

Thermal Comfort in Sight: Thermal Affordance and its Visual Assessment for Sustainable Streetscape Design

Sijie Yang^{a,b}, Adrian Chong^c, Pengyuan Liu^d, Filip Biljecki^{a,e,*}

^a*Department of Architecture, National University of Singapore, Singapore*

^b*School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, United States*

^c*Department of the Built Environment, National University of Singapore, Singapore*

^d*Future Cities Lab Global, Singapore-ETH Centre, Singapore*

^e*Department of Real Estate, National University of Singapore, Singapore*

Abstract

In response to climate change and urban heat island effects, enhancing human thermal comfort in cities is crucial for sustainable urban development. Traditional methods for investigating the urban thermal environment and corresponding human thermal comfort level are often resource intensive, inefficient, and limited in scope. To address these challenges, we (1) introduce a new concept named thermal affordance, which formalizes the integrated inherent capacity of a streetscape to influence human thermal comfort based on its visual and physical features; and (2) an efficient method to evaluate it (visual assessment of thermal affordance — VATA), which combines street view imagery (SVI), online and in-field surveys, and statistical learning algorithms. VATA extracts five categories of image features from SVI data and establishes 19 visual-perceptual indicators for streetscape visual assessment. Using a multi-task neural network and elastic net regression, we model their chained relationship to predict and comprehend thermal affordance for Singapore. VATA predictions are validated with field-investigated OTC data, providing a cost-effective, scalable, and transferable method to assess the thermal comfort potential of urban streetscape. Moreover, we demonstrate its utility by generating a geospatially explicit mapping of thermal affordance, outlining a model update workflow for long-term urban-scale analysis, and implementing a two-stage prediction and inference approach (IF-VPI-VATA) to guide future streetscape improvements. This framework can inform streetscape design to support sustainable, liveable, and re-

*Corresponding author

Email address: filip@nus.edu.sg (Filip Biljecki)

silient urban environments.

Keywords: urban design, urban planning, perception, human-centric GeoAI, sustainability, heat mitigation

1. Introduction

1.1. Urban heat environment and outdoor thermal comfort (OTC)

The growing effect of urban heat island (UHI) has brought the thermal environment of cities and their adaptive strategies into the spotlight [1, 2, 3, 4, 5]. UHI is detrimental to the built environment – it drives extreme climate changes [6, 7], increases energy consumption for cooling [8], and deteriorates health due to poor air quality and decreased thermal comfort [9].

Outdoor thermal comfort (OTC), the subjective evaluation of one’s satisfaction with urban thermal conditions, is a key area of study [10, 11]. Poor OTC that makes people feel hotter can raise skin temperatures, increasing the risk of heat-related illness [12, 13], and can even affect mental health [14, 15]. The growing concern around OTC has driven the need for effective OTC evaluation methods [16]. However, traditional methods, which focus on field surveys and environmental measurements, remain costly and limited in scalability and precision [17, 18, 19]. The research by [20] offers a promising example of high-resolution thermal climate mapping by integrating urban morphology and meteorological data. Nevertheless, it lacks detailed consideration of the street environment and does not incorporate survey or measurement validation of human comfort at the street level. These constraints, amid climate change’s growing impact, call for improved methods to assess urban-scale OTC, for fostering sustainable street designs that enhance thermal comfort and urban livability.

Urban environments, characterised by their intricate and complex nature, deviate significantly from the finite conditions of indoor thermal comfort (ITC) evaluation. The multitude of diverse variables present in urban environments, such as solar radiation and shade availability, complicates climatic parameter control in OTC evaluations [21]. This complexity hinders comprehensive OTC assessments in urban areas [22]. To address this challenge, [23] developed the physiological equivalent temperature (PET) for estimating outdoor thermal conditions comparable to indoor environments. Another model applied for OTC is the predicted mean vote (PMV), developed by [24], which predicts the average thermal sensation on a scale from -3 (cold) to +3 (hot), based on six environmental factors: air temperature, radiant temperature, relative humidity, air speed, clothing insulation, and metabolic rate. Recently, the universal thermal climate index (UTCI) has gained attention for its dynamic consideration of human physiological responses, making

it particularly suitable for urban environments [25, 26]. PET, PMV, and UTCI are now widely used for OTC evaluation, though the complexity of urban meteorological variables and human factors makes the process challenging [22].

Traditional OTC measurements based on these indexes rely on meteorological instruments, field investigations, and surveys [27, 28], but these methods are resource-intensive, limited in spatial coverage, and lack geographical precision [16]. Such limitations highlight the need for a scalable, cost-effective framework for urban-scale OTC assessment that integrates multidimensional data while ensuring precision and applicability in sustainable urban planning. Moreover, recent advances in ITC research have introduced active learning approaches to thermal comfort modelling, enabling more adaptive and data-efficient solutions [29, 30, 31]. These developments, coupled with the growing demand for responsive planning frameworks, have paved the way for enhanced thermal comfort assessment systems that not only integrate changing environmental factors but also align with dynamic urban planning needs [32, 33]. Together, these trends underline the shift toward a more efficient, responsive and adaptive framework for urban-scale OTC evaluation, which is essential for sustainable streetscape design and human-centric urban development in the digital era.

1.2. Visual assessment of streetscape quality based on street view imagery (SVI)

Street view imagery (SVI) is a crucial tool for assessing the quality of streetscapes and their social impacts in urban environments [34, 35, 36, 37, 38, 39]. The visual composition of streetscapes provides a wealth of multidimensional information, enabling researchers to understand the socioeconomic implications of built environments and guide improvements in urban design [40]. SVI, paired with computer vision and GIS, has become a standard method for large-scale streetscape analysis [41, 42, 43, 44, 45, 46]. It supports both objective analysis of streetscape elements, including **image features (IF)** such as geometry, colour, composition, and texture, and captures subjective **visual-perceptual indicators (VPI)** such as visual comfort, enclosure, complexity, through online surveys [47, 48, 49]. IF refers to the objective, measurable aspects of an image (SVI in urban studies) that can be extracted and analysed through computational techniques [50, 51, 52, 53, 54], while VPI, on the other hand, refers to the subjective, human-centred interpretations of streetscapes, capturing how people perceive and experience the urban environment [42, 55, 56].

SVI captures a variety of IFs that can be extracted using advanced computer vision techniques. Algorithms like PSPNet [57] and FCN [58], trained on datasets such as Cityscapes [59] and ADE20K [60], enable precise semantic segmentation of SVI. These algorithms allow researchers to quantify the proportions of elements

such as buildings, trees, and roads, and open-source packages are developed for semantic SVI data analytics in urban planning [61, 62, 63]. In addition to streetscape indices based on element percentages, other metrics have been developed to capture additional IFs of urban streetscapes in SVI. These include vehicle and pedestrian counts [64], the architectural style and building characteristics [65, 66, 67], urban scene types [68], and weather conditions [69]. These metrics, derived from specialised computer models and datasets, provide valuable spatial data for assessing urban environment.

Beyond objective IFs derived from computer vision and SVI datasets, human visual assessments of streetscapes involve subjective VPIs, which are often interrelated [36, 70]. These subjective qualities can be evaluated using SVI-based VPI questionnaires [55], typically conducted online. Participants rate or compare SVIs based on classic subjective VPIs such as imageability, enclosure, human scale, transparency, complexity, and safety, reflecting their perceptions of urban spaces [71, 36, 48]. Many studies have invited participants to perform pairwise comparisons or grading-based (such as score range 1-10) visual assessments of SVI to evaluate streetscape quality and subjective visual perception they evoke [42, 43, 72].

As the urban thermal environment and corresponding thermal comfort are closely linked to the objective composition of streetscapes and subjective visual perceptions [73, 74, 75, 76, 77], SVI has proven to be an important instrument in this domain. For instance, [76] uses 3D morphological data from SVI for street-level thermal mapping, while [77] combines SVI features with urban social-physical data for seasonal thermal assessments. When it comes to using streetscape elements and visual perception for thermal comfort evaluation, [74] identifies three key ways street greenery enhances thermal comfort across different street types, based on SVI and micrometeorological measurements: lowering radiant temperatures, improving streetscape aesthetics, and positively influencing conscious perceptions of microclimate. Similarly, after the evaluation of streetscape design through satellite images and SVI, local climate conditions, and ENVI-met modelling data, [75] suggests that thermal comfort can be improved through open spaces and thoughtful planting choices. Therefore, comprehensive visual assessments of streetscape quality not only enhance aesthetics and functionality but also improve thermal comfort and urban liveability.

1.3. Affordance information from SVI in urban studies

The theory of affordance, proposed by [78], highlights how the ecological environment shapes perception and behaviour. Applied to the built environment, affordances represent environmental features that promote human perception and behaviour [78, 79]. According to [78], the environment acts as a ‘surface’ medium

through which humans perceive and interact with their surroundings, offering clues to potential actions. The informational content and environmental characteristics embedded within the 'surfaces' constitute the 'value' and 'meaning' of a place. These features, defined as affordances and inferred from environmental attributes, indicate likely human perceptions and behaviours even without direct experience. [78] further emphasises that this inherent affordance of the environment can be deduced through visual assessment.

Affordance theory has been employed to elucidate the relationship between the built environment and human mobility behaviour, utilising geospatial indicators such as greenery index, street network morphology metrics, and urban function density, to comprehend how environments support activities like walking, biking, and running [80, 79, 72, 81]. The built environment provides a foundation for urban life, influencing social activities, events, human movement, and exploration. Understanding these affordances is crucial for improving city conditions. SVI, as a direct and human-centred data source, offers valuable insights into the built environment's affordances. Until now, in conjunction with SVI, affordance theory has been used to explain the evaluations of playability [82], runnability [79], and spatial navigation [83], which has been shown to be a robust theoretical framework for explaining the relationship between urban attributes and human perception or behaviour [79].

Both objective SVI indices and its subjective VPIs derived from human visual assessment offer insights into the affordances of urban streetscapes. For example, greenery indices derived from these images link higher green views to increased pedestrian activity and improved mental well-being [84, 85], while visual complexity affects perceived safety and comfort [86]. By combining objective SVI analysis with subjective SVI assessment surveys, we can explore various urban spaces and their perceptual and behavioural potentials across different urban scales [87, 70]. This affordance analysis framework, using both objective and subjective metrics, effectively evaluates urban environments and streetscape design [88, 89]. These assessments can reveal a variety of human behaviours and urban outcomes, including walking and cycling patterns [90], property values [89, 70], and public health indicators [91, 92]. Superior streetscape design improves the visual experience, walkability, bikeability, environmental appeal, and public activities, boosting the value of urban streets [93, 94, 95].

1.4. Research gaps and our response

The interrelation between thermal comfort and streetscape design constitutes a significant area of study [96]. Based on existing literature [74, 42, 75, 76, 77], we posit that both the objective composition of streetscape elements and subjective

visual assessment could serve as proxies for evaluating a streetscape’s capacity to influence human thermal comfort. However, key research gaps persist.

Firstly, there is no clear concept for assessing the built environment’s objective properties that influence thermal comfort in streetscape design. Although many studies have explored associations between streetscape elements, outdoor thermal environment, and OTC [97, 98, 99, 77, 76], a theoretical framework that explains these relationships remains underdeveloped. This limits the creation of a convincing framework for sustainable streetscape design that improves outdoor thermal comfort.

To address this gap, we introduce the concept of **thermal affordance** (see details in Section 2), inspired by Gibson’s theory of environmental affordance [78]. Our novel concept highlights the potential of urban design to promote thermal comfort through streetscape element configurations, forming the foundation for a data-driven framework to analyse their impact on OTC.

Secondly, the use of human visual assessment in evaluating streetscapes’ thermal conditions and their capability to impact thermal comfort is rarely studied. While some research has touched on the interaction between visual information from SVI and outdoor thermal environments [74, 100, 101, 102, 76], they have not applied human visual assessment to examine how streetscape element composition influences thermal comfort. Some studies predict thermal environments using SVI [77, 76], but they overlook OTC and lack validation through field experiments. Therefore, a methodological approach incorporating SVI and human visual assessment is needed for urban-scale OTC evaluation, validated by in-field OTC measurements. Moreover, while it is widely accepted that street greenery improves OTC [74, 96, 103], the influence of various streetscape elements on perceived thermal comfort remains underexplored, especially regarding buildings, roads, and vehicles.

Our proposed **visual assessment of thermal affordance (VATA)** framework fills this gap by using SVI data and online questionnaires to quantify thermal affordance by integrating **image features (IF)** and human **visual-perceptual indicators (VPI)** from online surveys. VATA provides a novel approach to urban OTC evaluation.

Third, the systematic application of multi-source urban data, such as SVI, human visual-perceptual surveys, and in-field comfort and physiological indicators, has not been fully adopted as an efficient workflow for assessing OTC in urban environments. Although some studies have linked SVI features with the urban thermal environment [100, 104, 105, 76, 106], a verifiable, replicable, and street-level methodology for urban-scale OTC evaluation supported by multi-source urban data has not yet been built.

To address this methodological shortfall, our VATA framework introduces a

data-driven workflow that integrates SVI, human visual assessments, and in-field comfort data. This approach enables the development and validation of prediction and inference models based on our proposed concept of thermal affordance, facilitating scalable and precise OTC evaluation for sustainable urban planning.

1.5. Research question

Given the identified research gaps, we propose the following research questions:

1. How can the concept of thermal affordance be utilised as a theoretical framework to describe and assess urban street environment properties that promote relative thermal comfort potential?
2. How can the visual assessment of thermal affordance (VATA) framework be constructed based on elements that can be extracted from street view images (SVI) — image features (IF) and human visual-perceptual indicators (VPI) — to perform prediction and inference models to efficiently assess and understand urban-scale streetscapes' OTC potential (thermal affordance) for sustainable urban planning?
3. How can we develop a multi-source data analysis workflow, incorporating in-field outdoor thermal comfort (OTC) investigation data, to validate the VATA framework?
4. What analytical insights and urban planning frameworks can be derived from the data outputs of the VATA framework to guide sustainable streetscape design?

2. The idea of thermal affordance

2.1. Proposed definition of thermal affordance

As Gibson's theory of affordance suggests [78], environments contain inherent values and information that shape human perceptions and behaviours. Thermal comfort, as part of human perceptions, reflects both subjective satisfaction and objective factors such as air temperature, air humidity, and wind speed [24, 107, 108, 109, 110]. We introduce the concept of '**thermal affordance**' to describe the inherent capability of an environment to impact thermal comfort. This concept integrates various environmental factors, indicating the possible thermal comfort experienced. Studies have shown the connection between thermal comfort and integrated environmental factors such as walkability [97] and streetscape design [75], supporting the validity of thermal affordance.

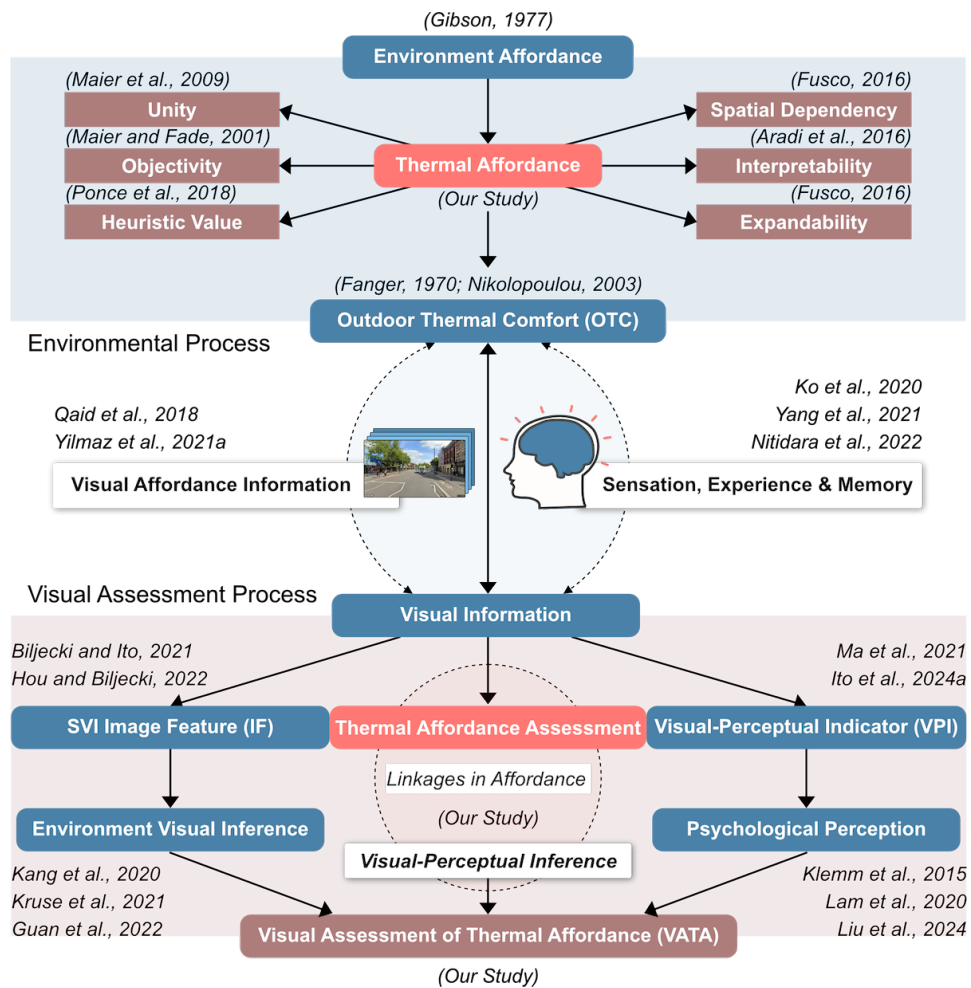


Figure 1: VATA conceptual framework: from thermal affordance to its visual assessment. Source of the SVI: Google Street View.

Currently, the literature on thermal comfort presents various interconnected concepts, each providing a different lens to understand how individuals experience thermal environments [111]. For instance, thermal comfort is defined by the ASHRAE Standard 55 as a 'condition of mind that expresses satisfaction with the thermal environment', emphasising a subjective evaluation [112]. Meanwhile, thermal neutrality refers to the conditions where the majority of individuals feel neither warm nor cold [113, 112]. Concepts such as the thermoneutral zone and thermal sensation further highlight the physiological and psychological dimensions of thermal comfort [114], with the latter focussing on how people consciously per-

ceive temperature [115, 112].

These established concepts offer significant insights, but they remain largely subjective, relying heavily on individuals' preferences or perceptions. In contrast, our proposed notion of thermal affordance aims to transcend individual subjectivity by focussing on the inherent characteristics of the environment itself and its potential to influence thermal comfort. While more concepts like thermal expectation or thermal satisfaction consider personal inclinations and thermal experiences [116, 107, 117], thermal affordance unifies objective environmental features, offering a more scalable and integrative framework for OTC improvement in urban planning and design.

The necessity of introducing thermal affordance lies in its ability to bridge the gap between subjective experiences of thermal environments and the objective environmental factors that shape them. Such an approach allows for a more comprehensive analysis, helping urban planners make informed decisions that improve thermal comfort by connecting measurable environmental properties with subjective or individual-based assessments.

We summarise several key characteristics of thermal affordance: unity, objectivity, heuristic value, spatial dependency, interpretability, and expandability, as illustrated in Figure 1.

- **Unity** indicates that thermal affordance aims to encompass all fixed variables that influence thermal comfort in the environment into a unified set for description, based on the integration and unity properties of affordance proposed by [118].
- **Objectivity** highlights the objective environmental information encapsulated in thermal affordance, unaffected by individuals' subjective perspectives, aligning with the idea of affordance as an integration of objectivity in the environment [119].
- **Heuristic value** of thermal affordance suggests that the information it encompasses can inspire understanding and analysis of thermal comfort, and the various dimensions of information within thermal affordance can mutually infer and enlighten one another. For instance, [120] suggests the possibility that visual data serve as heuristics for thermostats.
- **Spatial dependency** highlights variations in thermal comfort caused by different spatial environments. Differences in spatial or environmental properties have been shown to directly lead to a difference in affordance [80]. Thermal affordance does not stress absolute levels of thermal comfort, but rather emphasises spatial differences between environments in their capacity to promote thermal comfort.

- **Interpretability** means that thermal affordance has a clear inference relationship with environmental attributes, as environmental affordance is composed of multiple influencing parameters [121]. Comprehending this inference relationship can aid urban planners in enhancing thermal comfort through urban design.
- **Expandability** means that thermal affordance can continuously incorporate additional environmental variables to update information on an environment’s capacity to influence thermal comfort. More dimensions of information, such as lifestyle and dwelling mode [80], have been shown to allow for a more comprehensive evaluation and analysis of environmental affordance.

2.2. *Visual assessment of thermal affordance (VATA) framework*

This study identifies four key challenges in urban-scale OTC evaluation (Figure 2): (1) Measurement of thermal comfort requires substantial resources – labour, budget, and time [122]. (2) Current methods, relying on physical instruments and surveys, can only cover limited urban areas [123]. (3) Geographic precision is constrained for city-wide OTC assessments [124]. (4) Urban OTC measurements are affected by many uncontrollable environmental variables, unlike controlled indoor experiments [10]. To address these challenges, we developed the **VATA (visual assessment of thermal affordance)** framework to enhance urban-scale OTC assessments and improve understanding of urban thermal environments.

This research continues to explore the intrinsic connection between OTC in urban environments and visual data from SVI, shaped by affordance information and personal experiences. First, SVI visual data reflect objective characteristics of streetscapes, indicating streetscapes’ potential capability to promote thermal comfort and microclimate conditions [125, 126], which forms thermal affordance. Second, this visual information is linked to individuals’ thermal perception, influenced by memory and sensory experiences [127, 128, 129], which affects their thermal affordance assessment. These insights form the foundation of our VATA analysis, highlighting the role of visual characteristics and perceptual experiences in OTC assessment. Using machine learning, we combine SVI **image features (IF)** and **visual-perceptual indicator (VPI)** survey data to predict the VATA metric, allowing us to visually assess the urban environments’ capability to promote thermal comfort. Our model shows promise for streetscape assessment and VATA prediction, as illustrated in Figure 2.

Objective IF from SVI reflect the urban environment from an individual’s perspective in a three-dimensional, objective manner [35, 130], capturing elements such as greening rate [131], tree canopies[132], building information [65], urban function [133], and urban form [134]. These features, closely related to OTC, serve

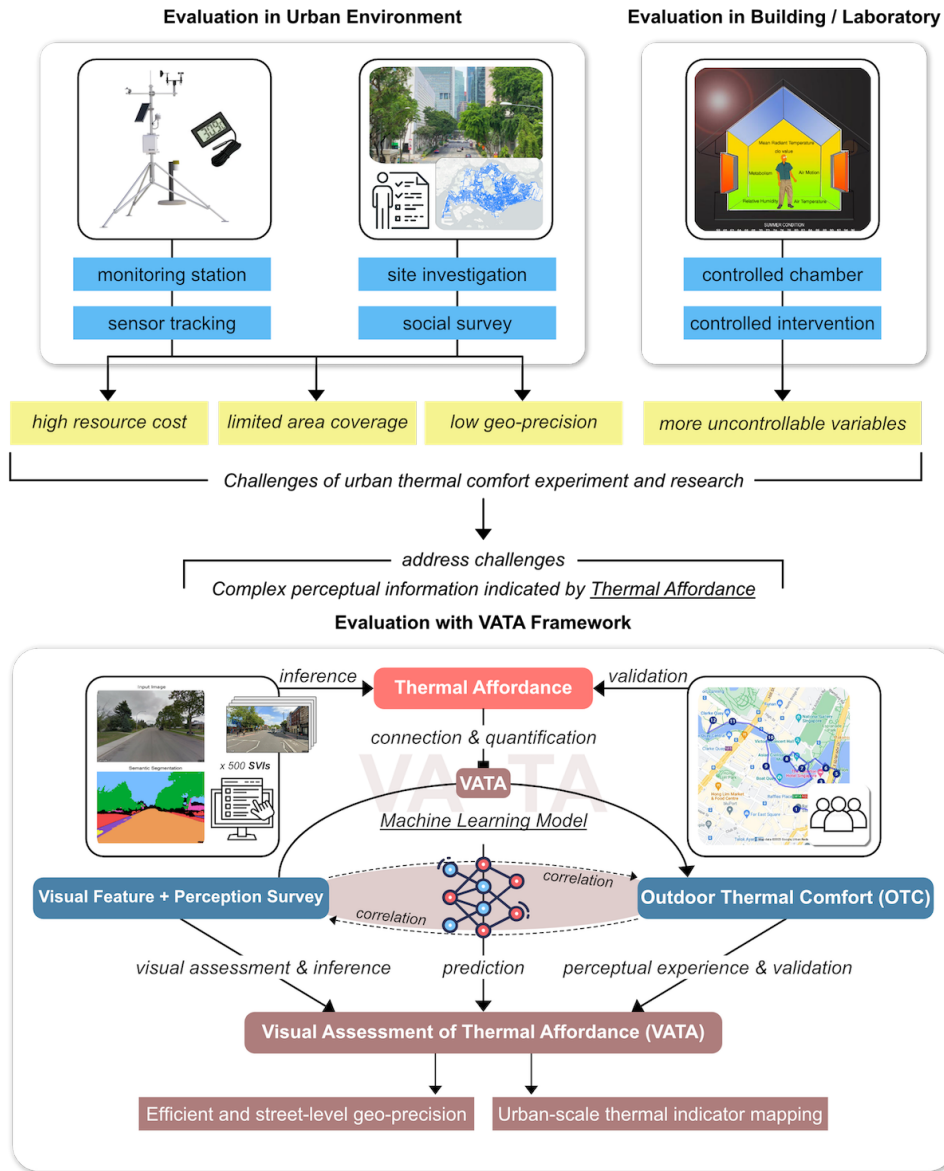


Figure 2: Visual Assessment of Thermal Affordance (VATA) framework: from challenges to a solution.

as affordance information since OTC is influenced by surrounding urban elements [135, 136], enabling visual-environmental inference in urban studies [42, 82, 137]. The VPI information obtained through the SVI visual assessment surveys further

accounts for urban perceptions of individuals [138, 56], impacting socioeconomic dimensions such as bikeability [90], running ability [139] and property prices [70]. VPI and OTC, as psychological perceptions, subtly influence each other [74, 140, 141, 142]. Figure 1 illustrates the relationships between IF, VPI, VATA and OTC, demonstrating that the use of SVI for the evaluation of VATA and OTC is grounded in both environmental and perceptual principles. These relationships are not directly derived or scored from the survey data, as the problem remains a “black box” with only separate scores available for IF, VPI, VATA, and OTC. Accordingly, machine learning techniques—including both predictive and inferential modelling—are used to uncover and quantify these underlying relationships.

From another perspective, the intricate relationship between VATA and OTC can be deeply rooted in individuals’ long-term memory and experiential knowledge, as people’s sensory experience in urban environments is a long-term process. This kind of environmental experience goes beyond short-term observations and involves complex reasoning and memory processes in the brain [143, 144, 145]. For example, the association of greenery with shading and cooling stems from prolonged human experiences [146, 147]. This knowledge enables individuals to infer OTC from visual cues, similar to how greenery is linked to cooler temperatures and better thermal comfort [148, 149, 150]. Thus, subjective streetscape visual assessments in our VATA framework are also grounded in the brain’s ability to interpret environmental cues based on past experiences.

While VATA provides an efficient means of quantifying thermal affordance using SVI data, it is important to acknowledge that integrating meteorological information, detailed geospatial data, remote sensing imagery, and socio-cultural contexts can yield more comprehensive insights into how multiple factors collectively influence urban thermal comfort [27, 151, 152]. Incorporating additional environmental variables could further enhance the VATA framework by employing advanced data fusion and machine learning techniques, ultimately improving predictive accuracy and informing more effective urban design strategies.

3. Methodology

3.1. Research framework and study area

Figure 3 presents the research framework for this study, built on the VATA framework. We conducted an online SVI visual assessment survey, evaluating VATA and 19 other VPIs based on 500 SVIs. Five classes of IFs from SVI data, along with survey-based VATA and VPI data, were used to develop datasets for statistical VATA prediction and inference models. Using machine learning models, we can predict and infer relationships between IFs, VPIs, and VATA based on their respective scoring data, all drawn from the same set of 500 SVIs used in the

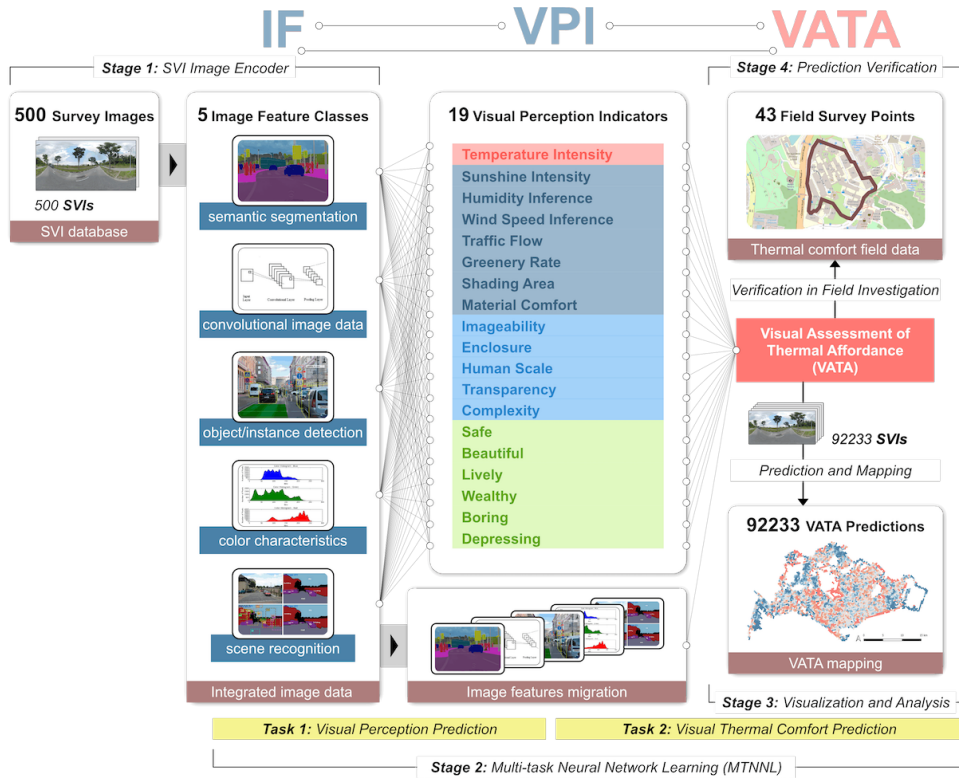


Figure 3: Research framework based on IFs from SVI data, VPIs from online SVI visual assessment survey, and a two-stage MTNNL model for VATA prediction.

visual assessment survey. A multi-task neural network learning model (MTNNL) was constructed with two stages: predicting VPIs from IFs, and then predicting VATA from VPIs, using weighted loss values for iterative training. This model was applied to predict VATA for 92,233 SVIs in Singapore, and validated against real-world OTC data. A two-stage elastic net regression model (ENRM) was also used to interpret IF-VPI-VATA relationships, offering insights for streetscape design. SVIs are sourced from Google Street View [153].

Singapore serves as the study area due to its rich urban landscape design, diverse urban forms, consistent outdoor temperatures (25-31 °C year-round) and perennial need to ensure OTC as a tropical city [154]. Our study includes an on-line SVI visual assessment survey, model prediction and inference, and validation using OTC field data, aiming to propose a more efficient method for urban-scale OTC assessment.

3.2. Online SVI assessment survey setting

Table 1: Selected VPIs and associated questions for the online SVI assessment survey.

Visual-Perceptual Indicator (VPI)	Question
<i>Thermal Affordance Assessment</i> Visual Assessment of Thermal Affordance (VATA)	Which street view image do you perceive as having a more comfortable outdoor thermal environment for you?
<i>Microclimate Inference</i> Temperature Intensity / Sunshine Intensity / Humidity Inference / Wind Speed Inference	Which street view image do you perceive exhibits a higher outdoor temperature / sunlight intensity / humidity inference / wind speed inference?
<i>Environmental Evaluation</i> Traffic Flow / Greenery Rate / Shading Area / Construction Material Comfort	Which street view image do you think showcases higher (more) traffic flow / greenery rate / shading areas / construction material comfort?
<i>Design Quality</i> Imageability / Enclosure / Human-Scale / Transparency / Complexity	Which street view image stands out to you as a more impressive place / enclosed space / accommodating for human scale / transparent space / complex environment?
<i>Evoked Emotion</i> Safe / Beautiful / Lively / Wealthy / Boring / Depressing	Which street view image do you feel evokes a more safe / beautiful / lively / wealthy / boring / depressing atmosphere?

The study conducts an online SVI visual assessment survey on VATA and 19 other VPIs among 176 long-term Singapore residents using 500 SVIs. The survey aimed to capture residents’ subjective visual assessment of thermal affordance-related inferences and perceptions toward various streetscapes shown in random SVIs. These results, combined with IFs, were used to train the MTNNL model to predict the VATA metric, reflecting street-level thermal affordance evaluations.

Along with VATA, the 19 VPIs are grouped into four categories: microclimate inference, built environment evaluation, streetscape design quality, and evoked emotion (Table 1). These VPIs were chosen to capture participants’ holistic visual-perceptual impressions of streetscapes, rather than having them rate individual streetscape elements. Participants were presented with system-randomised pairs of SVIs from the online survey for each indicator, which they compared directly. This approach encouraged a more natural expression of their perceptions, free from the pressure of explicit grading tasks. The TrueSkill algorithm [155] was subsequently employed to analyse responses, generating rankings and scores for the 500 SVIs on VATA and each VPI. These scores were then normalised to a 0-5 scale for machine learning regression (Figure 4). The ranking algorithm consistently updates the ranking score of winners and losers from pairwise comparisons, utilising the survey results comprising 3168 pairs collected from all the 176 participants for

scoring calculation. Since one assumption of the TrueSkill algorithm is that the performance of all ranking elements (500 images in this case) follows a normal distribution, we can ultimately visualise the scoring results as a normal distribution within a range of 0 to 5 after normalisation, with the ideal population mean of 2.5. The TrueSkill process was repeated 20 times to obtain scores for VATA and each VPI. Details of the TrueSkill algorithm and survey results are introduced in Appendix A and Appendix B. These TrueSkill-derived scores enable us to employ predictive and inferential modelling to construct the IF–VPI–VATA relationship.

To enhance model training based on SVIs and corresponding indicator scores from the online SVI visual assessment survey, we apply six data augmentations—resizing, normalising, random clipping, rotating, colour jittering, and cropping—to the 500 images. These transformations enhance the model’s generalisation and robustness (Figure 4). For VATA and VPI scores, we set the standard deviation (std) of each scoring distribution to 1, balancing discriminative power and an even score distribution. As illustrated in Figure 4, a std of 0.5 centralises the scores, limiting range, while a std of 1.5 flattens the distribution, reducing distinction. Setting std to 1 strikes an optimal balance between clarity and evenness.

VATA and 19 VPIs scores for the 500 selected images were derived through an online pairwise visual comparison of SVI, using the TrueSkill ranking algorithm.

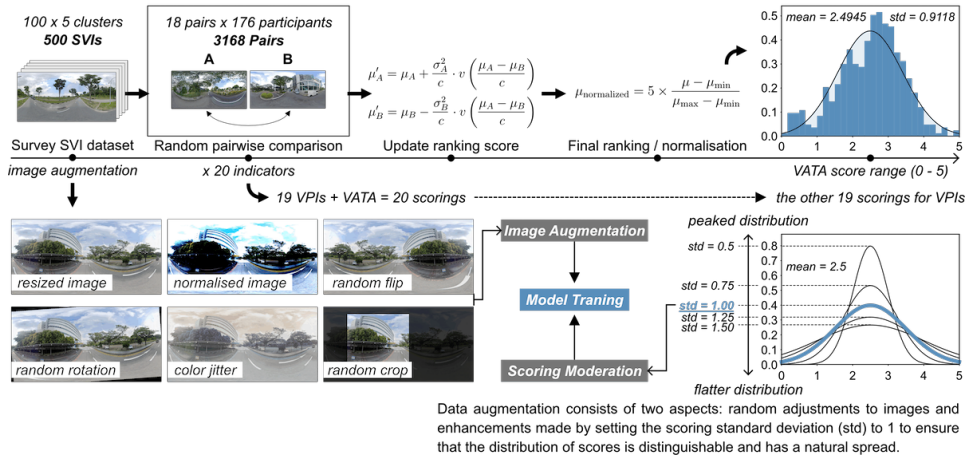


Figure 4: Details on indicator scoring and data augmentation in the SVI visual assessment survey.

To ensure diverse streetscape representation, 500 SVIs were selected from a dataset of 92,233 SVIs across Singapore through the following process:

1. *k*-means clustering of all SVIs based on 19 streetscape elements from semantic segmentation, following the streetscape classification method outlined by [37].

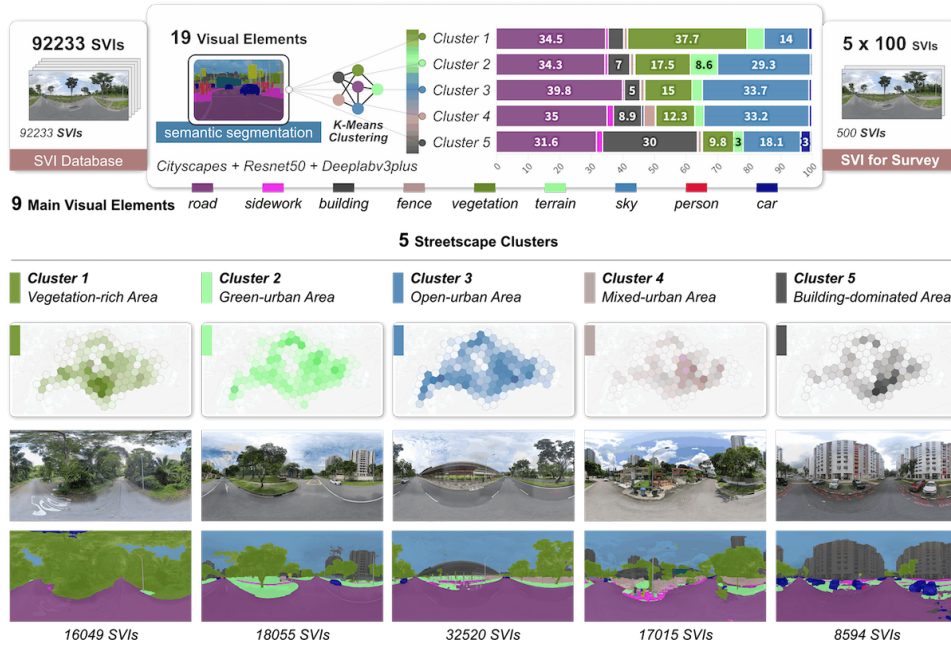


Figure 5: SVI clustering based on segmentation results, and selection process for 500 SVIs taken for the online SVI visual assessment survey.

2. Categorisation into five types of streetscape: vegetation-rich, green-urban, open-urban, mixed-urban, and building-dominated.
3. Random selection of 100 images from each category, totalling 500 images for the survey.

Figure 5 illustrates the streetscape element proportions and spatial distributions of these categories. This balanced selection ensures equitable representation of each streetscape type, enhancing the VATA model’s effectiveness by encompassing diverse urban environments. This approach underscores the importance of diversity and representativeness in training data for machine learning applications in complex urban settings.

3.3. IF extracted from SVI data for model training and prediction

SVI data captures key aspects of urban street environments, such as greening, architectural style, and overall urban scenes, which features delineate the objective streetscape characteristics that contribute to thermal affordance. By combining IFs from 500 SVIs with visual assessment survey results, we train a MTNNL to model 19 VPIs and VATA effectively.

As illustrated in Table 2, five categories of IF can be extracted from a single SVI i using various algorithms and datasets: semantic segmentation identifies streetscape elements (e.g., buildings, vegetation), convolutional data capture high-dimensional pixel representations, object recognition counts objects (e.g., people, cars), pixel feature analysis examines colour features, and scene recognition assesses the likelihood of displaying specific scene type (e.g., commercial, natural). Pre-trained models like DeepLabV3+ [156], ResNet-50 [157], Faster R-CNN [158], and OpenCV [159], along with datasets like Cityscape [59], COCO [160], and Places365 [161], are used to extract these IFs.

Each SVI ($h \times w$ pixels) undergoes feature extraction, yielding five tensors concatenated into a single tensor $X = (P, F, A, C, S)$. For a dataset of m images, this forms a tensor D_1 of size $N \times m$, where N is the total number of features in X_i . We also create dataset tensor D_2 for an ENRM-based VATA inference model, excluding convolutional features for better interpretability. Thus, this study provides two IF datasets for modelling: $D_1 = [P_i, F_i, A_i, C_i, S_i]$ and $D_2 = [P_i, A_i, C_i, S_i]$ for $i = 1, \dots, m$. D_1 serves as input for the MTNNL-based VATA prediction model, while D_2 trains the ENRM-based VATA inference model, showing interpretable correlations between IF, VPI, and VATA. Sub-features of IF under each category are listed in Appendix C.

Table 2: Models and datasets applied for IF extraction from SVI data.

Extraction Method	Model/Lib	Dataset	Symbol	Feature Extracted
Semantic Segmentation	DeepLabV3+	Cityscape	P_i	19 streetscape elements
Convolutional Data	ResNet-50	-	F_i	convolutional tensor
Object Detection	Fast R-CNN	COCO	A_i	91 object categories
Pixel Feature	OpenCV	-	C_i	12 pixel visual feature
Scene Recognition	ResNet-50	Places365	S_i	365 scene categories

3.4. Architecture of VATA prediction and inference model

VATA employs two main model architectures: the multi-task neural network learning (MTNNL) for predictive analysis and the elastic net regression model (ENRM) for inference analysis. MTNNL is used for predicting VATA outcomes based on SVI and human visual assessment results, leveraging its strong predictive power. ENRM, as a linear and interpretable model, addresses collinearity issues and is suited for VATA inference research. With separate scores obtained for IFs, VPIs, and VATA for each image in the sample, we can use MTNNL to predict VPIs and VATA scores from IF data, and employ ENRM to quantitatively determine how the visual environmental features encapsulated in IFs influence both VPIs and VATA. In this process, VPIs act as an intermediate layer that not

only improves VATA prediction accuracy but also deepens our understanding of the IF–VPI–VATA relationships.

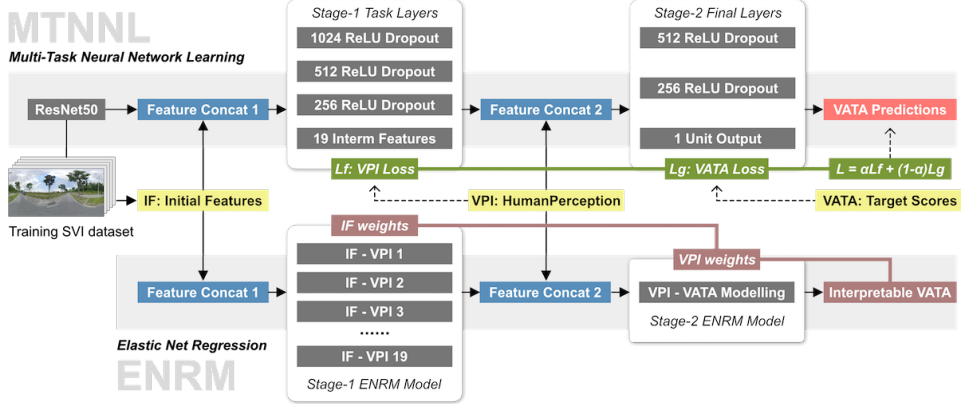


Figure 6: Overall architecture of VATA prediction and inference modelling

3.4.1. MTNNL Model for VATA prediction

A MTNNL model is employed as the VATA prediction model and is trained in two simultaneous stages, as shown in Figure 6. The model architecture is designed to allow for concurrent learning of two related tasks, each corresponding to a different stage of training. In the first stage, the model is trained on the dataset tensor D_1 , which contains five types of IFs in m SVI images. Simultaneously, an intermediate tensor V is created, representing human visual-perception rating data for the same m images, where each image is graded on a 0-5 scale across 19 human VPIs derived from our online SVI visual assessment survey.

$$D_1 = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \cdots \quad \mathbf{x}_m], V = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_m] \quad (1)$$

Here, \mathbf{x}_i is an n -dimensional feature vector for the i -th image, and \mathbf{v}_i is a 19-dimensional perception vector for the same image. The model learns a function $f : D_1 \rightarrow V$ mapping the image feature space to the human visual perception indicator space.

In the second stage, D_1 and V are combined to predict VATA scores, also on a 0-5 scale, learning a function $g : (D_1, V) \rightarrow Y$, where Y represents the VATA tensor for m images. The neural networks for these two stages are trained simultaneously, and overall learning process involves minimizing the following integrated loss function:

$$L = \alpha L_f + (1 - \alpha) L_g \quad (2)$$

where L_f is the loss for predicting the VPI tensor V from D_1 , and L_g is the loss for predicting VATA from D_1 and V . The parameter α controls the balance between the two tasks.

During the training phase, the visual IF data derived from 500 images across five feature extraction aspects, along with the VPI perceptual data across 19 metrics and the final VATA data, will be partitioned into a training set, a validation set, and a test set, adhering to a ratio of 60%, 20%, and 20%, respectively. The model's final predictive performance will be assessed based on its MAE, MSE, RMSE, and adjusted R^2 on the test set, while the validation set plays a critical role during the training process by helping to tune hyperparameters, select the best model configuration, and monitor overfitting. The details of the model training data and process are presented in Appendix E.

3.4.2. ENRM Model for VATA inference

The ENRM model is used to establish an interpretable relationship between IF in SVI data, human VPI, and VATA in our studies. Combining ridge (L2) and lasso (L1) regularizations, ENRM effectively handles feature collinearity and performs variable selection. The model coefficients indicate the importance of each feature in explaining VATA scores.

Given a dataset tensor D_2 containing visual feature data from m images, the ENRM can be mathematically formulated as:

$$\hat{y}_i = \mathbf{x}_i^\top \beta \quad , \quad \|\beta\|_1 = \sum_{j=1}^p |\beta_j| \quad , \quad \|\beta\|_2^2 = \sum_{j=1}^p \beta_j^2 \quad (3)$$

where \hat{y}_i is the VATA score for the i -th image, \mathbf{x}_i is the feature tensor of the i -th image, β is the coefficient tensor. Combining these elements, the ENRM can be mathematically formulated as:

$$\min_{\beta} \left\{ \frac{1}{2m} \sum_{i=1}^m (y_i - \mathbf{x}_i^\top \beta)^2 + \lambda_1 \|\beta\|_1 + \lambda_2 \|\beta\|_2^2 \right\} \quad (4)$$

where λ_1 and λ_2 are regularisation parameters for L1 and L2 penalties, respectively.

The L1 regularisation component helps in feature selection by shrinking irrelevant feature coefficients to zero, simplifying the model. The L2 regularisation component addresses multicollinearity among features, ensuring stable coefficient estimates. This model provides interpretable insights into the two-stage inference relationship between IF, VPI, and VATA scores (Figure 6). By identifying key visual features and perceptions contributing to VATA in urban settings, the ENRM aids in developing informed urban design strategies.

3.5. VATA validation based on OTC dataset

To validate the effectiveness of the VATA framework in assessing human thermal comfort in urban street environments, we employed an on-the-ground OTC field survey dataset and data-driven statistical evaluations. The validation utilised OTC measurements from field investigations conducted on the campus of the National University of Singapore (NUS) [162]. These field investigations were carried out independently from both the SVI-based visual assessments and the subsequent VATA modelling efforts, ensuring that there was no pre-established alignment or dependency between the datasets. This relative independence enhances the robustness of the validation process by minimizing the risk of bias and reinforcing the credibility of the comparative analysis.

The survey involved 15 NUS students and 43 investigation points along a continuous path (Figure 7). The participants walked the path, providing comfort ratings (1-10 scale) at each point. They wore Apple Watches with the Cozie app [163] to collect subjective feedback and heart rate data. Sound and solar intensity were measured using a sound metre and a smartphone, respectively. These 'thermal walk' experiments [164] were carried out three times daily (10 am, 2 pm, 5 pm) over five consecutive days in July 2022. The resulting OTC dataset includes average comfort scores, heart rate values, sunlight intensity, noise levels, and altitude measurements. We specifically focused on these parameters—readily captured via portable devices such as smartphones and wearable sensors—due to their potential for rapid variability across short distances. This enables fine-grained assessment of thermal comfort differences along the walking route, without the added complexity of factors like humidity or air temperature that tend to remain relatively uniform over these small spatial scales. While including additional factors such as humidity or air temperature measurements could enhance the comprehensiveness of comfort modelling, doing so is not the central focus of this study. Additionally, although Singapore provides publicly available air temperature data, the uneven spatial distribution and limited number of weather stations (where sensors are located) create significant uncertainties when interpolating localized conditions. Furthermore, the lack of measurements at very short intervals further reduces the granularity of these data for our fine-scale analysis.

In addition, Google Street View images corresponding to each of the 43 investigation points were collected to facilitate visual-based VATA predictions and validations. The Google Street View imagery used was captured in the same year as the field investigation (2022), thus ensuring temporal synchronization. This alignment in timing allowed for a more reliable comparison and modelling between the VATA prediction scores and the on-site OTC measurements. The validation methodology, described in Figure 7, comprises two parts: 1) comparing VATA with OTC to assess thermal affordance effectiveness by directly comparing SVI-predicted VATA

at 43 data points with corresponding OTC values; and 2) using a multivariate linear modelling approach combining VATA and HSNA indicators (heart rate, solar intensity, noise level, and altitude) to evaluate VATA reliability in predicting OTC.

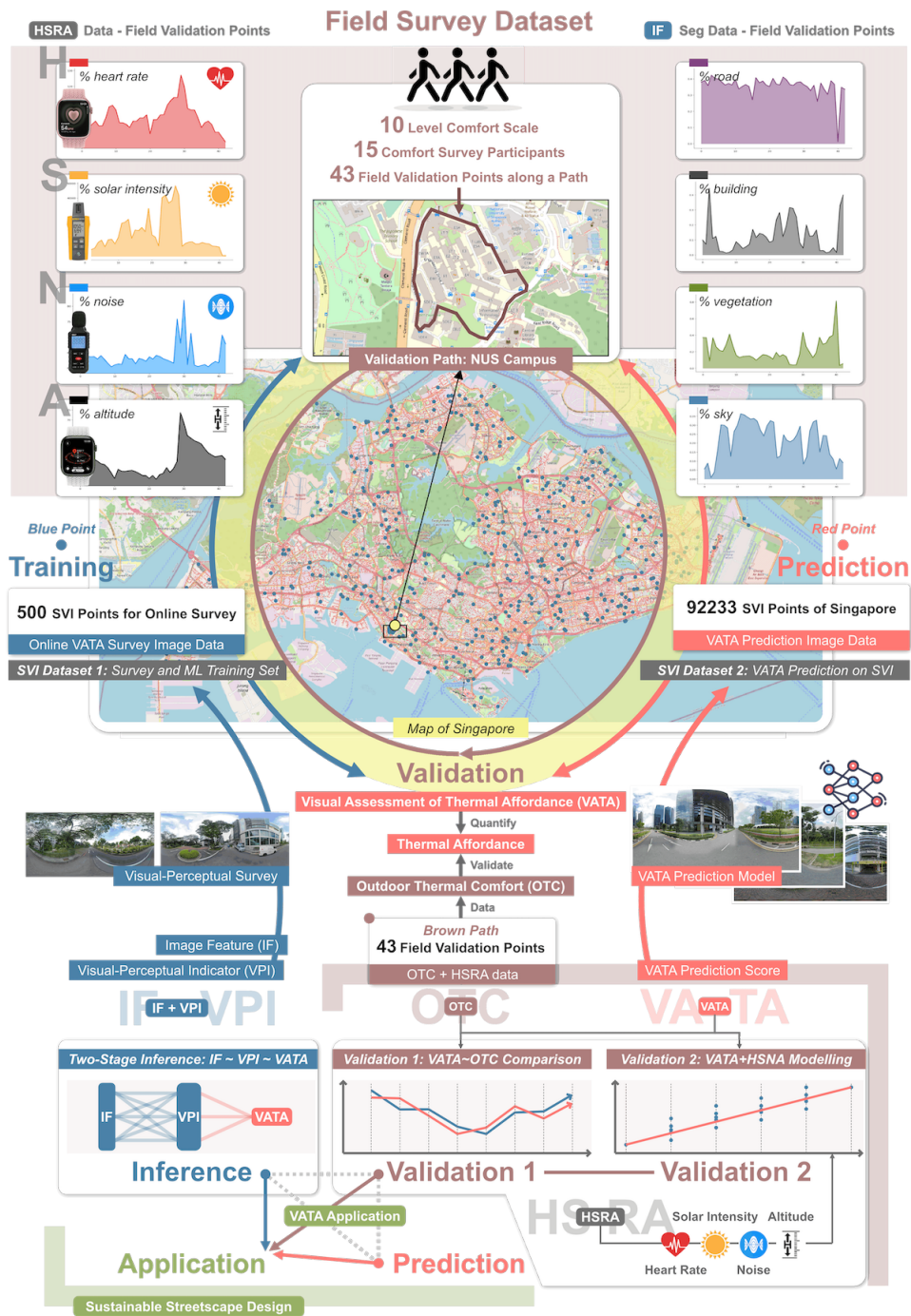


Figure 7: Details of applying OTC field survey dataset as a validation method for VATA framework

4. Results

4.1. VATA prediction model performance

Table 3: VATA prediction model performance.

Model	MAE	MSE	RMSE	adjusted R^2
Decision Tree	0.4509	0.7430	0.8620	0.1357
KNN Regression	0.4192	0.6340	0.7963	0.2625
Support Vector Regression	0.3929	0.5858	0.7654	0.3186
ENRM (Lasso + Ridge)	0.2402	0.3159	0.5621	0.6325
LightGBM	0.2422	0.3106	0.5573	0.6388
Random Forest	0.2161	0.2919	0.5403	0.6604
Extreme Gradient Boosting	0.2282	0.2906	0.5391	0.6619
CatBoost	0.2260	0.2888	0.5374	0.6640
MTNNL	0.1960	0.2339	0.4836	0.7316

The MTNNL model has demonstrated superior performance in predicting VATA on the test set. As presented in Table 3, a comparative analysis was conducted against various models, including Decision Tree, KNN Regression, Support Vector Regression, ENRM, LightGBM, Random Forest, Extreme Gradient Boosting, and CatBoost. The results show that MTNNL consistently outperforms these models, achieving the highest adjusted R^2 (0.7316) and the lowest MAE (0.1960). This improved performance is attributed to the model’s two-stage, two-task prediction architecture, which leverages the advantages of deep learning for complex prediction tasks. Instead of directly predicting VATA scores from IF, the MTNNL model incorporates an intermediary variable layer, VPI, to account for human visual assessments, as shown in Figure 3. This two-stage approach enhances the model’s ability to accurately predict VATA scores.

4.2. VATA framework validation results

4.2.1. SVI-predicted VATA score vs. surveyed comfort data

Using our OTC field survey dataset, we analysed the relationship between SVI-predicted VATA scores and surveyed subjective comfort data across 43 discrete points along a path. To account for non-contiguous data points, we applied an exponential moving average (EMA) to smooth the VATA and OTC data, with the smoothing formula defined as:

$$S_i = \alpha \cdot X_i + (1 - \alpha) \cdot S_{i-1} \quad (5)$$

where S_i is the smoothed value at the i -th data point, α is the smoothing factor (between 0 and 1), X_i is the actual value at the i -th data point, and S_{i-1} is the

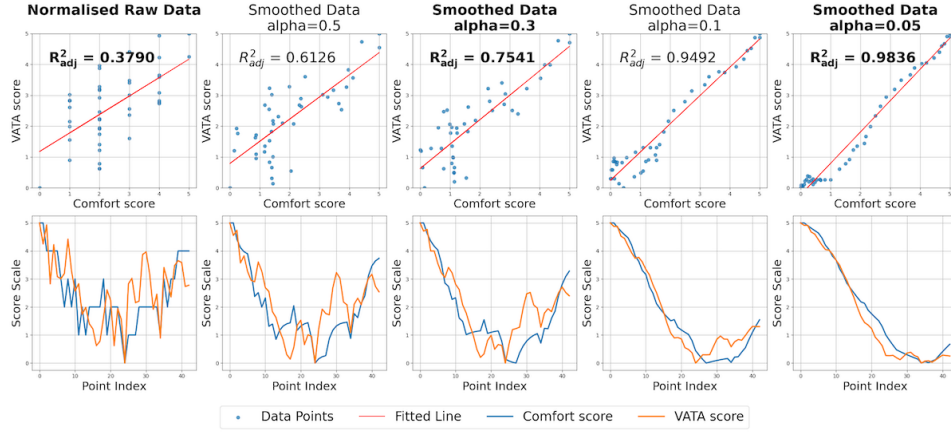


Figure 8: VATA predictions fitting surveyed thermal comfort data along the path ($p < 0.001$).

smoothed value at the $i - 1$ -th data point. Figure 8 present the results for different smoothing factors.

The raw data yielded an adjusted R^2 of 0.3790, indicating that VATA explained 37.90% of the variance in comfort data. As the smoothing factor decreased from 0.5 to 0.05, the adjusted R^2 steadily improved, approaching 1, indicating a stronger fit; this suggests that our VATA prediction effectively captures the trend in the comfort data. Compared to the smoothing factor of 0.05, α equal to 0.3 can be a more balanced smoothing factor in VATA evaluation practice, which considers both continuous trends and actual measurement values, with an adjusted R^2 of 0.7541. These results confirm the validity of VATA as a significant predictor of in-field OTC variations, though additional factors like physiological parameters and sunlight intensity also influence thermal comfort.

4.2.2. VATA-assisted comfort modelling vs. IF-assisted comfort modelling

We analysed the 43-point OTC field survey comfort data using HSNA parameters (heart rate, solar intensity, noise level, altitude). The performance of VATA+HSNA, IF+HSNA, and individual heart rate and solar intensity models is summarised in Figure 9. The VATA+HSNA model outperformed all others, achieving an adjusted R^2 of 0.8893, surpassing IF+HSNA (adjusted $R^2 = 0.7784$). This highlights VATA's superior ability to model thermal comfort compared to IF-based models. One of the potential reasons for VATA's improved performance can be its ability to incorporate human visual-perceptual evaluations to construct the perception logic chain IF-VPI-VATA-OTC, which more accurately capture comfort perceptions.

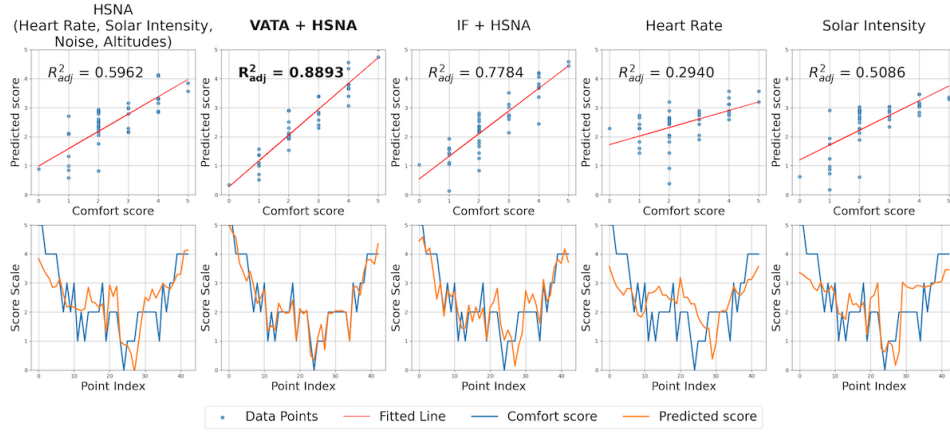


Figure 9: Multivariate comfort modelling performance based on HSNA, VATA, and IF data ($p < 0.001$).

HSNA alone explained 59.62% of the variance in comfort data, with heart rate and solar intensity models achieving adjusted R^2 values of 0.2940 and 0.5086, respectively. While VATA is a strong explanatory variable for thermal comfort, accurate prediction of OTC requires its integration with HSNA and additional environmental factors.

4.3. VATA mapping and spatial analysis

The trained VATA prediction model assigns a VATA value to each of the 92,233 SVIs and aggregates them as average scores within hexagonal spatial units (Figure 10). Using the H3 geospatial indexing system at resolution level 9, each hexagonal unit covers approximately 0.1 square kilometres, balancing precision and scalability for urban planning-orientated geospatial analytics. The VATA framework offers a valuable tool for enhancing urban quality of life by informing sustainable streetscape planning and design. It provides a comprehensive analysis of thermal affordance at the urban scale. By calculating and visually representing the average VATA values for each hexagonal spatial cell, the model effectively communicates the visual evaluation of thermal affordance across urban streetscapes in Singapore. Hexagonal spatial cells are colour coded to reflect varying levels of thermal affordance: red indicates high VATA values ($3.24 < \text{VATA} \leq 5$), suggesting superior thermal conditions due to shading and vegetation; blue represents low values ($0 \leq \text{VATA} \leq 1.76$), indicating poor conditions from sun exposure and limited greenery; grey signifies moderate thermal affordance ($1.76 < \text{VATA} \leq 3.24$).

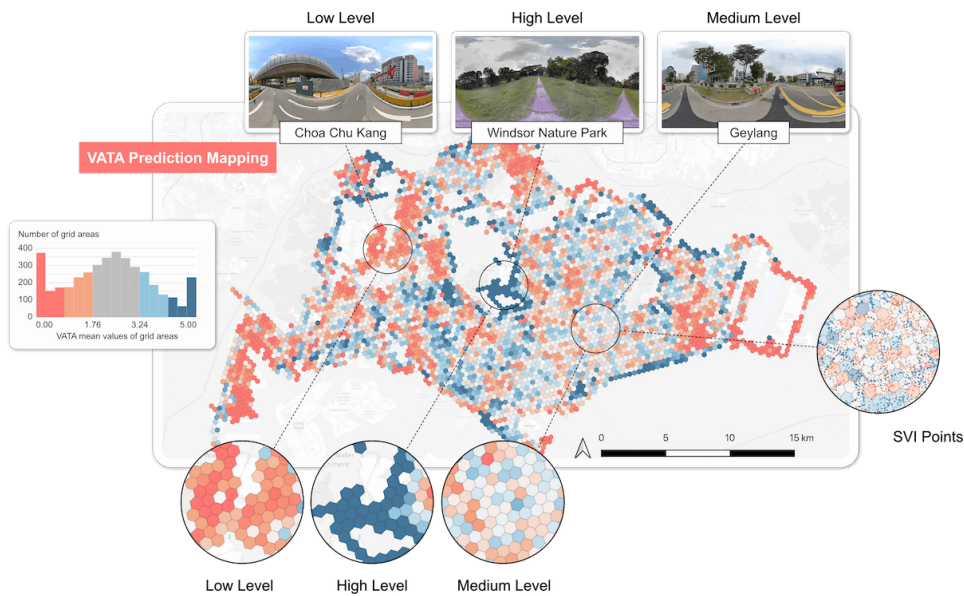


Figure 10: VATA prediction result mapping for Singapore streetscapes.

The VATA mapping presented in Figure 10 effectively visualises thermal affordance across urban streetscapes. Low VATA areas, such as Choa Chu Kang, lack continuous walking paths, greenery, and shaded spaces, resulting in increased sun exposure and reduced comfort. In contrast, high VATA areas like Windsor Nature Park feature well-defined pedestrian paths and abundant vegetation, offering better thermal affordance. Medium-rated areas such as Geylang neighbourhood provide moderate comfort. Despite continuous pedestrian paths and shaded areas, proximity to high-traffic zones or dense housing areas introduces additional artificial heat sources that diminish overall thermal affordance.

Figures 11 and 12 illustrate the spatial distribution of high and low VATA values, respectively. High-VATA areas, particularly city parks such as Garden of Peace, Windsor Nature Park, and Bukit Timah Nature Reserve, offer superior thermal comfort due to abundant green spaces, highlighting their importance for urban planning and public health. The exceptional VATA performance of the area around East Coast Park further underscores the cooling effects of water bodies and greenery in urban heat mitigation. In contrast, Figure 12 reveals low-VATA regions in the northwestern and eastern parts of Singapore, presenting opportunities for improving streetscape thermal affordance. Thus, the VATA framework serves as a valuable tool for urban planners and designers to identify and improve thermal affordance levels across different parts of a city, aiding the development of sustainable and

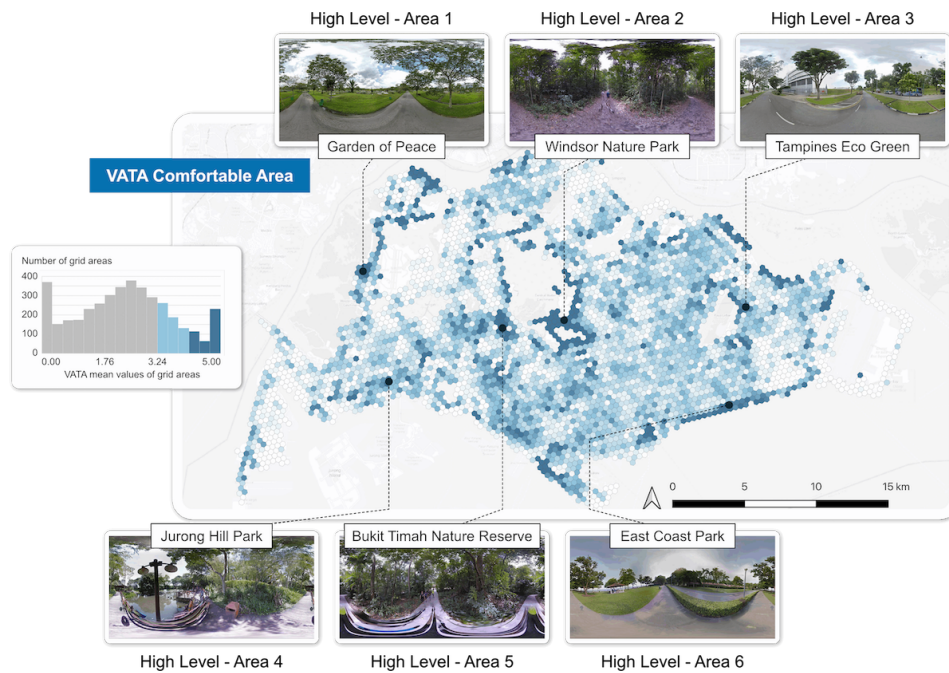


Figure 11: High level VATA area: Spatial distribution of VATA-comfortable streetscape area

liveable environments.

4.4. IF-VPI-VATA two-stage inference modelling results

After establishing the VATA prediction model using MTNNL, we developed an ENRM inference model based on 500 SVI-based online survey results to better understand the IF-VPI-VATA inference relationship, which helps future sustainable streetscape planning and design. Similarly to MTNNL, our ENRM-based inference modelling also employs a two-stage process: first, 52 IFs model 19 VPIs; second, the superposition of these two classes of features model VATA. The advantage of ENRM is that it efficiently addresses collinearity and identifies significant factors.

Table 4 presents the results for IF-VATA, VPI-VATA, and two-stage IF-VPI-VATA model. The two-stage model achieved the best performance with an adjusted R^2 of 0.7443, comparable to the MTNNL-based prediction model (adjusted $R^2 = 0.7316$). The feature weights derived from the two-stage IF-VPI-VATA model highlight the positive and negative factors influencing thermal affordance. A detailed discussion of these factors, including positive features (e.g., vegetation, shading area, human scale) and negative features (e.g., traffic elements, complexity, sunlight intensity), as well as design insights drawn from the inference modelling,

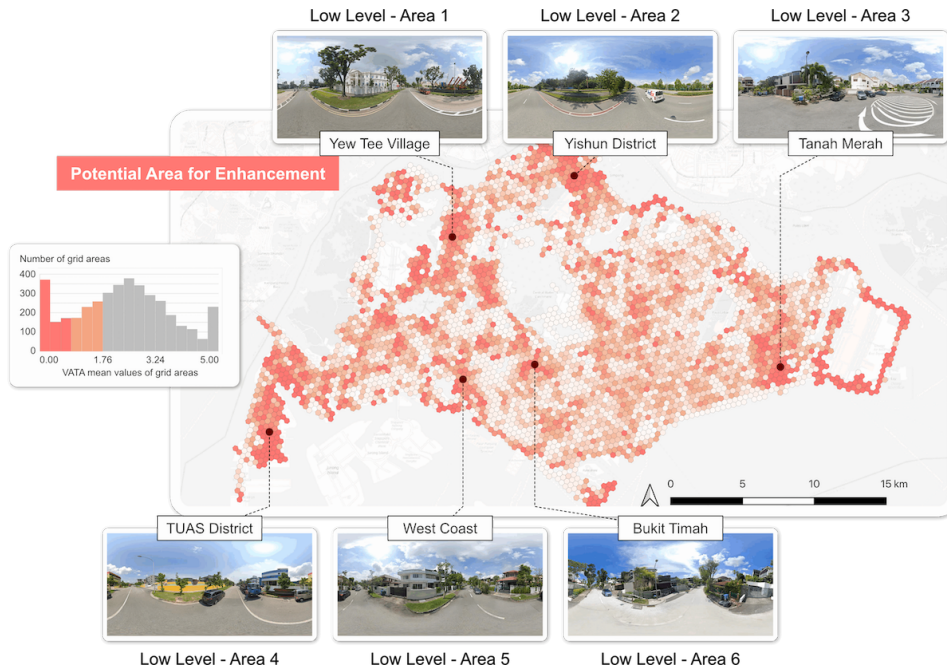


Figure 12: Low level VATA area: Spatial distribution of potential area for streetscape enhancement

is provided in Appendix G.

5. Discussion

5.1. VATA framework on sustainable streetscape planning and design

Our proposed VATA framework offers a novel approach for evaluating the thermal affordance of urban streetscapes at both urban-scale coverage and street-level resolution. Utilising SVIs across Singapore, computer vision algorithms to extract IF, and an online SVI visual assessment survey to investigate VPI, VATA supports sustainable streetscape design. The high-resolution urban-scale VATA maps, created after VATA prediction modelling and derived from averaging SVI scores within hexagonal spatial units, are instrumental in identifying urban areas with lower thermal affordance, helping urban planners prioritise regions for streetscape improvements and enhance sustainability. This targeted approach facilitates effective thermal affordance evaluations and informed urban planning decisions.

Furthermore, the inference modelling in our VATA framework is particularly valuable for helping urban planners and designers understand objective streetscape compositions (through 52 IFs), individuals' subjective perceptions (through 19

Table 4: ENRM models for IF-VATA, VPI-VATA, and IF+VPI-VATA inferences.

VATA Model	SVI Image Feature (IF)	Visual-Perceptual Indicator (VPI)	Two-stage Model (IF + VPI)
Model Parameter			
α	0.0289	0.0049	0.0070
L1 Ratio	0.5	0.5	0.5
Intercept	2.4945	2.4945	2.4945
Performance			
R ²	0.6587	0.6918	0.7576
Adjusted R²	0.6353	0.6823	0.7443
MAE	0.2804	0.2633	0.2184
MSE	0.2832	0.2557	0.2011
RMSE	0.5322	0.5057	0.4485
MAPE(%)	24.9150	23.4235	18.9456
AIC	-564.8017	-649.8209	-747.8569
BIC	-425.7196	-582.3871	-634.0625
Validation			
Cross-Validation Score	0.5000	0.4114	0.3816
Mean VIF	2.4843	2.4923	2.0383
Model Complexity			
Significant Coefficients	32	15	26
Total Features	52	19	71

VPIs), and their combined impact on thermal affordance. By analysing the inference relationship IF-VPI-VATA, urban planners and designers can strategically and organically design and arrange streetscape elements that not only meet thermal affordance standards but also elevate the overall quality of sustainable urban streetscape planning and design.

Figure 13 illustrates a continuous thermal affordance monitoring process and the corresponding sustainable streetscape planning and design supported by the evolving VATA framework. In each monitoring cycle, quarterly or monthly SVI data is collected, analysed using the VATA model, and further applied to develop the design guidelines to inform sustainable streetscape planning and design. This process updates the SVI dataset and VATA model for the next cycle, promoting a long-term sustainable and environmentally friendly streetscape planning and design paradigm in terms of thermal affordance.

The potential applications of VATA extend further. For residents, VATA maps are instrumental in improving street-level thermal affordance navigation. For urban planners and researchers, the open source nature of the VATA framework encour-

ages continuous improvement, integrating more diverse metrics for comprehensive urban thermal affordance analysis.

Importantly, while our current research focuses on the hot tropical climate of Singapore, the VATA framework and its underlying workflow can be methodologically transferred to other cities. To achieve reliable outcomes under different climatic and urban conditions, it is essential to conduct new local SVI visual assessment surveys and incorporate locally relevant geospatial data. Moreover, performing on-site thermal comfort validations in the target cities would be a crucial step to fine-tune the model’s parameters, ensuring the VATA approach remains robust and accurate. By doing so, the VATA workflow can guide sustainable streetscape planning and design across diverse environments while maintaining methodological consistency, adaptability, and scientific rigour.

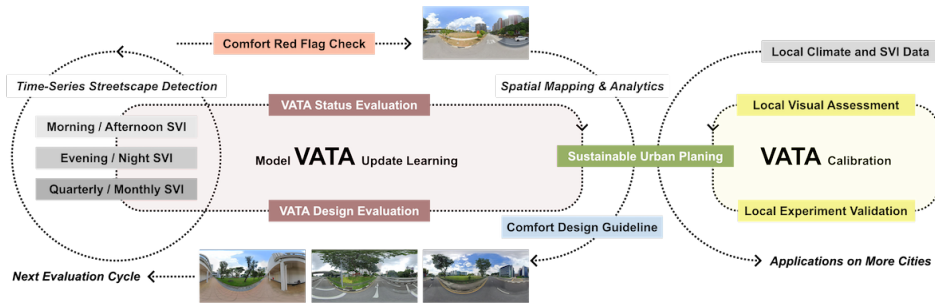


Figure 13: Long-term sustainable streetscape planning and design based on continuous thermal affordance evaluation supported by VATA framework.

5.2. Limitations and future work

The primary objective of this research is to leverage visual features in SVI and human assessment for a cost-effective, scalable approach to thermal affordance evaluation. However, certain limitations suggest opportunities for future study. First, temporal variance in SVI was not considered, omitting the possible impacts of time, weather conditions, and seasons on thermal affordance assessment. Future research could address this by sampling street views from the same locations at various times. Second, the validation was conducted along a single continuous path with 43 investigation points, which limits the diversity of environmental conditions captured. Expanding to a broader selection of field locations and incorporating microclimatic parameters (e.g. humidity, air temperature, and wind speed) for the validation would enhance both model robustness and generalizability.

Additionally, while visual cues provide valuable insights, they represent only one facet of urban conditions. Future research might incorporate complementary

data sources such as satellite imagery, remote sensing methods, or GIS-based analyses (e.g. land surface temperature or vegetation indices) to achieve a more comprehensive understanding of thermal affordance. Online surveys also pose limitations due to potential biases from participants' varied backgrounds and limited professional expertise; however, they remain a practical means of data collection within manageable constraints.

Finally, testing and validating this approach in multiple urban contexts, possibly spanning diverse climatic zones and geospatial conditions, would refine the framework and bolster its applicability. Through these efforts, future work can offer a more robust, scalable, and universally relevant tool for assessing and improving thermal comfort in urban environments worldwide.

6. Conclusions and Contributions

This research introduces the concept of thermal affordance and develops the visual assessment of thermal affordance – VATA method to evaluate built environments' capability to enhance thermal comfort. Key conclusions and contributions include:

1. Introducing thermal affordance, which describes the built environment's hidden information and inherent capability to promote thermal comfort. This new concept represents objective and fixed environmental influences on thermal comfort perception.
2. Pioneering the VATA framework, which integrates objective image features (IF) derived from street view imagery (SVI) with subjective visual-perceptual indicators (VPI) from survey data to build a two-stage neural network model for SVI-based cost-effective urban-scale thermal affordance assessment.
3. Exploring the IF–VPI–VATA relationship through inference modelling, contributing to the understanding of how visual compositions and human perceptions jointly shape thermal affordance, informing future sustainable streetscape planning and design to improve thermal comfort perception.
4. Generating a spatially detailed urban-scale mapping of thermal affordance across Singapore's streetscape, which is validated by in-field thermal comfort data and transferable to other cities, prioritising regions for future sustainable streetscape renovation.
5. Demonstrating that the validity of the VATA framework is supported by in-field comfort data. Comfort modelling based on VATA metrics outperforms models relying solely on IF metrics, highlighting the key role of VPI in the two-stage modelling process.

Author Contributions

Sijie Yang: Conceptualisation, Methodology, Investigation, Formal analysis, Software, Validation, Visualisation, Writing - Original Draft. **Adrian Chong:** Conceptualisation, Methodology, Writing - Review & Editing. **Pengyuan Liu:** Data Curation, Resources, Methodology, Writing - Review & Editing. **Filip Biljecki:** Conceptualisation, Methodology, Resources, Writing - Review & Editing, Supervision, Project administration, Funding acquisition.

Data and Code Availability

The computed VATA data for thermal affordance in Singapore can be accessed at <https://thermal-affordance.ual.sg/>. Data for additional cities will be updated regularly. The code and calculation methods will be open-sourced at <https://github.com/Sijie-Yang/Thermal-Affordance>.

Acknowledgements

We gratefully acknowledge the participants of the experiment. We thank our colleagues at the NUS Urban Analytics Lab for the discussions. The research involves human participants. This experiment has been approved by the Ethical Review Committee (ERC) of the NUS College of Design and Engineering on 6 February 2024 (reference code 2024/DOA/002). This research is supported by NUS Research Scholarship (NUSGS-CDE DO IS AY22&L GRSUR0600042). This research is part of the project 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. This research is part of the project 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

We utilised Grammarly and Writefull for language refinement and grammar checks during the writing process. While the tools helped ensure language accuracy and clarity, all scientific insights, conclusions, and content were generated solely by the authors. The final manuscript was thoroughly reviewed and edited by the authors to maintain accuracy and integrity.

Appendix A. Introduction on TrueSkill algorithm in this study

The TrueSkill algorithm, developed by Microsoft Research, is used in this study to rank street view images (SVIs) based on pairwise comparisons made by survey participants. Each image's performance in a comparison is modelled as a normal distribution, where the observed performance p_i is the sum of its true skill level s_i and some random noise ϵ_i . This relationship can be expressed as follows:

$$p_i = s_i + \epsilon_i \quad \text{where} \quad \epsilon_i \sim \mathcal{N}(0, \beta^2)$$

Here, ϵ_i is normally distributed with a mean of zero and variance β^2 , representing uncertainties in human perception during the comparison process.

After each comparison, the skill estimates of the two images involved are updated using Bayesian inference. If image i is rated higher than image j , their skill levels are adjusted based on the following formulas:

$$\mu'_i = \mu_i + \frac{\sigma_i^2}{c} \cdot \Delta, \quad \sigma_i'^2 = \sigma_i^2 \left(1 - \frac{\sigma_i^2}{c^2} \right)$$

Here, μ'_i and $\sigma_i'^2$ are the updated mean and variance for the image i , and Δ represents the performance difference between the two images. The term c is a normalisation factor that accounts for the uncertainty in both images' skill estimates.

TrueSkill operates under the assumption that all images' performances follow a normal distribution, which allows for an iterative refinement of the skill estimates through multiple pairwise comparisons. As more comparisons are made, the algorithm becomes more confident in the ranking, reflected in reduced variance for each image.

Once all comparisons are completed, the raw TrueSkill scores are normalised to a 0-5 scale to ensure consistency across VATA and the 19 VPIs. The normalisation process is performed using the formula:

$$S_i = \frac{s_i - \min(s)}{\max(s) - \min(s)} \times 5$$

where S_i is the normalised score for image i , and s_i is the raw skill score. This normalisation ensures that the scores are comparable across different indicators and suitable for further analysis, such as machine learning regression tasks.

Appendix B. Result of the online SVI visual assessment survey

The results of the SVI visual assessment survey (Figure B.14) reveal clear discrepancies in the scoring across all indicators, including VATA and the other 19 VPIs, when comparing the five streetscape clusters. Notable examples include indicators such as humidity inference, greenery rate, and imageability. Streetscapes rich in vegetation tend to receive higher average scores for humidity inference and greenery rate, whereas streetscapes dominated by buildings score higher on average for imageability. On the other hand, some indicators, like sunshine intensity and traffic flow, show much smaller discrepancies across the different clusters.

The visualization of the online SVI visual assessment survey reveals discrepancies in scoring between VATA and 19 VPIs for five streetscape clusters.

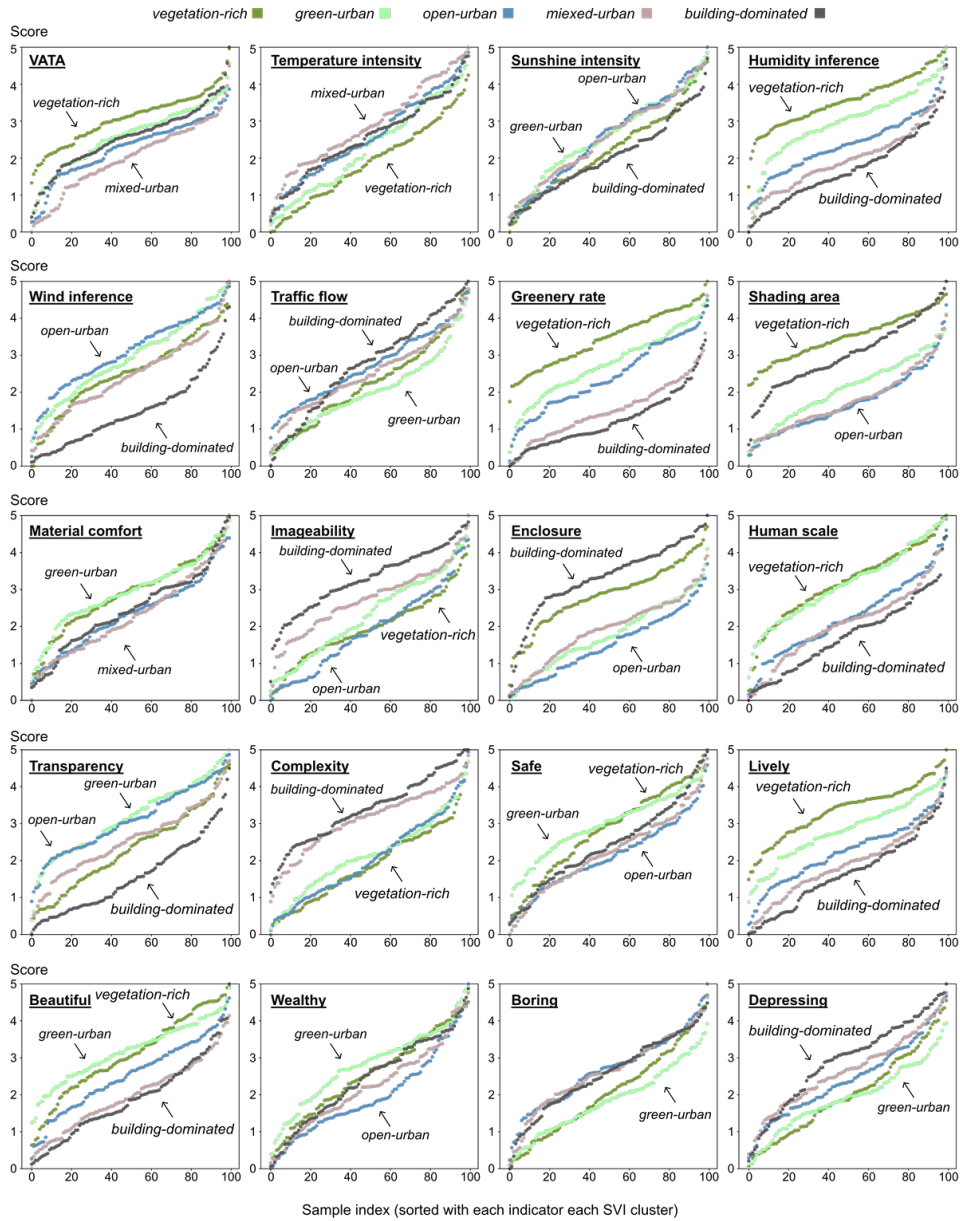


Figure B.14: Visualisation of online SVI visual assessment survey result.

Appendix C. Subfeatures of IF datasets in VATA modelling

Table C.5: Main sub-features applied for modelling and analytics.

Visual Feature Type	Sub-feature	Definition	
Semantic Segmentation P_i			
Built Env Element	P_{i-B}	B_1 building	The proportion of building structures like houses and offices.
		B_2 wall	The proportion of wall, vertical barriers made of concrete.
		B_3 fence	The proportion of fences, barriers made of wood or metal.
Transport Element	P_{i-T}	T_1 road	The proportion of roads, paved pathways for vehicular traffic.
		T_2 traffic_infra	The proportion of poles, traffic lights, and traffic signs.
		T_3 vehicle	The proportion of cars, trucks, buses, motorcycles, and trains.
Nature Element	P_{i-N}	N_1 sky	The proportion of sky in the image.
		N_2 vegetation	The proportion of plants, trees, and bushes.
		N_3 terrain	The proportion of terrain, ground features like soil or rocks.
Human-scale Element	P_{i-H}	H_1 person	The proportion of people walking or standing.
		H_2 sidewalk	The proportion of pathways for pedestrians.
		H_3 bicycle_rider	The proportion of bicycles and riders.
Object Detection A_i			
Infrastructure Object	A_{i-I}	I_1 traffic_light	Number of traffic lights detected in the image.
		I_2 stop_sign	Number of stop signs detected in the image.
		I_3 fire_hydrant	Number of fire hydrants detected in the image.
Vehicular Object	A_{i-V}	V_1 car	Number of cars detected in the image.
		V_2 bus	Number of buses detected in the image.
		V_3 truck	Number of trucks detected in the image.
		V_4 motorcycle	Number of motorcycles detected in the image.
Human-scale Object	A_{i-H}	H_1 person	Number of persons detected in the image.
		H_2 bench	Number of benches detected in the image.
		H_3 bicycle_rider	Number of bicycles and riders detected in the image.
Pixel Feature C_i			
Pixel Detail Complexity	C_{i-D}	D_1 sharpness	A measure of the texture clarity of the image.
		D_2 entropy	A measure of the texture randomness within the image.
		D_3 canny_edge	The ratio of edge pixels to overall pixels in the image.
Pixel Aesthetic Vibrancy	C_{i-A}	A_1 colorfulness	An indicator of the richness and variety of colors in the image.
		A_2 hue_std	The standard deviation of the hue values in the image.
		A_3 saturation_std	The standard deviation of the saturation values in the image.
		A_4 lightness_std	The standard deviation of the lightness values in the image.
		A_5 contrast	The intensity difference between the light and dark areas.
		A_6 image_variance	The variance of the pixel values in the grayscale image.
Pixel Average Perception	C_{i-P}	P_1 hue	The average hue value of the pixels in the image.
		P_2 saturation	The average saturation value of the pixels in the image.
		P_3 lightness	The average lightness value of the pixels in the image.
Scene Recognition S_i			
Commercial Scene	S_{i-B}	B_1 downtown	Probability of the image being central business district.
		B_2 office_building	Probability of the image being building used for business or work.
Residential Scene	S_{i-R}	R_1 apartment	Probability of the image being outdoor view of apartments.
		R_2 neighborhood	Probability of the image being area with houses and apartments.
		R_3 food_court	Probability of the image being dining area with multiple eateries.
Transportation Scene	S_{i-T}	T_1 parking_lot	Probability of the image being area for parking vehicles.
		T_2 driveway	Probability of the image being private road leading to a house
		T_3 highway	Probability of the image being main road for fast travel.
Public Space Scene	S_{i-P}	P_1 plaza	Probability of the image being open public square.
		P_2 market/outdoor	Probability of the image being outdoor market for shopping.
		P_3 campus	Probability of the image being grounds of a university or college.
		P_4 promenade	Probability of the image being public area for walking.
Nature and Park Scene	S_{i-N}	N_1 field/wild	Probability of the image being natural open land.
		N_2 forest_path	Probability of the image being path through a forest.
		N_3 forest_leaf	Probability of the image being forest with broadleaf trees.
		N_4 park	Probability of the image being public green area.
Construction Scene	S_{i-C}	C_1 construction	Probability of the image being where buildings are being built.
		C_2 industrial_area	Probability of the image being industrial factory area.

Appendix D. Rationale explanation of the survey and training sample sizes

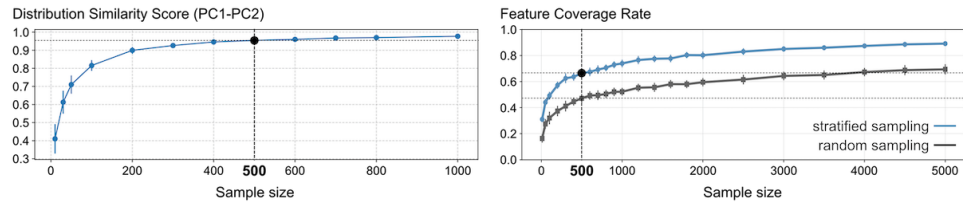
Given the constraints on resources and participant availability for the visual evaluation, it was essential to carefully determine an appropriate sample size. An excessively large number of images would risk incomplete evaluations, leading to inaccurate results, while too small a sample would fail to comprehensively represent the diversity of street scenes. Thus, a balanced approach was needed to ensure both feasibility and representativeness.

Our solution was to employ stratified sampling, incorporating equal representation from five street scene categories identified through a k -means clustering algorithm. From each category, 100 images were randomly selected, yielding a total of 500 images. We invited 176 participants to perform 18 pairwise comparisons for each of the 19 VPI and VATA indicators. On average, each indicator was compared 6.34 times per image, resulting in 3168 pairwise comparison results per indicator. The methodology is detailed in Section 3.2, with relevant visuals provided in Figures 4 and 5.

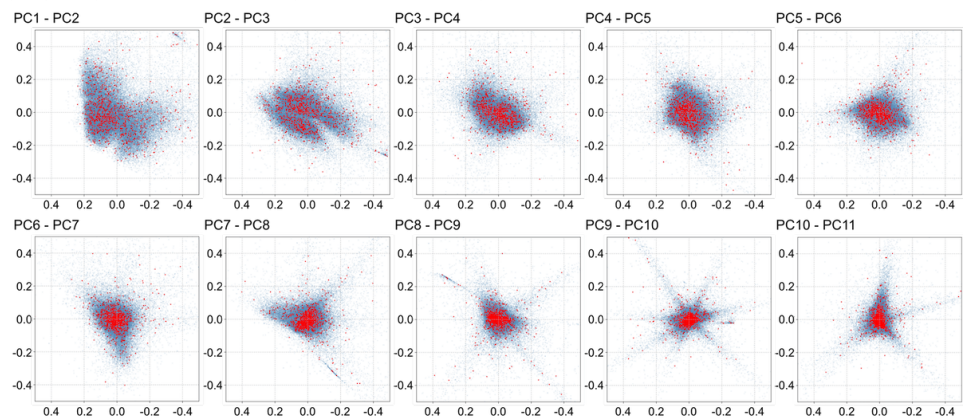
A sample size of 500 images successfully captures the diversity of street scene features, as demonstrated in Figure D.15. Subfigure (a) highlights that a sample of 500 images captures the overall distribution of street scenes across the first two principal components (PC1 and PC2) with a similarity of 0.95, ensuring comprehensive representation. Beyond this size, the marginal benefit of additional samples diminishes significantly. Stratified sampling further improved coverage of the feature space, increasing it from 0.47 to 0.67. This performance matches what would typically require random sampling of approximately 4,000 images. Subfigure (b) visualizes this improvement, showing that the 500 sampled images (red dots) uniformly represent the overall distribution (blue dots) across the first 11 principal components.

Figure D.15 Subfigures (c) and (d) demonstrate that a training sample size of 300 images is sufficient to estimate the average VATA score within a 95% confidence interval (CI), with an error margin of only 5%, while 200 images from the survey sample are allocated for use in the validation and test sets. This estimation assumes an ideal population mean of 2.5 (based on the normal distribution assumption of the TrueSkill algorithm) and a data-augmented population standard deviation of 1. Additionally, increasing the sample size does not result in a significant reduction in CI width, as shown by the graph illustrating the second derivative of the CI width. Moreover, as the sample size increases from 300 to 500, the distribution of VATA scores remains largely unchanged (K-S statistic < 0.05), further suggesting that a representative and convergent sample of 300 images is convergent and sufficient for training a VATA prediction model.

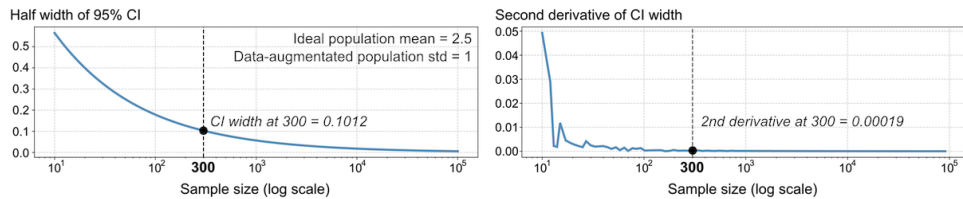
a Survey image set with 500 samples has a high degree of similarity with the overall distribution in the PC1-PC2 principal component space, with up to 95% similarity. The stratified sampling design significantly enhances the coverage of small samples in the overall feature interval (from 47% to 67%).



b The survey image set with 500 samples can uniformly capture the overall features in the first 11 principal components. It showcases the comprehensiveness of feature capture.



c Training set with 300 samples reaches 95% confidence interval (CI) to reach the ideal population mean of VATA score within 5% error. And there is no significant CI change if we increase the sample size.



d There is no significant change on VATA score distribution if we add the sample size from 300 images to 500 images (relative error = 4.06% < 5%, K-S statistic = 0.0273 < 0.05).

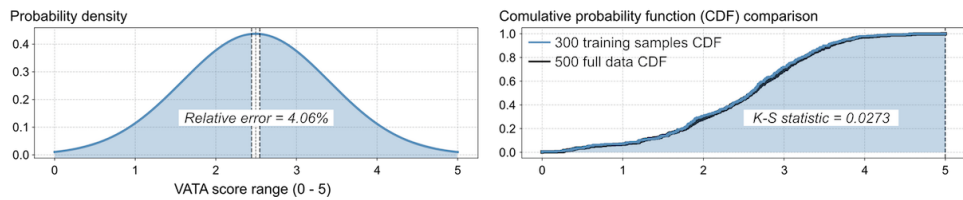
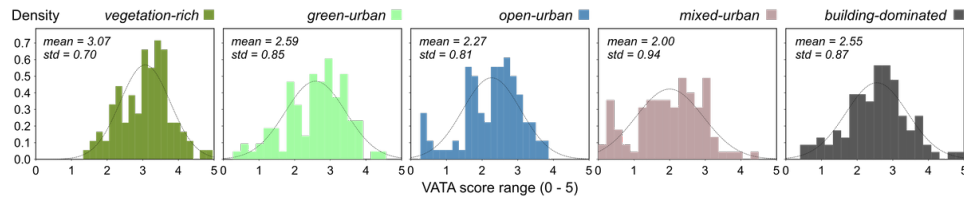


Figure D.15: Rationale analytics of survey sample size (500) and training sample size (300)

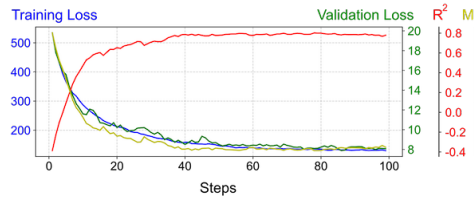
Appendix E. Details on MTNNL model training data and process

Figures E.16.a and E.16.c demonstrate that different streetscape categories can influence their VATA scores obtained from the online survey. During the division of train/validation/test sets, it is ensured that an equal quantity of each of the five streetscape categories is included in every fold. Figure E.16.b illustrates the training process for fold 1 in k-fold cross-validation model training. Both validation loss and train loss consistently decrease and stabilise, indicating that no overfitting occurred during training.

a VATA score distributions differ on the five streetscape clusters, with cluster 1, vegetation-rich, having the highest mean value of 3.07, while cluster 4, mixed-urban, the lowest mean value 2.00.



b The decrease in training loss matched the decrease in validation loss, MAE, and the increase in R2 without overfitting (Fold 1).



c Division of train / validation / test set in K-fold cross-validation averagely covers each cluster of streetscapes.

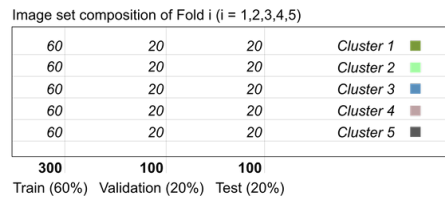


Figure E.16: Train-validation-test set division in K-folds and training process metrics.

Appendix F. Details on stage-1 ENRM models

From the stage-1 ENRM models shown in Figure F.6, we can observe that the R^2 values for IF-VPI inference modelling in the microclimate inference and environmental assessment sections—except for ‘material comfort’—exceed 0.6. This suggests that these VPIs are highly predictable and can be effectively modelled using IFs. Additionally, the VPIs ‘enclosure’, ‘transparency’, and ‘complexity’ in the design quality section demonstrate strong consistency, while in the evoked emotion section, only the R^2 for ‘lively’ exceeds 0.6. This indicates that most VPIs related to evoked emotions vary significantly between individuals. It is crucial to consider both IF and VPI dimensions in VATA evaluation, as visual-perceptual indicators capture human visual-perceptual interpretations that are not fully reflected in SVI image features.

Table F.6: Stage-1 ENRM models for IF-VPI inference

Perception Model	adjusted R^2	MAE	MSE	Significant Coefficients (out of 52 IFs)
Microclimate Inference				
temp_intensity	0.6011	0.4091	0.5279	24
sun_intensity	0.6179	0.3852	0.4887	24
humidity_inference	0.6788	0.2983	0.3571	38
wind_inference	0.6417	0.3557	0.4610	27
Environment Assessment				
traffic_flow	0.6571	0.3289	0.4020	37
greenery_rate	0.7994	0.2263	0.2573	29
shading_area	0.8943	0.1017	0.1123	43
material_comfort	0.4069	0.4576	0.6333	29
Design Quality				
imageability	0.5105	0.4379	0.5660	37
enclosure	0.7927	0.2237	0.2515	26
human_scale	0.5209	0.4446	0.6198	25
transparency	0.6183	0.3843	0.4701	30
complexity	0.7375	0.2805	0.3275	41
Evoked Emotion				
safe	0.4235	0.4827	0.6568	32
lively	0.7102	0.2563	0.3125	32
beautiful	0.5598	0.3830	0.5093	34
wealthy	0.4072	0.5116	0.7259	35
boring	0.4378	0.4462	0.6072	35
depressing	0.4733	0.4526	0.6158	34

Appendix G. Feature weight analytics in the ENRM inference modelling

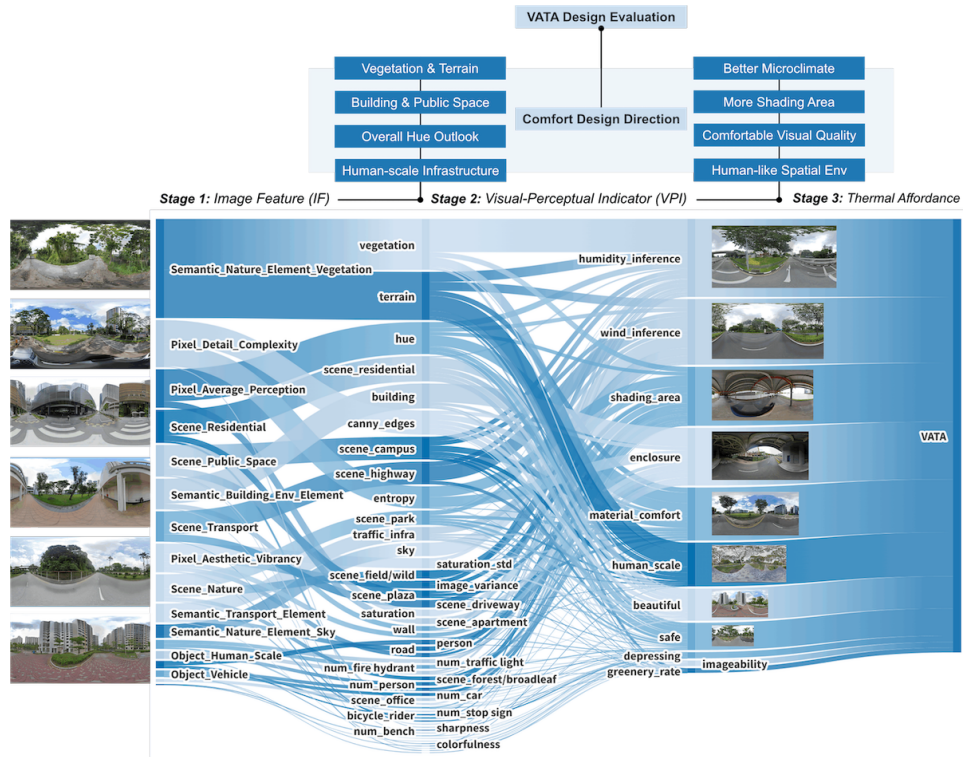


Figure G.17: Sankey diagram of positive feature weight proportion in the IF-VPI-VATA two-stage inference modelling.

The weights from the IF-VPI-VATA two-stage modelling are normalised and the proportional distributions of positive and negative weights are visualised as Sankey diagrams (Figure G.17 and Figure G.18). Figure G.17 highlights the proportional distribution of the weights of positive features, where IFs and VPIs with higher weights significantly enhance the VATA level of streetscapes. Conversely, Figure G.18 shows the proportional distribution of the weights of negative features, where IFs and VPIs with higher absolute weights more strongly reduce the VATA level of streetscapes.

From all positive features, semantic elements like nature (vegetation, terrain, sky), open urban scenes (residential, campus, park), and colourful pixel features (hue) dominate the IF. Key VPIs include microclimate inference (humidity, wind), environmental assessment (shading area, material comfort), design quality (enclosure, human scale), and emotional response (beautiful, safe). We summarise

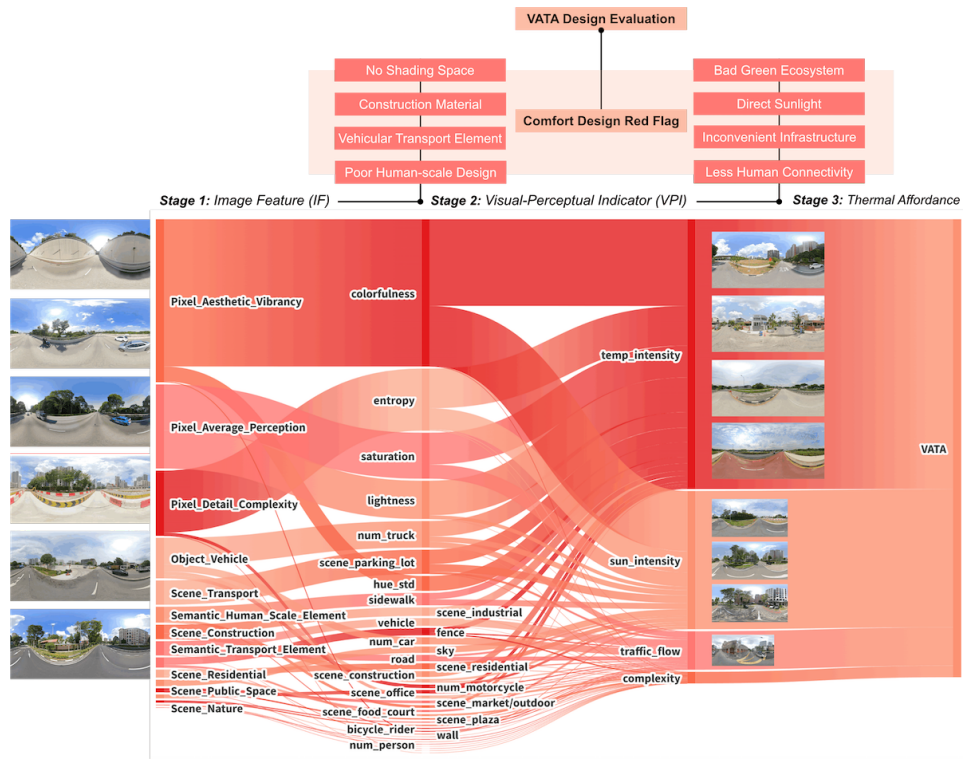


Figure G.18: Sankey diagram of negative feature weight proportion in the IF-VPI-VATA two-stage inference modelling

four important aspects of IF and VPI that are conducive to improving the thermal affordance level of streetscapes, as shown in Figure G.17: 1) higher vegetation and terrain proportions can bring a better street microclimate; 2) more buildings and public spaces can increase the proportion of shading area; 3) a better overall hue outlook of streetscapes can improve visual comfort; 4) more human-scale infrastructures can make streetscape spatial environments more human-friendly and comfortable.

From all negative features, pixel-level characteristics, traffic elements, and complex urban scenes dominate the IF, while essential VPIs include temperature intensity, sunlight intensity, traffic flow, and complexity. Four key aspects of IF and VPI that reduce thermal affordance (Figure G.18) are: 1) less shading area results in more direct sunlight and thus higher temperature; 2) a higher proportion of construction materials and complicated urban scenes (transport, construction) reduce the proportion of green ecosystems in streetscapes; 3) more vehicular transport elements imply fewer pedestrian connections; 4) lack of human-scale designs indicate

decreased infrastructure convenience.

In addition to the inference logic chains hidden in the IF-VPI-VATA two-stage modelling mentioned above, more design implications can be discovered and summarised on the basis of the VATA framework in the long-term process of sustainable streetscape planning and design.

Appendix H. Further analytics on feature-wise relationships

Appendix H.1. Correlation among VATA and visual-perceptual indicators (VPI)

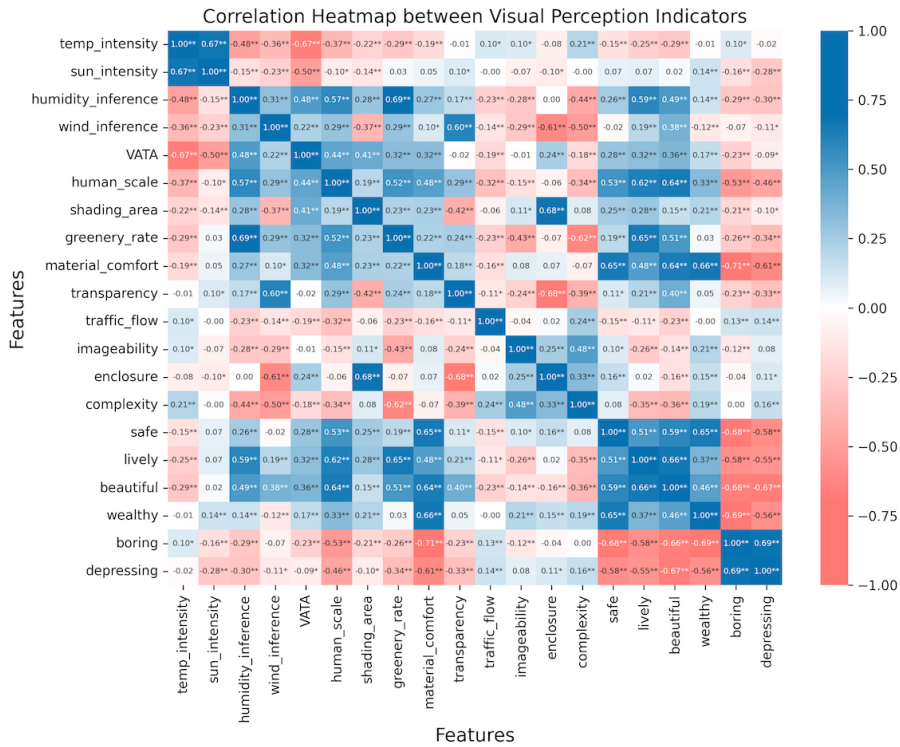


Figure H.19: Correlation coefficient among VATA and visual-perceptual indicators (VPI)

Figure H.19 shows the cross-correlation matrix between VATA and 19 VPIs. The results reveal strong positive relationships between VATA and several key VPIs, including humidity inference, wind inference, human scale, shading area, greenery rate, material comfort, and transparency. These correlations highlight the importance of green spaces and human-scale considerations in improving streetscape

thermal affordance. Conversely, traffic flow, imageability, enclosure, and complexity exhibit negative correlations with VATA, suggesting that these elements may detract from the overall thermal affordance of urban environments.

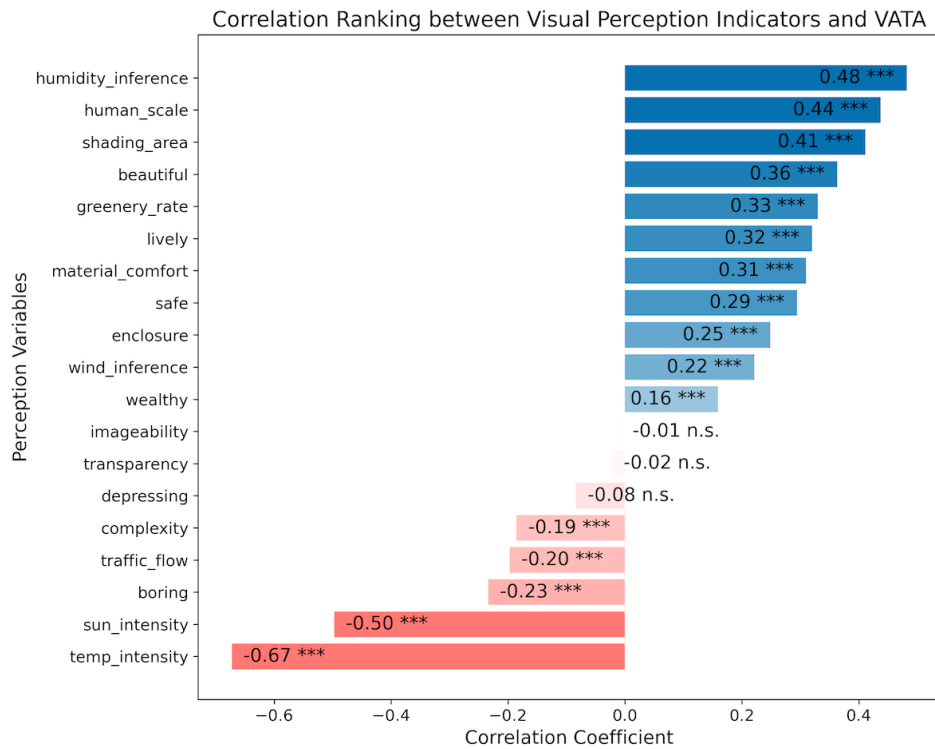


Figure H.20: Ranking of VPIs' correlation with VATA.

Figure H.20 ranks VPIs' correlations with VATA. Factors like humidity inference, human scale, shading area, 'beautiful' emotion, and greenery rate are strongly positively correlated with VATA, significantly contributing to thermal affordance. In contrast, temperature intensity, sunshine intensity, 'boring' emotion, traffic flow, and complexity are significantly negatively correlated with VATA, indicating their detrimental impact. Notably, imageability, transparency, and 'depressing' emotion show no significant correlations with VATA, suggesting that these aspects, while important in urban design, may not directly influence thermal affordance assessment. This insight invites further research into the nuanced role of urban form and aesthetics in shaping human thermal affordance perception.

Appendix H.2. Relationship between image features (IF) and visual-perceptual indicators (VPI)

The exploration of correlations between IFs and VPIs reveals nuanced interactions that significantly influence urban thermal affordance assessment. This study incorporates a comprehensive analysis of 52 IFs and 19 VPIs, offering a detailed cross-correlation evaluation. The methodology prioritises the six most relevant IFs for each VPI providing a ranked evaluation as depicted in Figure H.21.

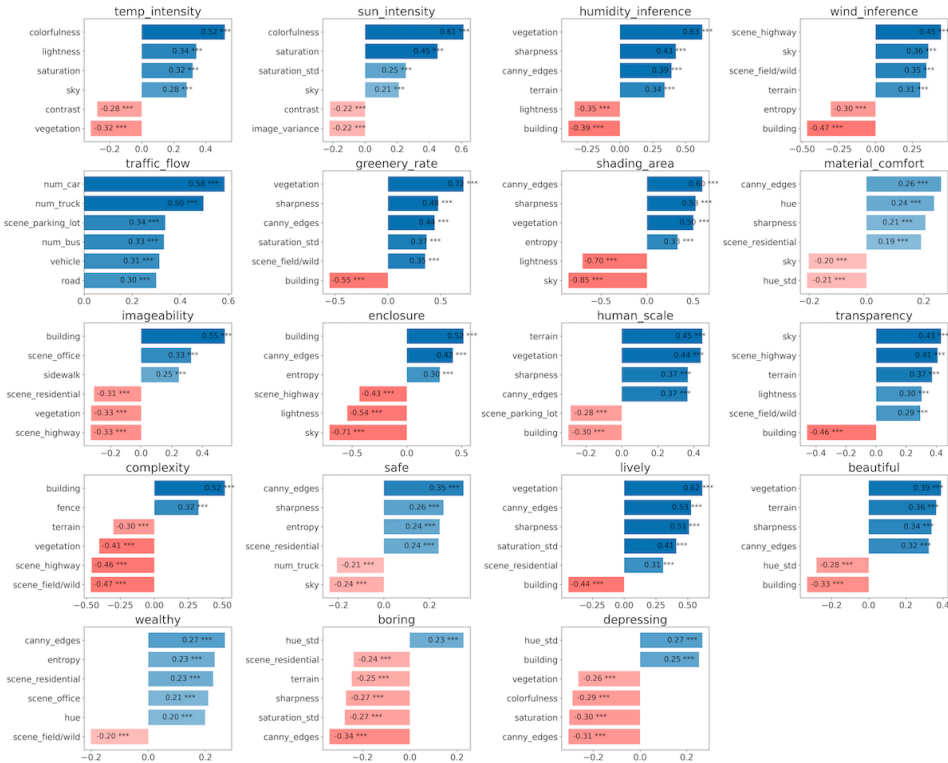


Figure H.21: Top six correlated IFs for each VPI.

Vegetation emerges as a pivotal element, showing strong positive correlations with humidity inference, greenery rate, shading area, and human scale. It significantly influences VATA, as shown in H.22. However, vegetation and terrain negatively correlate with temperature intensity, traffic flow, imageability, and complexity, highlighting the complex interplay between natural and built elements in urban settings and emphasising the critical role of vegetation and terrain in shaping thermally sustainable streetscapes. Buildings present a mixed influence on VATA and VPIs. They generally negatively impact VATA evaluation, correlating

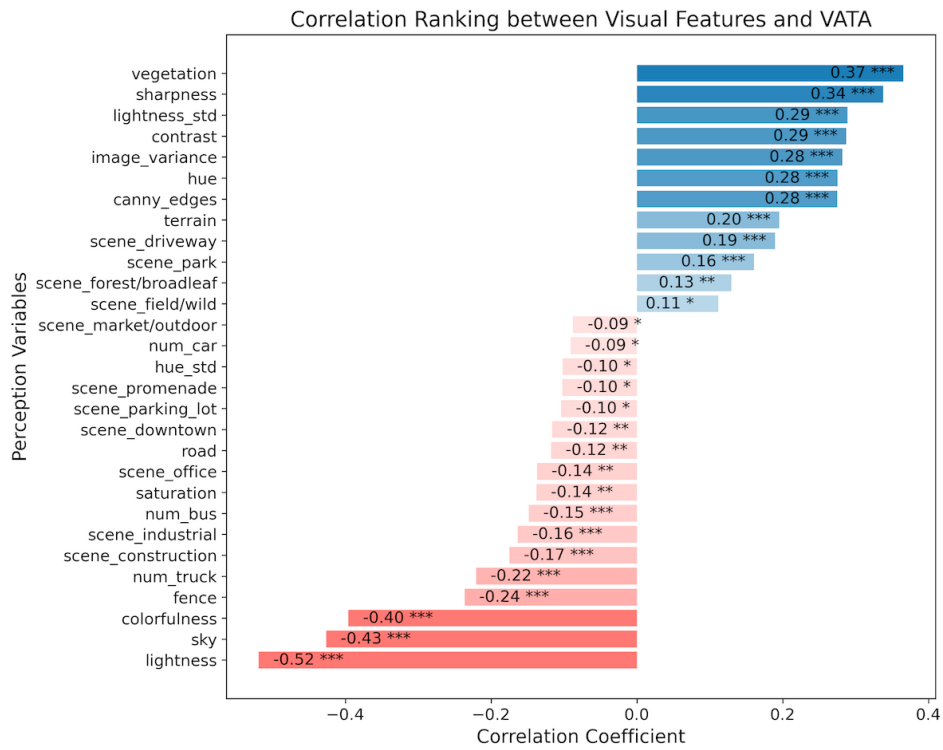


Figure H.22: Ranking of top twenty IFs' correlation with VATA.

positively with enclosure and complexity, and negatively with humidity inference, wind speed inference, greenery rate, and human scale. However, buildings contribute positively to the shading area, aligning with vegetation to enhance VATA.

This intricate web of correlations underscores the multifaceted nature of streetscape elements in determining visual assessment of urban thermal comfort. It also calls attention to the necessity of a balanced integration of natural and built components in urban planning to foster sustainable and comfortable living environments.

References

- [1] M. J. Alcoforado, H. Andrade, Global warming and the urban heat island, *Urban ecology: An international perspective on the interaction between humans and nature* (2008) 249–262.
- [2] G. Li, X. Zhang, P. A. Mirzaei, J. Zhang, Z. Zhao, Urban heat island effect of a typical valley city in china: Responds to the global warming and rapid urbanization, *Sustainable cities and society* 38 (2018) 736–745.
- [3] H. Takebayashi, H. Takebayashi, *Improvement measures of urban thermal environment*, Springer, 2015.
- [4] D. Lai, W. Liu, T. Gan, K. Liu, Q. Chen, A review of mitigating strategies to improve the thermal environment and thermal comfort in urban outdoor spaces, *Science of the Total Environment* 661 (2019) 337–353.
- [5] A. Tseliou, I. Koletsis, K. Pantavou, E. Thoma, S. Lykoudis, I. Tsiros, Evaluating the effects of different mitigation strategies on the warm thermal environment of an urban square in athens, greece, *Urban Climate* 44 (2022) 101217.
- [6] K. Chen, A. J. Newman, M. Huang, C. Coon, L. A. Darrow, M. J. Strickland, H. A. Holmes, Estimating heat-related exposures and urban heat island impacts: A case study for the 2012 chicago heatwave, *GeoHealth* 6 (1) (2022) e2021GH000535.
- [7] B. Steensen, L. Marelle, Ø. Hodnebrog, G. Myhre, Future urban heat island influence on precipitation, *Climate Dynamics* 58 (11-12) (2022) 3393–3403.
- [8] F. B. Errebai, D. Strebel, J. Carmeliet, D. Derome, Impact of urban heat island on cooling energy demand for residential building in montreal using meteorological simulations and weather station observations, *Energy and Buildings* 273 (2022) 112410.
- [9] J. Tan, Y. Zheng, X. Tang, C. Guo, L. Li, G. Song, X. Zhen, D. Yuan, A. J. Kalkstein, F. Li, et al., The urban heat island and its impact on heat waves and human health in shanghai, *International journal of biometeorology* 54 (2010) 75–84.
- [10] L. Chen, E. Ng, Outdoor thermal comfort and outdoor activities: A review of research in the past decade, *Cities* 29 (2) (2012) 118–125.

- [11] S. Shooshtarian, C. K. C. Lam, I. Kenawy, Outdoor thermal comfort assessment: A review on thermal comfort research in australia, *Building and Environment* 177 (2020) 106917.
- [12] N. da Silva Espinoza, C. A. C. dos Santos, M. B. L. de Oliveira, M. T. Silva, C. A. G. Santos, R. M. da Silva, M. Mishra, R. R. Ferreira, Assessment of urban heat islands and thermal discomfort in the amazonia biome in brazil: A case study of manaus city, *Building and Environment* 227 (2023) 109772.
- [13] R. S. Kovats, S. Hajat, Heat stress and public health: A critical review, *Annual Review of Public Health* 29 (1) (2008) 41–55, pMID: 18031221. arXiv:<https://doi.org/10.1146/annurev.publhealth.29.020907.090843>, doi:10.1146/annurev.publhealth.29.020907.090843. URL <https://doi.org/10.1146/annurev.publhealth.29.020907.090843>
- [14] A. Piracha, M. T. Chaudhary, Urban air pollution, urban heat island and human health: a review of the literature, *Sustainability* 14 (15) (2022) 9234.
- [15] A. R. Hwong, M. Wang, H. Khan, D. N. Chagwedera, A. Grzenda, B. Doty, T. Benton, J. Alpert, D. Clarke, W. M. Compton, Climate change and mental health research methods, gaps, and priorities: a scoping review, *The Lancet Planetary Health* 6 (3) (2022) e281–e291.
- [16] R. Aghamolaei, M. M. Azizi, B. Aminzadeh, J. O’Donnell, A comprehensive review of outdoor thermal comfort in urban areas: Effective parameters and approaches, *Energy & Environment* (2022) 0958305X221116176.
- [17] E. Johansson, S. Thorsson, R. Emmanuel, E. Krüger, Instruments and methods in outdoor thermal comfort studies—the need for standardization, *Urban climate* 10 (2014) 346–366.
- [18] G. Evola, V. Costanzo, C. Magrì, G. Margani, L. Marletta, E. Naboni, A novel comprehensive workflow for modelling outdoor thermal comfort and energy demand in urban canyons: Results and critical issues, *Energy and Buildings* 216 (2020) 109946.
- [19] K. Xiong, Z. Yang, B.-J. He, Spatiotemporal heterogeneity of street thermal environments and development of an optimised method to improve field measurement accuracy, *Urban Climate* 42 (2022) 101121.

- [20] X. Li, G. Wang, B. Zaitchik, A. Hsu, T. Chakraborty, Sensitivity and vulnerability to summer heat extremes in major cities of the united states, *Environmental Research Letters* 19 (9) (2024) 094039.
- [21] J. Spagnolo, R. De Dear, A field study of thermal comfort in outdoor and semi-outdoor environments in subtropical sydney australia, *Building and environment* 38 (5) (2003) 721–738.
- [22] T. Honjo, Thermal comfort in outdoor environment, *Global environmental research* 13 (2009) (2009) 43–47.
- [23] H. Mayer, P. Höppe, Thermal comfort of man in different urban environments, *Theoretical and applied climatology* 38 (1987) 43–49.
- [24] P. O. Fanger, et al., Thermal comfort. analysis and applications in environmental engineering., *Thermal comfort. Analysis and applications in environmental engineering.* (1970).
- [25] G. Jendritzky, R. de Dear, G. Havenith, UtcI—why another thermal index?, *International journal of biometeorology* 56 (2012) 421–428.
- [26] P. Bröde, E. L. Krüger, F. A. Rossi, D. Fiala, Predicting urban outdoor thermal comfort by the universal thermal climate index utci—a case study in southern brazil, *International journal of biometeorology* 56 (2012) 471–480.
- [27] J. A. Acero, K. Herranz-Pascual, A comparison of thermal comfort conditions in four urban spaces by means of measurements and modelling techniques, *Building and Environment* 93 (2015) 245–257.
- [28] P. A. Mirzaei, Recent challenges in modeling of urban heat island, *Sustainable cities and society* 19 (2015) 200–206.
- [29] M. M. Abdelrahman, A. Chong, C. Miller, Personal thermal comfort models using digital twins: Preference prediction with bim-extracted spatial-temporal proximity data from build2vec, *Building and Environment* 207 (2022) 108532.
- [30] Z. D. Tekler, Y. Lei, Y. Peng, C. Miller, A. Chong, A hybrid active learning framework for personal thermal comfort models, *Building and Environment* 234 (2023) 110148.
- [31] E. Ono, K. Mihara, K. P. Lam, A. Chong, The effects of a mismatch between thermal comfort modeling and hvac controls from an occupancy perspective, *Building and Environment* 220 (2022) 109255.

- [32] B. N. Silva, M. Khan, C. Jung, J. Seo, D. Muhammad, J. Han, Y. Yoon, K. Han, Urban planning and smart city decision management empowered by real-time data processing using big data analytics, *Sensors* 18 (9) (2018) 2994.
- [33] M. Batty, Digital twins in city planning, *Nature Computational Science* 4 (3) (2024) 192–199.
- [34] C. Gilge, Google street view and the image as experience, *GeoHumanities* 2 (2) (2016) 469–484.
- [35] F. Biljecki, K. Ito, Street view imagery in urban analytics and gis: A review, *Landscape and Urban Planning* 215 (2021) 104217.
- [36] W. Qiu, Z. Zhang, X. Liu, W. Li, X. Li, X. Xu, X. Huang, Subjective or objective measures of street environment, which are more effective in explaining housing prices?, *Landscape and Urban Planning* 221 (2022) 104358.
- [37] X. Liang, T. Zhao, F. Biljecki, Revealing spatio-temporal evolution of urban visual environments with street view imagery, *Landscape and Urban Planning* 237 (2023) 104802.
- [38] X. Liang, J. H. Chang, S. Gao, T. Zhao, F. Biljecki, Evaluating human perception of building exteriors using street view imagery, *Building and Environment* 263 (2024) 111875.
- [39] M. Zhao, Y. Zhang, Exploring the dose-response of landscape preference: A case study in singapore, *Applied Geography* 170 (2024) 103357.
- [40] J. Lee, S. Park, Current design guidelines' streetscape improvement for visual perception and walkability: A case study of sejong city, republic of korea, *Frontiers of Architectural Research* 12 (3) (2023) 423–443.
- [41] L. Cheng, S. Chu, W. Zong, S. Li, J. Wu, M. Li, Use of tencent street view imagery for visual perception of streets, *ISPRS International Journal of Geo-Information* 6 (9) (2017) 265.
- [42] Y. Kang, F. Zhang, S. Gao, H. Lin, Y. Liu, A review of urban physical environment sensing using street view imagery in public health studies, *Annals of GIS* 26 (3) (2020) 261–275.
- [43] A. Larkin, X. Gu, L. Chen, P. Hystad, Predicting perceptions of the built environment using gis, satellite and street view image approaches, *Landscape and urban planning* 216 (2021) 104257.

- [44] D. Ki, K. Park, Z. Chen, Bridging the gap between pedestrian and street views for human-centric environment measurement: A gis-based 3d virtual environment, *Landscape and Urban Planning* 240 (2023) 104873.
- [45] C. Hou, Y. Li, F. Zhang, Sensing urban physical environment with geoai and street-level imagery, in: *Handbook of Geospatial Approaches to Sustainable Cities*, CRC Press, 2024, pp. 3–30.
- [46] Z. Fan, F. Biljecki, Nighttime street view imagery: A new perspective for sensing urban lighting landscape, *Sustainable Cities and Society* 116 (2024) 105862.
- [47] N. Naik, J. Philipoom, R. Raskar, C. Hidalgo, Streetscore-predicting the perceived safety of one million streetscapes, in: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2014, pp. 779–785.
- [48] X. Xu, W. Qiu, W. Li, X. Liu, Z. Zhang, X. Li, D. Luo, Associations between street-view perceptions and housing prices: Subjective vs. objective measures using computer vision and machine learning techniques, *Remote Sensing* 14 (4) (2022) 891.
- [49] T. Zhao, X. Liang, W. Tu, Z. Huang, F. Biljecki, Sensing urban soundscapes from street view imagery, *Computers, Environment and Urban Systems* 99 (2023) 101915.
- [50] J. Xiao, L. Quan, Multiple view semantic segmentation for street view images, in: *2009 IEEE 12th international conference on computer vision*, IEEE, 2009, pp. 686–693.
- [51] S. Nagata, T. Nakaya, T. Hanibuchi, S. Amagasa, H. Kikuchi, S. Inoue, Objective scoring of streetscape walkability related to leisure walking: Statistical modeling approach with semantic segmentation of google street view images, *Health & Place* 66 (2020) 102428.
- [52] W. Yap, J.-H. Chang, F. Biljecki, Incorporating networks in semantic understanding of streetscapes: Contextualising active mobility decisions, *Environment and Planning B: Urban Analytics and City Science* 50 (6) (2023) 1416–1437.
- [53] X. Wei, F. Guan, X. Zhang, N. Van de Weghe, H. Huang, Integrating planar and vertical environmental features for modelling land surface temperature based on street view images and land cover data, *Building and Environment* 235 (2023) 110231.

- [54] K. Fujiwara, M. Khomiakov, W. Yap, M. Ignatius, F. Biljecki, Microclimate vision: Multimodal prediction of climatic parameters using street-level and satellite imagery, *Sustainable Cities and Society* 114 (2024) 105733.
- [55] Y. Ogawa, T. Oki, C. Zhao, Y. Sekimoto, C. Shimizu, Evaluating the subjective perceptions of streetscapes using street-view images, *Landscape and Urban Planning* 247 (2024) 105073.
- [56] K. Ito, Y. Kang, Y. Zhang, F. Zhang, F. Biljecki, Understanding urban perception with visual data: A systematic review, *Cities* 152 (2024) 105169.
- [57] H. Zhao, J. Shi, X. Qi, X. Wang, J. Jia, Pyramid scene parsing network, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2881–2890.
- [58] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.
- [59] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, B. Schiele, The cityscapes dataset for semantic urban scene understanding, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 3213–3223.
- [60] B. Zhou, H. Zhao, X. Puig, S. Fidler, A. Barriuso, A. Torralba, Scene parsing through ade20k dataset, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 633–641.
- [61] W. Yap, P. Janssen, F. Biljecki, Free and open source urbanism: Software for urban planning practice, *Computers, Environment and Urban Systems* 96 (2022) 101825.
- [62] W. Yap, R. Stouffs, F. Biljecki, Urbanity: automated modelling and analysis of multidimensional networks in cities, *npj Urban Sustainability* 3 (1) (2023) 45.
- [63] K. Ito, Y. Zhu, M. Abdelrahman, X. Liang, Z. Fan, Y. Hou, T. Zhao, R. Ma, K. Fujiwara, J. Ouyang, et al., Zensvi: An open-source software for the integrated acquisition, processing and analysis of street view imagery towards scalable urban science, *arXiv preprint arXiv:2412.18641* (2024).
- [64] R. Goel, L. M. Garcia, A. Goodman, R. Johnson, R. Aldred, M. Murugesan, S. Brage, K. Bhalla, J. Woodcock, Estimating city-level travel patterns using

street imagery: A case study of using google street view in britain, *PLoS one* 13 (5) (2018) e0196521.

- [65] J. Kang, M. Körner, Y. Wang, H. Taubenböck, X. X. Zhu, Building instance classification using street view images, *ISPRS journal of photogrammetry and remote sensing* 145 (2018) 44–59.
- [66] H. Xu, H. Sun, L. Wang, X. Yu, T. Li, Urban architectural style recognition and dataset construction method under deep learning of street view images: A case study of wuhan, *ISPRS International Journal of Geo-Information* 12 (7) (2023) 264.
- [67] B. Lei, P. Liu, N. Milojevic-Dupont, F. Biljecki, Predicting building characteristics at urban scale using graph neural networks and street-level context, *Computers, Environment and Urban Systems* 111 (2024) 102129.
- [68] Y. Su, Y. Zhong, Q. Zhu, J. Zhao, Urban scene understanding based on semantic and socioeconomic features: From high-resolution remote sensing imagery to multi-source geographic datasets, *ISPRS Journal of Photogrammetry and Remote Sensing* 179 (2021) 50–65.
- [69] M. R. Ibrahim, J. Haworth, T. Cheng, WeatherNet: Recognising weather and visual conditions from street-level images using deep residual learning, *ISPRS International Journal of Geo-Information* 8 (12) (2019) 549.
- [70] S. Yang, K. Krenz, W. Qiu, W. Li, The role of subjective perceptions and objective measurements of the urban environment in explaining house prices in greater london: A multi-scale urban morphology analysis, *ISPRS International Journal of Geo-Information* 12 (6) (2023) 249.
- [71] R. Ewing, S. Handy, Measuring the unmeasurable: Urban design qualities related to walkability, *Journal of Urban design* 14 (1) (2009) 65–84.
- [72] K. Ito, M. Quintana, X. Han, R. Zimmermann, F. Biljecki, Translating street view imagery to correct perspectives to enhance bikeability and walkability studies, *International Journal of Geographical Information Science* 38 (12) (2024) 2514–2544.
- [73] M. Nikolopoulou, Outdoor thermal comfort, *Frontiers in Bioscience-Scholar* 3 (4) (2011) 1552–1568.
- [74] W. Klemm, B. G. Heusinkveld, S. Lenzholzer, B. van Hove, Street greenery and its physical and psychological impact on thermal comfort, *Landscape and urban planning* 138 (2015) 87–98.

- [75] S. Yilmaz, B. E. Mutlu, A. Aksu, E. Mutlu, A. Qaid, Street design scenarios using vegetation for sustainable thermal comfort in erzurum, turkey, *Environmental Science and Pollution Research* 28 (2021) 3672–3693.
- [76] Y. Hu, F. Qian, H. Yan, A. Middel, R. Wu, M. Zhu, Q. Han, K. Zhao, H. Wang, F. Shao, et al., Which street is hotter? street morphology may hold clues-thermal environment mapping based on street view imagery, *Building and Environment* 262 (2024) 111838.
- [77] J. Yu, Q. Hu, J. Li, Combining visual intelligence and social-physical urban features facilitates fine-scale seasonality characterization of urban thermal environments, *Building and Environment* (2024) 112088.
- [78] J. J. Gibson, *The theory of affordances*, Hilldale, USA 1 (2) (1977) 67–82.
- [79] A. Shashank, N. Schuurman, R. Copley, S. Lear, Creation of a rough runnability index using an affordance-based framework, *Environment and Planning B: Urban Analytics and City Science* 49 (1) (2022) 321–334.
- [80] G. Fusco, Beyond the built-up form/mobility relationship: Spatial affordance and lifestyles, *Computers, Environment and Urban Systems* 60 (2016) 50–66.
- [81] C. Yang, S. M. Lo, R. Ma, H. Fang, The effect of the perceptible built environment on pedestrians’ walking behaviors in commercial districts: Evidence from hong kong, *Environment and Planning B: Urban Analytics and City Science* 51 (2) (2024) 329–346.
- [82] J. Kruse, Y. Kang, Y.-N. Liu, F. Zhang, S. Gao, Places for play: Understanding human perception of playability in cities using street view images and deep learning, *Computers, Environment and Urban Systems* 90 (2021) 101693.
- [83] L. Gregorians, H. J. Spiers, Affordances for spatial navigation, *Affordances in everyday life: A multidisciplinary collection of essays* (2022) 99–112.
- [84] R. Wang, M. Helbich, Y. Yao, J. Zhang, P. Liu, Y. Yuan, Y. Liu, Urban greenery and mental wellbeing in adults: Cross-sectional mediation analyses on multiple pathways across different greenery measures, *Environmental research* 176 (2019) 108535.
- [85] H. He, X. Lin, Y. Yang, Y. Lu, Association of street greenery and physical activity in older adults: A novel study using pedestrian-centered photographs, *Urban forestry & urban greening* 55 (2020) 126789.

- [86] L. Kawshalya, U. Weerasinghe, D. Chandrasekara, The impact of visual complexity on perceived safety and comfort of the users: A study on urban streetscape of sri lanka, *Plos one* 17 (8) (2022) e0272074.
- [87] Y. Wang, W. Qiu, Q. Jiang, W. Li, T. Ji, L. Dong, Drivers or pedestrians, whose dynamic perceptions are more effective to explain street vitality? a case study in guangzhou, *Remote Sensing* 15 (3) (2023) 568.
- [88] Q. Song, M. Li, W. Qiu, W. Li, D. Luo, The coherence and divergence between the objective and subjective measurement of street perceptions for shanghai, in: *International Conference on Advanced Data Mining and Applications*, Springer, 2022, pp. 244–256.
- [89] W. Qiu, W. Li, X. Liu, Z. Zhang, X. Li, X. Huang, Subjective and objective measures of streetscape perceptions: Relationships with property value in shanghai, *Cities* 132 (2023) 104037.
- [90] K. Ito, F. Biljecki, Assessing bikeability with street view imagery and computer vision, *Transportation research part C: emerging technologies* 132 (2021) 103371.
- [91] A. Larkin, P. Hystad, Evaluating street view exposure measures of visible green space for health research, *Journal of exposure science & environmental epidemiology* 29 (4) (2019) 447–456.
- [92] R. Wang, Y. Lu, J. Zhang, P. Liu, Y. Yao, Y. Liu, The relationship between visual enclosure for neighbourhood street walkability and elders' mental health in china: Using street view images, *Journal of Transport & Health* 13 (2019) 90–102.
- [93] S. Park, E. Deakin, J. S. Lee, Perception-based walkability index to test impact of microlevel walkability on sustainable mode choice decisions, *Transportation Research Record* 2464 (1) (2014) 126–134.
- [94] Z. Gan, M. Yang, Q. Zeng, H. J. Timmermans, Associations between built environment, perceived walkability/bikeability and metro transfer patterns, *Transportation Research Part A: Policy and Practice* 153 (2021) 171–187.
- [95] S. J. Strath, M. J. Greenwald, R. Isaacs, T. L. Hart, E. K. Lenz, C. J. Dondzila, A. M. Swartz, Measured and perceived environmental characteristics are related to accelerometer defined physical activity in older adults, *International Journal of Behavioral Nutrition and Physical Activity* 9 (2012) 1–9.

- [96] E. Jamei, P. Rajagopalan, M. Seyedmahmoudian, Y. Jamei, Review on the impact of urban geometry and pedestrian level greening on outdoor thermal comfort, *Renewable and Sustainable Energy Reviews* 54 (2016) 1002–1017.
- [97] K. Labdaoui, S. Mazouz, M. Moeinaddini, M. Cools, J. Teller, The street walkability and thermal comfort index (swtci): A new assessment tool combining street design measurements and thermal comfort, *Science of the total environment* 795 (2021) 148663.
- [98] A. R. Sadeghi, Y. Bahadori, Urban sustainability and climate issues: The effect of physical parameters of streetscape on the thermal comfort in urban public spaces; case study: Karimkhan-e-zand street, shiraz, iran, *Sustainability* 13 (19) (2021) 10886.
- [99] Y. Zheng, Q. Han, G. Keeffe, An evaluation of different landscape design scenarios to improve outdoor thermal comfort in shenzhen, *Land* 13 (1) (2024) 65.
- [100] J. Urban, M. Píkl, F. Zemek, J. Novotný, Using google street view photographs to assess long-term outdoor thermal perception and thermal comfort in the urban environment during heatwaves, *Frontiers in Environmental Science* 10 (2022) 878341.
- [101] T. Zhang, B. Hong, X. Su, Y. Li, L. Song, Effects of tree seasonal characteristics on thermal-visual perception and thermal comfort, *Building and Environment* 212 (2022) 108793.
- [102] T. Yan, H. Jin, Y. Jin, The mediating role of emotion in the effects of landscape elements on thermal comfort: A laboratory study, *Building and Environment* 233 (2023) 110130.
- [103] M. B. Motie, M. Yeganeh, M. Bemanian, Assessment of greenery in urban canyons to enhance thermal comfort & air quality in an integrated seasonal model, *Applied Geography* 151 (2023) 102861.
- [104] H. Yan, F. Wu, X. Nan, Q. Han, F. Shao, Z. Bao, Influence of view factors on intra-urban air temperature and thermal comfort variability in a temperate city, *Science of The Total Environment* 841 (2022) 156720.
- [105] Y. Kim, D. Li, Y. Xu, Y. Zhang, X. Li, L. Muhlenforth, S. Xue, R. Brown, Heat vulnerability and street-level outdoor thermal comfort in the city of houston: Application of google street view image derived svfs, *Urban Climate* 51 (2023) 101617.

- [106] N. Klimenko, K. Zhao, R. Hilland, F. Zhang, J. Voogt, C. Ratti, Instant infrared: Estimating urban surface temperatures from street view imagery, *Building and Environment* (2024) 112122.
- [107] M. Nikolopoulou, K. Steemers, Thermal comfort and psychological adaptation as a guide for designing urban spaces, *Energy and buildings* 35 (1) (2003) 95–101.
- [108] J. Van Hoof, M. Mazej, J. L. Hensen, Thermal comfort: research and practice, *Frontiers in Bioscience* 15 (2) (2010) 765–788.
- [109] J. F. Nicol, S. Roaf, *Rethinking thermal comfort* (2017).
- [110] F. J. de Feijter, B. J. van Vliet, Housing retrofit as an intervention in thermal comfort practices: Chinese and dutch householder perspectives, *Energy Efficiency* 14 (1) (2021) 2.
- [111] D. Lai, Z. Lian, W. Liu, C. Guo, W. Liu, K. Liu, Q. Chen, A comprehensive review of thermal comfort studies in urban open spaces, *Science of the Total Environment* 742 (2020) 140092.
- [112] ANSI/ASHRAE Standard 55, *Thermal environmental conditions for human occupancy* (2023).
- [113] S. Shooshtarian, P. Rajagopalan, Study of thermal satisfaction in an australian educational precinct, *Building and Environment* 123 (2017) 119–132.
- [114] J. Bligh, K. G. Johnson, Glossary of terms for thermal physiology., *Journal of Applied Physiology* 35 (6) (1973) 941–961.
- [115] K. Parsons, *Human thermal environments: the effects of hot, moderate, and cold environments on human health, comfort and performance*, CRC press, 2007.
- [116] R. De Dear, G. Schiller Brager, The adaptive model of thermal comfort and energy conservation in the built environment, *International journal of biometeorology* 45 (2001) 100–108.
- [117] Y. Dzyuban, G. N. Ching, S. K. Yik, A. J. Tan, S. Banerjee, P. J. Crank, W. T. Chow, Outdoor thermal comfort research in transient conditions: A narrative literature review, *Landscape and Urban Planning* 226 (2022) 104496.

- [118] J. R. Maier, G. M. Fadel, D. G. Battisto, An affordance-based approach to architectural theory, design, and practice, *Design Studies* 30 (4) (2009) 393–414.
- [119] J. R. Maier, G. M. Fadel, Affordance: the fundamental concept in engineering design, in: *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Vol. 80258, American Society of Mechanical Engineers, 2001, pp. 177–186.
- [120] P. Ponce, D. Balderas, T. Peffer, A. Molina, Deep learning for automatic usability evaluations based on images: A case study of the usability heuristics of thermostats, *Energy and Buildings* 163 (2018) 111–120.
- [121] R. Aradi, K. Halvorsen Thorén, I. Fjørtoft, The urban landscape as affordance for adolescents' everyday physical activity, *Landscape research* 41 (5) (2016) 569–584.
- [122] P. Kumar, A. Sharma, Study on importance, procedure, and scope of outdoor thermal comfort—a review, *Sustainable Cities and Society* 61 (2020) 102297.
- [123] I. A. Balogun, M. T. Daramola, The outdoor thermal comfort assessment of different urban configurations within akure city, nigeria, *Urban Climate* 29 (2019) 100489.
- [124] G. Mutani, S. Beltramino, et al., Geospatial assessment and modeling of outdoor thermal comfort at urban scale, *International Journal of Heat and Technology* 40 (4) (2022) 871–878.
- [125] A. Qaid, H. B. Lamit, D. R. Ossen, M. H. Rasidi, Effect of the position of the visible sky in determining the sky view factor on micrometeorological and human thermal comfort conditions in urban street canyons, *Theoretical and applied climatology* 131 (2018) 1083–1100.
- [126] S. Yilmaz, E. A. Külekçi, B. E. Mutlu, I. Sezen, Analysis of winter thermal comfort conditions: street scenarios using envi-met model, *Environmental Science and Pollution Research* 28 (2021) 63837–63859.
- [127] W. H. Ko, S. Schiavon, H. Zhang, L. T. Graham, G. Brager, I. Mauss, Y.-W. Lin, The impact of a view from a window on thermal comfort, emotion, and cognitive performance, *Building and Environment* 175 (2020) 106779.
- [128] Y. Yang, L. Hu, R. Zhang, X. Zhu, M. Wang, Investigation of students' short-term memory performance and thermal sensation with heart rate vari-

ability under different environments in summer, *Building and Environment* 195 (2021) 107765.

- [129] N. P. A. Nitidara, J. Sarwono, S. Suprijanto, F. N. Soelami, The multisensory interaction between auditory, visual, and thermal to the overall comfort in public open space: A study in a tropical climate, *Sustainable Cities and Society* 78 (2022) 103622.
- [130] Y. Hou, F. Biljecki, A comprehensive framework for evaluating the quality of street view imagery, *International Journal of Applied Earth Observation and Geoinformation* 115 (2022) 103094.
- [131] X. Li, C. Zhang, W. Li, R. Ricard, Q. Meng, W. Zhang, Assessing street-level urban greenery using google street view and a modified green view index, *Urban Forestry & Urban Greening* 14 (3) (2015) 675–685.
- [132] K. Fujiwara, K. Ito, M. Ignatius, F. Biljecki, A panorama-based technique to estimate sky view factor and solar irradiance considering transmittance of tree canopies, *Building and Environment* 266 (2024) 112071.
- [133] Y. Zhang, P. Liu, F. Biljecki, Knowledge and topology: A two layer spatially dependent graph neural networks to identify urban functions with time-series street view image, *ISPRS Journal of Photogrammetry and Remote Sensing* 198 (2023) 153–168.
- [134] F. Biljecki, T. Zhao, X. Liang, Y. Hou, 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 (2023) 103385.
- [135] F. Yang, S. S. Lau, F. Qian, Thermal comfort effects of urban design strategies in high-rise urban environments in a sub-tropical climate, *Architectural Science Review* 54 (4) (2011) 285–304.
- [136] X. Deng, W. Nie, X. Li, J. Wu, Z. Yin, J. Han, H. Pan, C. K. C. Lam, Influence of built environment on outdoor thermal comfort: A comparative study of new and old urban blocks in guangzhou, *Building and Environment* 234 (2023) 110133.
- [137] F. Guan, Z. Fang, L. Wang, X. Zhang, H. Zhong, H. Huang, Modelling people’s perceived scene complexity of real-world environments using street-view panoramas and open geodata, *ISPRS Journal of Photogrammetry and Remote Sensing* 186 (2022) 315–331.

- [138] B. Ma, R. J. Hauer, C. Xu, W. Li, Visualizing evaluation model of human perceptions and characteristic indicators of landscape visual quality in urban green spaces by using nomograms, *Urban Forestry & Urban Greening* 65 (2021) 127314.
- [139] L. Dong, H. Jiang, W. Li, B. Qiu, H. Wang, W. Qiu, Assessing impacts of objective features and subjective perceptions of street environment on running amount: A case study of boston, *Landscape and urban planning* 235 (2023) 104756.
- [140] S. Shooshtarian, I. Ridley, The effect of physical and psychological environments on the users thermal perceptions of educational urban precincts, *Building and Environment* 115 (2017) 182–198.
- [141] C. K. C. Lam, H. Yang, X. Yang, J. Liu, C. Ou, S. Cui, X. Kong, J. Hang, Cross-modal effects of thermal and visual conditions on outdoor thermal and visual comfort perception, *Building and Environment* 186 (2020) 107297.
- [142] Y. Liu, Z. Yu, Y. Song, X. Yu, J. Zhang, D. Song, Psychological influence of sky view factor and green view index on daytime thermal comfort of pedestrians in shanghai, *Urban Climate* 56 (2024) 102014.
- [143] M. Knauff, A neuro-cognitive theory of deductive relational reasoning with mental models and visual images, *Spatial Cognition & Computation* 9 (2) (2009) 109–137.
- [144] H. Ögmen, M. H. Herzog, A new conceptualization of human visual sensory-memory, *Frontiers in psychology* 7 (2016) 830.
- [145] C. Wang, F. Zhang, J. Wang, J. K. Doyle, P. A. Hancock, C. M. Mak, S. Liu, How indoor environmental quality affects occupants’ cognitive functions: A systematic review, *Building and environment* 193 (2021) 107647.
- [146] M. Bar, The proactive brain: memory for predictions, *Philosophical Transactions of the Royal Society B: Biological Sciences* 364 (1521) (2009) 1235–1243.
- [147] H. Eichenbaum, N. J. Fortin, The neurobiology of memory based predictions, *Philosophical Transactions of the Royal Society B: Biological Sciences* 364 (1521) (2009) 1183–1191.
- [148] K. Friston, Learning and inference in the brain, *Neural Networks* 16 (9) (2003) 1325–1352.

- [149] E. T. Berkman, E. B. Falk, Beyond brain mapping: Using neural measures to predict real-world outcomes, *Current directions in psychological science* 22 (1) (2013) 45–50.
- [150] X. Ma, M. Du, P. Deng, T. Zhou, B. Hong, Effects of green walls on thermal perception and cognitive performance: An indoor study, *Building and Environment* 250 (2024) 111180.
- [151] R. Sharma, L. Pradhan, M. Kumari, P. Bhattacharya, Assessing urban heat islands and thermal comfort in noida city using geospatial technology, *Urban Climate* 35 (2021) 100751.
- [152] S. Wu, W. Yu, B. Chen, Observed inequality in thermal comfort exposure and its multifaceted associations with greenspace in united states cities, *Landscape and Urban Planning* 233 (2023) 104701.
- [153] D. Anguelov, C. Dulong, D. Filip, C. Frueh, S. Lafon, R. Lyon, A. Ogale, L. Vincent, J. Weaver, Google street view: Capturing the world at street level, *Computer* 43 (6) (2010) 32–38.
- [154] W. Yang, N. H. Wong, S. K. Jusuf, Thermal comfort in outdoor urban spaces in singapore, *Building and environment* 59 (2013) 426–435.
- [155] R. Herbrich, T. Minka, T. Graepel, Trueskill™: a bayesian skill rating system, *Advances in neural information processing systems* 19 (2006).
- [156] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, H. Adam, Encoder-decoder with atrous separable convolution for semantic image segmentation, in: *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 801–818.
- [157] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [158] S. Ren, K. He, R. Girshick, J. Sun, Faster r-cnn: Towards real-time object detection with region proposal networks, *Advances in neural information processing systems* 28 (2015).
- [159] G. Bradski, The opencv library., *Dr. Dobb’s Journal: Software Tools for the Professional Programmer* 25 (11) (2000) 120–123.

- [160] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, C. L. Zitnick, Microsoft coco: Common objects in context, in: *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part V 13*, Springer, 2014, pp. 740–755.
- [161] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva, A. Torralba, Places: A 10 million image database for scene recognition, *IEEE transactions on pattern analysis and machine intelligence* 40 (6) (2017) 1452–1464.
- [162] P. Liu, T. Zhao, J. Luo, B. Lei, M. Frei, C. Miller, F. Biljecki, Towards human-centric digital twins: Leveraging computer vision and graph models to predict outdoor comfort, *Sustainable Cities and Society* 93 (2023) 104480.
- [163] F. Tartarini, M. Frei, S. Schiavon, Y. X. Chua, C. Miller, Cozie apple: An ios mobile and smartwatch application for environmental quality satisfaction and physiological data collection, in: *Journal of Physics: Conference Series*, Vol. 2600, IOP Publishing, 2023, p. 142003.
- [164] M. Ignatius, J. Lim, B. Gottkehaskamp, K. Fujiwara, C. Miller, F. Biljecki, Digital twin and wearables unveiling pedestrian comfort dynamics and walkability in cities, *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences X-4/W5-2024* (2024) 195–202. doi:10.5194/isprs-annals-x-4-w5-2024-195-2024.