digital twins to achieve the sustainable development goals. *Nature Sustainability, 5*, 822–829.

Vuong, Q. H. (2023). *Mindsponge theory*. Walter de Gruyter GmbH.

Wang, H., & Wang, Y. (2024). Smart cities net zero planning considering renewable energy landscape design in digital twin. *Sustainable Energy Technologies and Assessments, 63*, Article 103629.

Wang, R., Liu, Y., Lu, Y., Zhang, J., Liu, P., Yao, Y., & Grekousis, G. (2019). Perceptions of built environment and health outcomes for older chinese in Beijing: A big data approach with street view images and deep learning technique. *Computers, Environment and Urban Systems, 78*, Article 101386.

Wang, S.m., & Vu, L. H. (2023). The integration of digital twin and serious game framework for new normal virtual urban exploration and social interaction. *Journal of Urban Management, 12*, 168–181.

Wang, Z., Ito, K., & Biljecki, F. (2024). Assessing the equity and evolution of urban visual perceptual quality with time series street view imagery. *Cities, 145*, Article 104704.

Weil, C., Bibri, S. E., Longchamp, R., Golay, F., & Alahi, A. (2023). Urban digital twin challenges: A systematic review and perspectives for sustainable smart cities. *Sustainable Cities and Society, 99*, Article 104862.

White, G., Zink, A., Codecá, L., & Clarke, S. (2021). A digital twin smart city for citizen feedback. *Cities, 110*, Article 103064.

Wu, Y., Liu, Q., Hang, T., Yang, Y., Wang, Y., & Cao, L. (2024). Integrating restorative perception into urban street planning: A framework using street view images, deep learning, and space syntax. *Cities, 147*, Article 104791.

Xia, H., Liu, Z., Efremochkina, M., Liu, X., & Lin, C. (2022). Study on city digital twin technologies for sustainable smart city design: A review and bibliometric analysis of geographic information system and building information modeling integration. *Sustainable Cities and Society, 84*, Article 104009.

Xu, H., Shao, Y., Chen, J., Wang, C., & Berres, A. (2023). Semi-automatic geographic information system framework for creating photo-realistic digital twin cities to support autonomous driving research. *Transportation Research Record, 03611981231205884*.

Xu, J., Xiong, Q., Jing, Y., Xing, L., An, R., Tong, Z., Liu, Y., & Liu, Y. (2023). Understanding the nonlinear effects of the street canyon characteristics on human perceptions with street view images. *Ecological Indicators, 154*, Article 110756.

Xu, Z., Jiang, T., & Zheng, N. (2022). Developing and analyzing eco-driving strategies for on-road emission reduction in urban transport systems-a vr-enabled digital-twin approach. *Chemosphere, 305*, Article 135372.

Yıldız, S., Kıvrak, S., Gültekin, A. B., & Arslan, G. (2020). Built environment design-social sustainability relation in urban renewal. *Sustainable Cities and Society, 60*, Article 102173.

Yin, Y., Zheng, P., Li, C., & Wang, L. (2023). A state-of-the-art survey on augmented reality-assisted digital twin for futuristic human-centric industry transformation. *Robotics and Computer-Integrated Manufacturing, 81*, Article 102515.

Yosifof, R., & Fisher-Gewirtzman, D. (2024). Hybrid quantitative mesoscale analyses for simulating pedestrians' visual perceptions: Comparison of three New York city streets. *Environment and Planning B: Urban Analytics and City Science, 51*, 140–156.

Yossef Ravid, B., & Aharon-Gutman, M. (2023). The social digital twin: The social turn in the field of smart cities. *Environment and Planning B: Urban Analytics and City Science, 50*, 1455–1470.

Yurtkulu, S. C., Şahin, Y. H., & Unal, G. (2019). Semantic segmentation with extended deeplabv3 architecture, in: 2019 27th signal processing and communications applications conference (SIU). *IEEE. pp., 1–4*.

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.

Zhang, W., & Zeng, H. (2024). Spatial differentiation characteristics and influencing factors of the green view index in urban areas based on street view images: A case study of futian district, Shenzhen, China. *Urban Forestry & Urban Greening, 93*, Article 128219.

Zhang, X., Han, L., Wei, H., Tan, X., Zhou, W., Li, W., & Qian, Y. (2022). Linking urbanization and air quality together: A review and a perspective on the future sustainable urban development. *Journal of Cleaner Production, 346*, Article 130988.

Zhang, Z., Wen, F., Sun, Z., Guo, X., He, T., & Lee, C. (2022). Artificial intelligence-enabled sensing technologies in the 5g/internet of things era: From virtual reality/augmented reality to the digital twin. *Advanced Intelligent Systems, 4*, 2100228.

Zhao, X., Lu, Y., & Lin, G. (2024). An integrated deep learning approach for assessing the visual qualities of built environments utilizing street view images. *Engineering Applications of Artificial Intelligence, 130*, Article 107805.