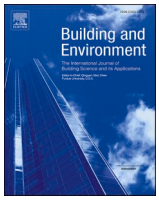
Building and Environment 293 (2026) 114322

Contents lists available at ScienceDirect

Building and Environment

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



Visual determinants of outdoor thermal comfort: integrating explainable AI and perceptual assessments

Lujia Zhu [a], Holly W. Samuelson [b], Filip Biljecki [a], Chun Liang Tan [a], Nyuk Hien Wong [a], Yu Qian Ang [a,*]

[a] College of Design and Engineering, National University of Singapore, Singapore
[b] Building Technology, Massachusetts Institute of Technology, USA

ARTICLE INFO

Keywords:
Outdoor thermal comfort
Visual perception
Explainable machine learning
Computer vision
Thermal sensation
SHAP

ABSTRACT

Outdoor thermal comfort is a crucial determinant of urban space quality. While research has developed various heat indices, such as the Universal Thermal Climate Index (UTCI) and the Physiological Equivalent Temperature (PET), these metrics fail to fully capture perceived thermal comfort. Beyond environmental and physiological factors, recent research suggests that visual elements significantly drive outdoor thermal perception. This study integrates computer vision, explainable machine learning, and perceptual assessments to investigate how visual elements in streetscapes affect thermal perception. To provide a comprehensive representation of diverse visual elements, we employed multiple computer vision models (viz. Segment Anything Model, ResNet-50, and Vision Transformer) and applied the Maximum Clique method to systematically select 50 representative ground-level images, each paired with a corresponding thermal image captured simultaneously. An outdoor, web-based survey among 317 students collected thermal sensation votes (TSV), thermal comfort votes (TCV), and element preference data, yielding 2,854 valid responses. The same survey was replicated in an indoor exhibition setting to provide a comparative reference against the outdoor experiment. A Random Forest classifier achieved 70% and 68% accuracy in predicting thermal sensation and comfort, respectively. Using Shapley Additive Explanations to interpret model outcomes, we uncovered that the colour magenta emerged as the most influential visual factor for thermal perception, while greenery – despite being participants' most preferred element for cooling – showed weaker correlation with actual thermal perception. These findings challenge conventional assumptions about visual thermal comfort and offer a novel framework for image-based thermal perception research, with important implications for climate-responsive urban design.

1. Introduction

1.1. Background

With the intensifying effects of global warming, maintaining thermal comfort in outdoor environments has become increasingly critical. Urban areas worldwide are experiencing more frequent and severe heat events [1]. In tropical countries such as Singapore, the risks are even more pronounced. As a highly urbanized and densely built city, Singapore faces additional challenges such as the urban heat island (UHI) effect, making the issue of reduced outdoor thermal comfort an increasingly urgent concern [2]. A central component in this topic is thermal perception – how people experience thermal conditions in outdoor spaces [3]. Past studies have indicated that human thermal perception in outdoor environments is strongly influenced by a wide range of physiological and psychological factors [4–6]. Heat indices such as the Universal Thermal Climate Index (UTCI) and Physiological Equivalent Temperature (PET) were developed to reveal the relationship between equivalent temperature and thermal sensation through standardized scales. These indices incorporate physiological factors (including air temperature, mean radiation temperature, wind speed, relative humidity, metabolic rate and clothing insulation). However, they do not fully capture the thermal comfort of a given location, as studies have shown that thermal perception is influenced not only by physiological conditions but also by psychological factors (Fig. 1). Recent research has revealed that visual elements also significantly affect perceived outdoor thermal comfort [6].

* Corresponding author.
E-mail address: yuqian@nus.edu.sg (Y.Q. Ang).

https://doi.org/10.1016/j.buildenv.2026.114322
Received 3 November 2025; Received in revised form 18 January 2026; Accepted 2 February 2026
Available online 5 February 2026

1.2. Related work

Visual elements act as psychological primers, shaping people's outdoor thermal perception and expectations. A substantial body of literature has examined the influence of visual cues on thermal experience, adopting diverse visual parameter selection and experimental settings. Some previous studies primarily focused on individual visual parameters, often investigated under controlled or semi-controlled conditions. These include illuminance [7], albedo and glare [8–10], surface materials [8,11], hue [12], and environmental elements such as tree shade, water features, and street furniture [9,13–15]. While these studies provide important insights into specific visual-thermal relationships, their limited-factor selection restrict their ability to capture the visual complexity of real outdoor environments.

Subsequent research extended these investigations to outdoor urban contexts by incorporating aggregated visual metrics, such as sky view factor, building view factor, and green visibility indices [6,16,17]. While these metrics provide scalable representations of urban visual environments, they typically reduce complex visual scenes to a limited set of summary indicators. In response to this limitation, more recent studies have adopted image-based and multi-feature frameworks to capture outdoor visual environments in a more holistic manner. For instance Yang et al. [18] introduced the concept of thermal affordance and proposed the Visual Assessment of Thermal Affordance (VATA) framework, which integrates multiple classes of image-derived features – including scene recognition, semantic segmentation, object detection, colour characteristics, and convolutional image representations – to predict thermal comfort potential at the urban scale. However, while such frameworks categorize visual features broadly, they often overlook the granular attributes that drive perception – such as surface material or fine-grained color distributions – which have been shown to influence thermal psychology in earlier studies.

1.3. Current methodological approaches

Researchers have developed various methodological approaches to study visual-thermal relationships. Three key methodological advances have emerged in the literature that (when integrated) offer promising pathways for more comprehensive analysis of how visual elements influence thermal perception.

1.3.1. Computer vision

Recent advances in computer vision have reshaped how researchers analyze visual environments. Studies increasingly utilize diverse image sources – from regular photographs and street-view images [18,19], to panoramic images [20] – as primary data for understanding urban visual characteristics. Models like PSPNet and FCN, trained on comprehensive datasets such as Cityscapes and ADE20K, enable automated extraction of complex visual parameters. These computational tools, combined with OpenCV and thermal imaging technologies, allow researchers to objectively quantify previously unmeasurable visual elements, including detailed colour distributions, material properties, and spatial configurations. However, while these methods excel at objective measurement, they cannot capture the subjective human experience of thermal environments.

1.3.2. Subjective thermal perception votes and assessments

To address the subjective dimension, researchers have established standardized approaches for measuring human thermal perception. Current understanding recognizes thermal perception as comprising two semantic dimensions: sensation and comfort. Thermal sensation, the objective or descriptive dimension, is most commonly assessed using the ASHRAE seven-point scale. Thermal comfort is the affective or hedonic component of thermal perception [21] – researchers working on these topics have widely adopted the Thermal Sensation Vote (TSV) and Thermal Comfort Vote (TCV) as standardized instruments for quantifying these subjective perceptions (and experiences) [22,23]. These dual assessment methods have proven valuable for capturing both physical sensations and emotional responses to outdoor environments in urban spaces. However, a significant challenge remains in linking these subjective assessments to objective visual measurements.

1.3.3. Explainable artificial intelligence/machine learning approaches

A recent methodological advancement has been the adoption of explainable machine learning techniques to bridge objective measurements with subjective perceptions. Shapley Additive Explanations (SHAP) – derived from game theory – has emerged as a powerful tool for interpreting model predictions [24]. Unlike traditional black-box models, SHAP quantifies each feature's contribution to predictions, revealing not just correlations but causal pathways – in this context, between visual elements and thermal perception. This interpretability is particularly valuable for thermal comfort research, as it allows researchers to move beyond simple prediction to understand which visual

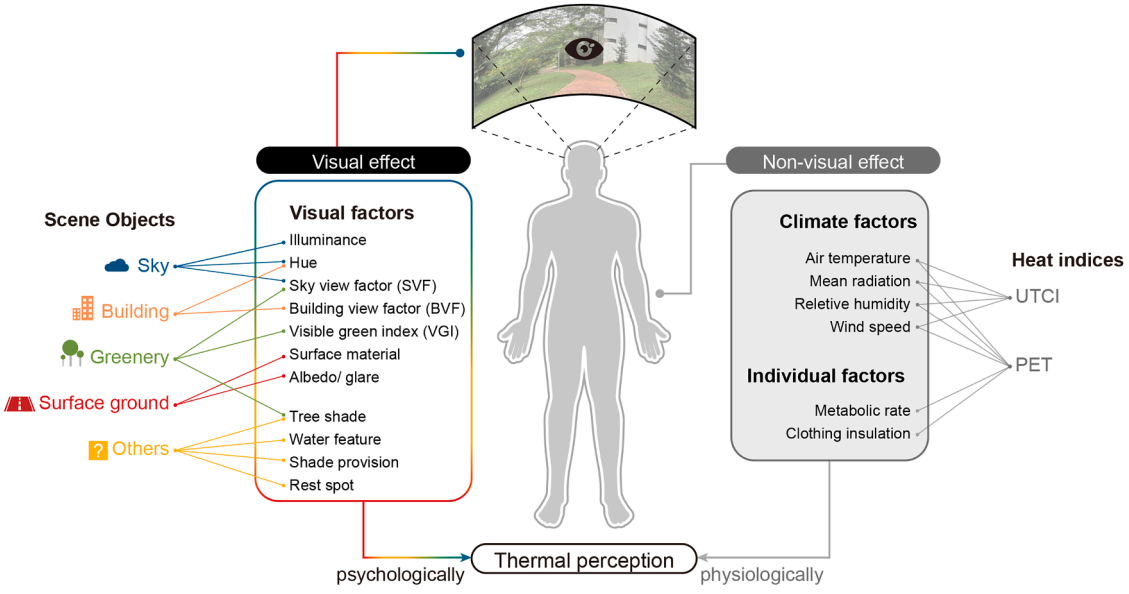


Fig. 1. Theoretical framework illustrating the interplay of visual and non-visual factors (physiological and psychological) in shaping outdoor thermal perception.

factors drive perception and why. Recent applications have demonstrated SHAP's potential for uncovering unexpected relationships between environmental features and comfort outcomes.

1.4. Research gap

Despite these methodological advances, most studies still apply them in isolation or focus on a limited set of visual parameters. To address this critical gap, our study compiles and analyzes a comprehensive set of visual parameters (building upon previous research) to investigate how multiple visual features collectively influence thermal perception. By integrating computer vision, perceptual assessment, and interpretable machine learning within a holistic framework, we aim to provide a more complete understanding of visual-thermal relationships in outdoor spaces – one that could reveal insights about the relative importance of different visual elements in shaping outdoor thermal comfort. Our model can be used for both inference and prediction, serving as a toolkit for human-centric urban design.

1.5. Research objectives and key contributions

This research comprehensively investigates how visual elements influence outdoor thermal perception by integrating computer vision, perceptual assessment, and explainable machine learning approaches. Through this integrated methodology, we seek to challenge existing assumptions about visual-thermal relationships and provide evidence-based insights for climate-responsive urban design.

The specific objectives of this research are:

1. To systematically quantify and compare the influence of diverse visual elements on outdoor thermal perception. Using computer vision techniques to extract comprehensive visual parameters from images, we aim to move beyond single-feature studies to understand the relative importance of colours, materials, vegetation, and spatial configurations in shaping thermal comfort and sensation.
2. To reveal and interpret the mechanisms through which visual features affect thermal perception. By applying explainable machine learning (ensemble tree-based models + SHAP) to bridge objective visual measurements with subjective thermal votes, we seek to uncover not just correlations but the underlying pathways of influence – potentially revealing unexpected relationships.
3. To develop and validate an integrated framework for image-based thermal perception assessment. Our framework will demonstrate how computer vision, standardized perception measures (TSV/TCV), and interpretable machine learning can be effectively combined to predict and explain thermal perception from visual data, providing a scalable methodology for future research and practical urban design applications.

Our study contributes a novel, integrated data-driven framework for understanding outdoor thermal comfort through visual perception. It integrates computer vision, explainable machine learning, and perceptual assessment data to identify key visual predictors of thermal sensation and comfort and help support actionable design insights that bridge physical and psychological aspects of thermal experiences in urban spaces.

2. Methodology

Holistically, our study employs a six-stage integrated methodology (Fig. 2) that progressively builds from data collection to interpretive analysis, ensuring comprehensive investigation of visual-thermal relationships. First, we conducted systematic on-site data collection across the campus of the National University of Singapore, capturing diverse outdoor environments comprehensively. Second, we employed computer vision models and a custom-designed dissimilarity analysis to select maximally diverse and representative images from our initial dataset. Third, we extracted and quantified 34 visual parameters from the selected images using various image processing and computer vision techniques. Fourth, we conducted a web-based survey where participants (situated in outdoor environments) provided thermal sensation votes, thermal comfort votes, and element preferences for each image. Fifth, we systematically evaluated multiple machine learning models – both regression and classification approaches – to identify the optimal model for capturing visual-thermal relationships. Finally, we applied SHAP, correlation analysis, and statistical testing to determine the relative importance and directional influence of visual features on thermal perception, revealing relationships between visual elements and thermal comfort.

2.1. Field data collection and study site

Field data collection was conducted across the National University of Singapore (NUS) campus in the period of 18–27 June 2025, during peak afternoon hours (14:00–17:00) when outdoor thermal stress is typically the highest. These temporal parameters were specifically chosen to capture visual environments under challenging thermal conditions, with ambient temperatures consistently ranging between 30–38.5 °C, as recorded by weather stations deployed across campus (Appendix A1). The NUS campus provided an advantageous study site due to its diverse outdoor environments, including varied topology, vegetation coverage, building densities, surface materials, and spatial configurations representative of tropical urban settings, and it has been subject of climate and comfort research for several years [25,26].

We systematically captured 135 photographs along pedestrian routes and activity areas throughout the campus, ensuring comprehensive coverage of different visual-thermal contexts – from open plazas with high sun exposure to tree-lined pathways with extensive shading. Each photograph was paired with corresponding thermal imagery captured using a thermal camera (Flir One® Pro, Appendix A1), which simultaneously records a photograph and its thermal image with a single shutter press. This ensured spatially aligned surface temperature information, providing ground-truth surface temperature distributions for validating the relationship between visual elements and actual surface temperatures. Unlike photographic images, thermal images cannot correspond to information that the human eye can directly perceive. Nevertheless, thermal images can partially reveal the (thermal) impressions people derive visually from their surroundings, providing an innovative proxy that bridges visual features and the underlying physical condition of the environment. This dual-imaging approach enriches the interpretation of how people perceive outdoor thermal environments.

2.2. Image selection

To ensure our analysis captured the full diversity of visual-thermal environments (while maintaining computational efficiency and a reasonable survey scope), we developed a novel image selection protocol that identified maximally representative images from our initial dataset of 135 photographs. This selection process was important for balancing comprehensive visual coverage with practical constraints of participant survey fatigue and computational resources.

We implemented a multi-model approach using three state-of-the-art computer vision models to capture different aspects of visual similarity: Segment Anything Model (SAM) for semantic understanding, ResNet-50 for general visual features, and Vision Transformer (ViT) for global image relationships. From each model, we extracted latent embeddings (high-dimensional representations capturing the essential visual characteristics of each image). We then computed pairwise cosine similarities between all image embeddings using the formula:

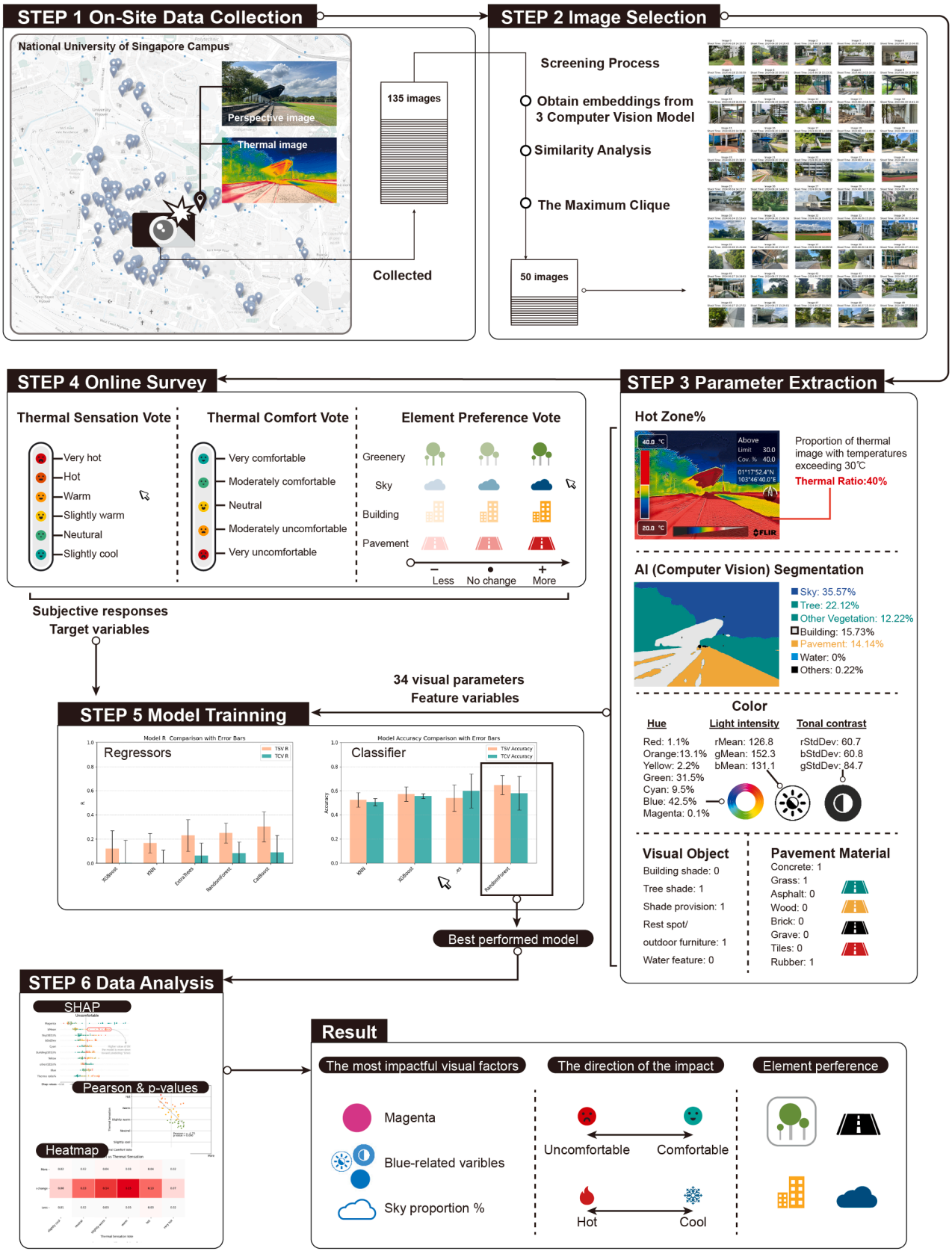


Fig. 2. Overall study methodology.

$$\text{Cosine Similarity} = \frac{\sum_{i=1}^{n} A_i \cdot B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \cdot \sqrt{\sum_{i=1}^{n} B_i^2}}$$

where A and B represent the embedding vectors for two images in the entire set

Fig. 3 shows the distribution of similarity scores across image pairs for each model. To establish an objective selection criterion, we analysed the *rate of change* in similarity score distributions and identified inflection points where the density of similar image pairs changed most rapidly. This analysis yielded model-specific thresholds: 0.40 for SAM, 0.33 for ResNet-50, and 0.45 for ViT (Vision Transformer), below which images were considered sufficiently dissimilar.

Using these thresholds, we applied the Maximum Clique algorithm – a graph theory method that identifies the largest subset of nodes (images) where all pairs meet a specified criterion (dissimilarity). In our implementation, the Maximum Clique represents the largest set of images that are all mutually dissimilar across all three models, ensuring no redundancy in visual content. This selection process yielded 50 images that optimally represent the visual diversity of campus environments, from heavily vegetated areas to exposed concrete plazas, providing a comprehensive (yet manageable dataset) for subsequent perceptual assessment.

2.3. Parameter extraction and selection

2.3.1. Feature extraction

We developed a systematic parameter extraction framework building upon established metrics from previous thermal comfort and urban perception studies, it comprehensively captures the visual elements that may influence thermal perception. Development of the parameter set was guided by three principles: (1) inclusion of established visual metrics proven relevant to environmental perception, (2) comprehensive coverage across different visual dimensions, and (3) computational feasibility for large-scale analysis. We organized these parameters into six distinct categories that capture different aspects of the visual environment: segmented objects, HSV colour distributions (both proportions and statistical properties), surface materials, special visual objects, pixel-level characteristics, and thermal properties. A key novelty in our parameter set, which records thermal properties, is the introduction of *HotZone%*, defined as the proportion of the thermal image where surface temperatures exceed 30°C. This parameter directly links visual analysis to thermal conditions, providing a quantitative measure of the (potentially) visible heat load in each scene. The 30°C threshold was selected based on established thermal comfort research indicating this as a critical temperature for human thermal stress in outdoor environments. Neurological studies have used 30°C as a baseline temperature for measuring human skin thermal responses [27], while thermal remote sensing studies frequently analyze this temperature as a critical threshold in urban heat island research [28].

2.3.2. Feature selection

The extracted visual parameters were subsequently treated as input features for the machine learning model. Prior to model training, an iterative model-informed feature screening process was conducted, guided by principles of recursive feature elimination/selection, model performance evaluation, and correlation analysis. Through this process, non-informative or redundant parameters were excluded, resulting in a final set of 34 visual features used for subsequent modelling. Pairwise correlation analysis was conducted among the retained features (Appendix A2). Although some feature groups exhibited relatively high correlations, these patterns are expected in image-based representations due to inherent compositional constraints. Importantly, correlated features are not mathematically redundant but capture complementary visual dimensions of the same scene elements, such as semantic presence, chromatic intensity, and visual heterogeneity.

Table 1 provides a detailed overview of all 34 parameters, their definitions, variable ranges, and extraction methods. Parameters were extracted using a combination of deep learning-based image segmentation (DeepLab v3), traditional computer vision techniques (OpenCV), manual coding for material identification, and thermal image analysis. This multi-method approach ensures robust quantification of visual elements ranging from basic colour properties to complex spatial and thermal characteristics, providing the comprehensive visual feature set necessary for understanding their relative influence on thermal perception.

2.4. Perceptual data collection through web-based survey

2.4.1. Questionnaire

We designed a web-based survey (Fig. 4) to systematically collect subjective thermal perception data [31]. The survey utilized our 50 representative images selected through the dissimilarity analysis, presenting each image with corresponding assessment questions. To improve ecological validity, we required all 317 participants to complete the entire survey being physically present in an outdoor location of their choice on campus, ensuring consistency – i.e. their thermal responses reflected their thermal perceptions while in warm/hot outdoor conditions rather than in conditioned indoor spaces.

The survey instrument comprised three distinct assessment components designed to capture different dimensions of thermal-visual perception (Fig. 4). Questions 1 and 2 assessed Thermal Sensation Votes and Thermal Comfort Votes respectively, employing established psychometric approaches for quantifying subjective thermal perception. Question 3 introduced an element preference test, asking participants to indicate desired changes (more/less/no change) for four key urban

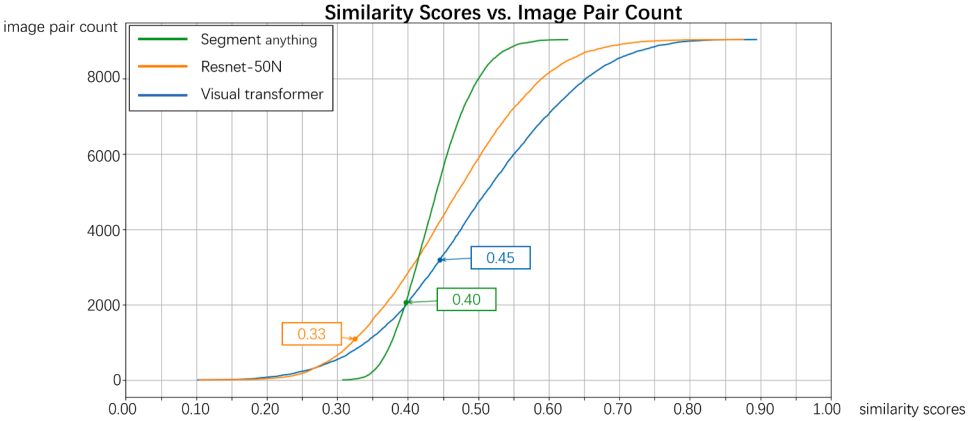


Fig. 3. Similarity scores and threshold.

Table 1
Visual parameters and extraction method.

Parameter category	Visual parameters	Description	Variable range	Measurement method
Segmented object	*Sky*	Proportion of sky	0–100 %	Image segmentation [29] (Deeplab v3)
	Building	Proportion of Building	0–100 %	
	Pavement	Proportion of pavement	0–100 %	
	Vegetation	Proportion of vegetation (tree/ shrub/ turf)	0–100 %	
	Others	Proportion of vegetation (tree/ shrub/ turf)	0–100 %	
HSV*	*Red*	Proportion of red pixels	0–100 %	Python code (OpenCV)
	Orange	Proportion of orange pixels	0–100 %	
	Yellow	Proportion of yellow pixels	0–100 %	
	Green	Proportion of green pixels	0–100 %	
	Cyan	Proportion of cyan pixel	0–100 %	
	Blue	Proportion of blue pixels	0–100 %	
	Magenta	Proportion of magenta pixels	0–100 %	
	Red mean value (rMean)	The average intensity of red colour	0–255	
	Green mean value (gMean)	The average intensity of green colour	0–255	
	Blue mean value (bMean)	The average intensity of blue colour	0–255	
	rStdDev	Measures the variation in the intensity of red	0–127.5	
	gStdDev	Measures the variation in the intensity of green	0–127.5	
	bStdDev	Measures the variation in the intensity of blue	0–127.5	
Surface material	-	Material of pavement: *Concrete; Grass; Asphalt; Wood; Brick; Gravel; Tiles; Rubber*	0 or 1	Image segmentation (Manual check)
Special Visual object	-	Special visual objects that might make people feel cooler *Building shade; Tree shade; Shade provision* (i.e., *canopies, awnings); Rest spot/ outdoor furniture; Water feature*	0 or 1	Manual
Thermal image	*Hot Zone%*	Proportion of thermal image with temperatures exceeding 30°C	0–100 %	Thermal studio software

* HSV colour classification follows established colour perception methodologies [30]. Images were converted from RGB to HSV colour space and pixels categorized into seven colour groups based on OpenCV hue ranges: Red [0–10, 160–179], Orange [11–25], Yellow [26–35], Green [36–85], Cyan [86–100], Blue [101–130], and Magenta [131–159].

elements: greenery, buildings, sky, and pavement. This preference data provides insights into perceived thermal mitigation strategies and priorities for climate-responsive urban design.

Responses for TSV and TCV were numerically encoded for quantitative analysis. Specifically, TSV responses ranging from "Slightly cool" to "Very hot" were encoded on a scale from -1 to 4, while TCV responses from "Very uncomfortable" to "Very comfortable" were encoded from -2 to 2, facilitating standardized data handling for regression and classification modelling.

2.4.2. Adaptation of thermal scales for tropical context

A critical methodological consideration was adapting standard thermal assessment scales to Singapore's consistently warm tropical climate. The questionnaire design was informed by two widely used standards: ISO 10551:2019 [32], which provides general guidance for the design of subjective judgment scales in environmental perception research, and ASHRAE 55 [33], which is commonly applied in the assessment of physical thermal environments and thermal sensation. Following these standards, the initial survey design was based on a 9-point TSV scale and a 5-point TCV scale. However, local validation studies have shown that cold-end sensation categories (very cold, cold, cool) are psychologically irrelevant in tropical outdoor contexts where air temperatures rarely fall below 25°C [34]. Based on this empirical evidence, these cold-related categories were removed, resulting in a reduced 6-point TSV scale ranging from slightly cool to very hot ("very cold", "cold", and "cool" were removed). This approach is consistent with prior literature [34,35].

2.4.3. Sample size

To determine the minimum required sample size for this study, a standard statistical approach for finite population sampling was used. The total target population was estimated to be ~50,000 individuals. We selected a 90 % confidence level and a 5 % margin of error, balancing statistical robustness with practical feasibility. Initially assuming the most conservative scenario ($p = 0.5$) for maximum variability, the preliminary sample size for an infinite population was calculated using the formula:

$$n_0 = \frac{Z^2 \cdot p \cdot (1-p)}{e^2}$$

where Z is the Z-score for a 90 % confidence level (1.645 for 90 %), p is the estimated proportion of the attribute being measured (0.5), and e is the margin of error (0.05). This calculation yielded an initial sample size of approximately 271.

$$n_0 = \frac{1.645^2 \cdot 0.5 \cdot (1-0.5)}{0.05^2} \approx 270.6$$

To account for the finite population size, we applied the finite population correction (FPC) formula:

$$n = \frac{n_0 \cdot N}{N - 1 + n_0}$$

where N = 50,000 and n_0=270.6, the corrected sample size was approximately 270. This sample size provides confidence that the findings will accurately represent the broader population within the specified confidence and margin of error.

2.5. Model development and evaluation

2.5.1. Preliminary evaluation of (explainable) machine learning models

Prior to model development, we conducted an initial evaluation of several machine learning algorithms to identify the most effective approach for predicting thermal comfort. Guided by previous literature on urban environmental modelling and thermal perception prediction and avoiding an a priori commitment to a single modelling paradigm,

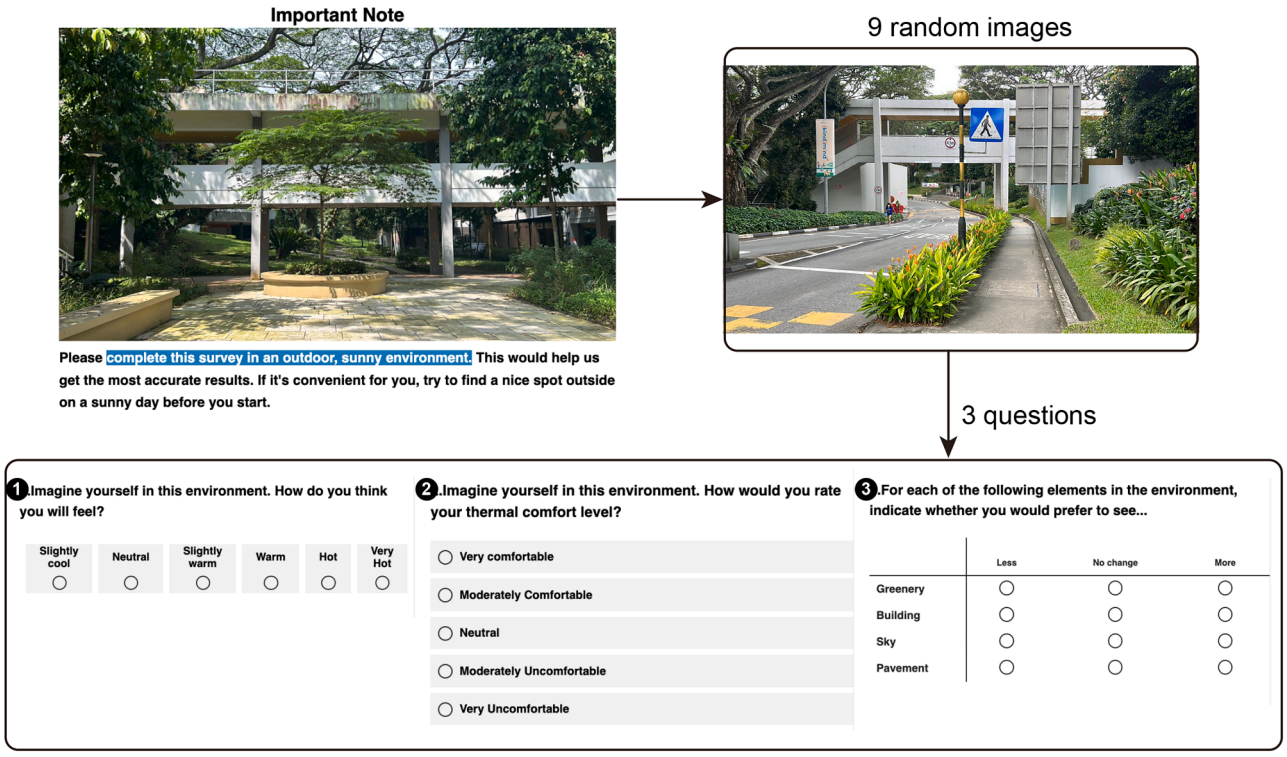


Fig. 4. Web-based voting platform.

we selected a set of regression and classification algorithms for comparison, including Random Forest, XGBoost, ExtraTrees, Catboost, and K-Nearest Neighbours (KNN) [30,36–38]. These are mostly tree-based (ensemble) models that balance interpretability with predictive prowess, with the exception of KNN.

All candidate models were optimized using grid search with 10-fold cross-validation, and their performance was evaluated using task-appropriate metrics (R^2 for regression models and accuracy and macro-averaged F1-score for classification models). After systematically comparing the cross-validated performance across all models, we selected the best-performing model for subsequent detailed analysis and interpretation.

2.5.2. Equal frequency binning

Further data preprocessing was conducted to ensure robust and interpretable results. Equal frequency binning was employed to discretize thermal sensation votes into three categories (i.e. Cool, Neutral, and Hot) based on percentile distributions. Similarly, thermal comfort votes were discretized into three corresponding categories (i.e. Uncomfortable, Neutral, and Comfortable).

This approach aligns with prior studies (Ballantyne et al., 1977; Schweiker et al., 2017, 2020) and is particularly important because thermal comfort perception scales are inherently non-equidistant; individuals do not necessarily perceive the difference between scale points (e.g., from 1 to 2, 2 to 3, or neutral to warm, warm to hot) uniformly, highlighting the need for categorization based on distribution rather than equal interval assumptions. We then implemented a 10-fold cross-validation procedure.

2.6. Data analysis

2.6.1. Model selection and interpretation

Based on model selection results, we employed the Random Forest Classifier, with post-hoc application of SHAP to provide an interpretable means to model predictions [39,40]. SHAP values quantify the impact of each input feature on model predictions, enabling clear interpretation of feature importance.

To explore the relationships between participants' TSV, TCV, and preferences for visual elements (sky, building, pavement, and vegetation), we employed heatmap visualization combined with correlation analysis. Preferences were encoded as ordinal variables (More = 1, No change = 0, Less = −1).

Pearson correlation coefficients were computed between element preferences and TSV/TCV to assess linear relationships. The strength and direction of these relationships were quantified, and significance was determined using p-values. A positive correlation indicates that a higher preference for a particular visual element corresponds to warmer thermal sensations or increased discomfort, whereas a negative correlation suggests the opposite. Statistically significant correlations ($p < 0.05$) indicate meaningful psychological/subjective impacts of visual elements on thermal perception.

3. Results

3.1. Selection of representative images

As shown in Fig. 5, similarity scores varied across the three computer vision models due to distinct feature extraction methods (and their associated latent spaces) inherent to each model. To achieve a comprehensive and representative image selection, the Maximum Clique method was individually applied to each model. Specifically, 28 images from the Segment Anything Model, 14 from ResNet-50, and 21 from the Vision Transformer were selected. After combining these selections and eliminating duplicates, a final set of 50 unique images was utilized for the subsequent web-based survey.

3.2. Extracted visual parameters

From the finalized set of images, we then extracted a set of 34 parameters, categorized into several key groups. These parameters included surface materials (i.e. *concrete, grass, asphalt, wood, brick, gravel, tiles, rubber*), special visual objects (i.e. *building shade, tree shade,*

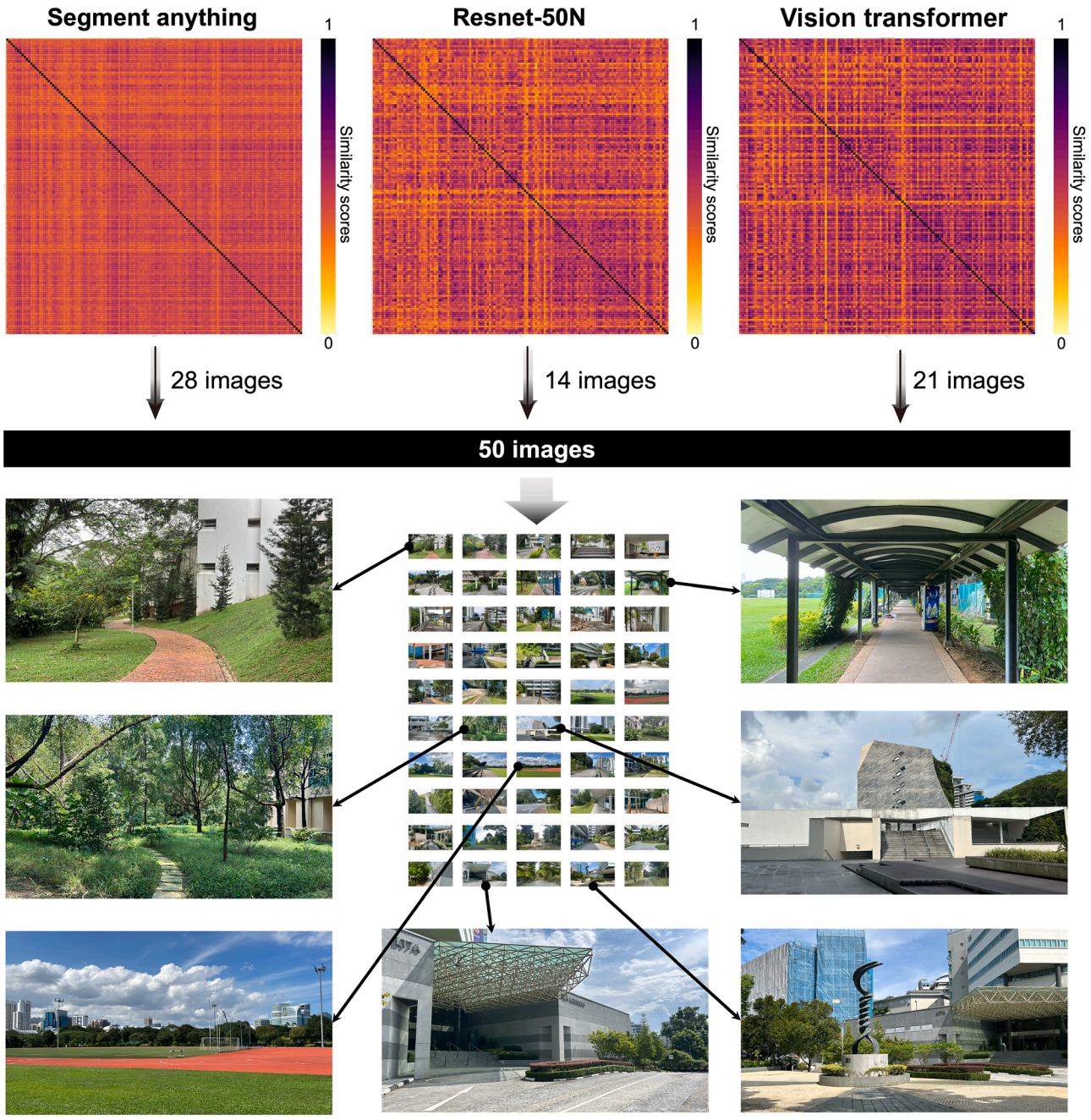


Fig. 5. Similarity scores metrices and selected images.

other shading provision, rest spot/outdoor furniture, water feature), the proportion of the image with a temperature greater than 30°C ('*HotZone %*'), hue (i.e. *red, orange, yellow, green, cyan, blue, magenta*), light intensity (i.e. *rMean, gMean, bMean*), tonal contrast (i.e. *rStdDev, gStdDev, bStdDev*) [41–43] and image segmentation proportions (i.e. sky, pavement, tree, other vegetation, building, water and others). These parameters constituted the input features for subsequent machine learning analyses.

3.3. Analysis of thermal (Sensation and comfort) votes

A total of 2,854 valid responses (Appendix A3, Appendix A4) were collected from 317 participants. Fig. 6 displays the average TSV and TCV scores for each of the 50 selected images, ranked according to perceived thermal sensation from coolest to hottest. The ranking was derived from mean TSV scores calculated from participant responses.

Overall, TCV scores closely aligned with TSV rankings, indicating that images perceived as hotter were typically rated as less comfortable. Nevertheless, minor discrepancies emerged between TSV and TCV scores, suggesting variations in individual comfort perceptions despite similar thermal sensations. To quantitatively assess the relationship between TSV and TCV, a Pearson correlation analysis was conducted, revealing a strong and statistically significant negative correlation ($r = -0.97, p < 0.001$). This finding confirms that higher thermal sensations (feeling hotter) were closely associated with decreased comfort levels.

The heatmap in Fig. 6 further illustrates the distribution of TCV responses within each TSV category, showing a distinct diagonal pattern. For example, 70 % of respondents experiencing "very hot" sensations reported feeling "very uncomfortable," while 57 % of those feeling "slightly cool" indicated being "comfortable," with an additional 30 % rating themselves as "very comfortable." Nonetheless, notable variability was observed within moderate thermal sensation categories

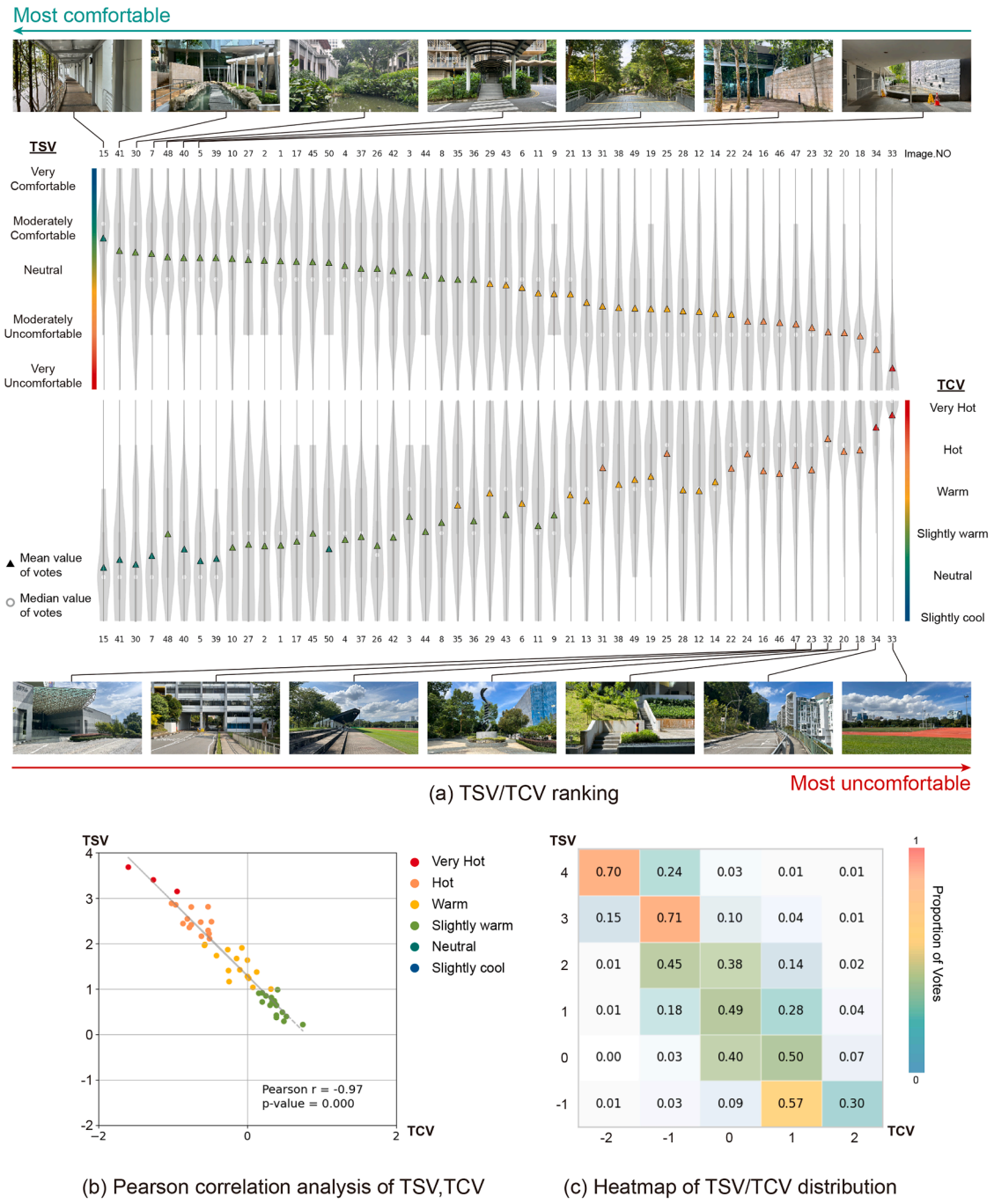


Fig. 6. TSV/TCV response distribution, Pearson correlation analysis and heatmap.

("neutral," "slightly warm," and "warm"), indicating subjective variability in perceived comfort.

Qualitative visual inspection of the Top 5 hottest and Top 5 coolest images revealed distinguishable visual characteristics. Hotter images typically featured extensive sky exposure, high colour contrast, and minimal shading. Conversely, images perceived as cooler often had dense vegetation and visible shade structures. These visual distinctions underscore potential *drivers* of thermal perception, necessitating detailed statistical validation in subsequent analyses.

Overall, the findings support a robust correlation between thermal sensation and comfort while highlighting meaningful perceptual variability under moderate thermal conditions. Future analysis should further examine these divergences to develop a nuanced understanding of subjective thermal experiences.

3.4. Model training and optimal model selection

For both regression and classification models, hyperparameter optimization was conducted using grid search combined with cross-validation. For regression models, the encoded thermal responses were treated as continuous target variables. We evaluated multiple (ensemble tree-based) regression algorithms. Hyperparameter optimization for the regressors focused primarily on model depth and ensemble size, with max_depth ranging from 1 to 15 and n_estimators ranging from 50 to 250. For KNN regression, the number of neighbours (k) was varied between 3 and 15. All models were initialized with a fixed random seed (42) to ensure reproducibility.

For the classification models, the encoded thermal responses were first transformed into three discrete classes using an equal-frequency

binning strategy. This binning approach was adopted to address the non-equidistant perceptual spacing of ordinal thermal sensation scales [44]. We evaluated a range of classification algorithms, and hyperparameters for the tree-based classifiers were optimized by varying max_depth (1–15) and n_estimators (50–250), while k was tuned for the KNN classifier. All classification models were also trained with the same fixed random seed for consistency. All models were evaluated using 10-fold cross-validation.

As shown in Fig. 7, across all tested regression algorithms and hyperparameter configurations, the coefficient of determination (R^2) consistently remained below 0.4, indicating limited explanatory power when modelling thermal perception as a continuous variable. For classification models, performance was primarily assessed using overall accuracy for intuitive model comparison. The classification models achieved average accuracy exceeding 0.6, indicating substantially stronger predictive performance compared to regression approaches.

This performance difference between the regressors and classifiers is likely attributable to the fact that regression approaches assume a continuous dependent variable with equidistant numerical encoding, which may not fully capture the discrete or ordinal nature of thermal perception responses. In contrast, classification models can handle categorical or ordinal outcomes, making them better suited to the data characteristics in this study.

Among the tested classifiers, the Random Forest Classifier (random_state = 42, n_estimators = 150, max_depth = 8) emerged as the most effective model, achieving accuracy scores of 70 % for TSV and 68 % for TCV.

3.5. Model outcome interpretation and feature importance

Based on the SHAP analysis, both feature importance plots and SHAP beeswarm plots were generated for TSV and TCV. These plots outline the overall contribution of each predictor, offering a global perspective on feature importance. Meanwhile, the SHAP beeswarm plot provides more detailed local interpretability by showing how variations in each feature's values influence individual prediction outcomes.

The SHAP summary plots (Fig. 8a, c) highlighted *Magenta* as the most influential visual feature in predicting thermal comfort and sensation. Other significant predictors included, sky coverage(*Sky(SEG)%*), blue-related variables (i.e. *Blue, bMean, bStdDev*) and building coverage. Similarly, the detailed SHAP beeswarm plots (Fig. 8b, d) revealed that increased *Magenta* strongly predicted cooler and more comfortable perceptions, whereas greater *Blue* and *Sky(SEG)%* predicted hotter and less comfortable conditions, thus demonstrating clear directional influences of visual parameters on thermal perception. Another noteworthy predictor is *HotZone%,* which represents the proportion of an image area with surface temperatures exceeding 30°C. Although not the most influential predictor, its effect exhibited a consistent pattern: higher *HotZone%* values were associated with warmer thermal perceptions. Importantly, this indicator also reflects an aspect of the physical environment, suggesting that visual perception may partially encompasses cues regarding underlying physical conditions – that is, people may infer certain environmental properties, such as surface temperature, through visual scenes, thereby invoking corresponding thermal perceptions.

3.5.1. Comparative feature category analysis

Fig. 9 presents two polar charts that summarize the SHAP values of major visual feature categories. The cumulative SHAP value for each category was calculated based on the grouped features under six categories.

1. Hue: red, orange, blue, yellow, green, cyan, magenta
2. Light intensity: rMean, gMean, bMean
3. Tonal contrast: bStdDev, rStdDev, gStdDev
4. Segmentation Objects: sky(SEG)%, pavement(SEG)%, water(SEG)%, building(SEG)%, tree(SEG)%, otherVegetation(SEG)%, other(SEG)%
5. Surface Temperature: HotZone%
6. Surface Material: concrete, grass, asphalt, wood, brick, gravel, tiles, rubber

Hue and Segmentation emerged as the most influential categories overall. Specifically, Hue had a more pronounced influence on TCV predictions compared to segmentation features, whereas the impact of these two categories was relatively balanced for TSV predictions. Surface Temperature and Surface Material categories exhibited the lowest influence across both models. Additionally, Tonal Contrast showed a notably stronger impact on TSV, indicating it plays a critical role in perceiving how hot or cool an environment feels. Conversely, Light Intensity had a greater impact on TCV, suggesting it significantly influences perceptions of environmental comfort.

3.6. Element preferences and thermal votes

To investigate the association between TSV, TCV and preferences for urban elements (viz. greenery, pavement, sky, buildings), Pearson correlations and heatmap analyses were conducted (Fig. 10). Sky visibility showed the strongest correlation with thermal perceptions, exhibiting a significant negative correlation with TSV ($r = -0.75$, $p < 0.001$) and a significant positive correlation with TCV ($r = 0.72$, $p < 0.001$). This indicates participants associated higher sky exposure with cooler sensations and enhanced comfort. In contrast, correlations for greenery, pavement, and buildings were considerably weaker, all with low r ($r < 0.31$) and the association were not significant (p-value>0.05)

Heatmap analyses revealed a pronounced preference for increased greenery, particularly under warmer conditions, despite its weaker statistical correlation with TSV and TCV. Preferences for sky, buildings, and pavement were predominantly neutral, reflecting less distinct attitudes toward altering these elements. These findings suggest participants strongly value greenery despite the stronger statistical (and model) linkage of sky visibility with thermal perception. This result may be influenced by a commonly held belief that vegetation contributes to reducing ambient temperatures [45].

3.7. Survey group comparison

This study's dataset primarily consisted of responses from students aged 20-30, collected through a web-based survey conducted outdoors. While this approach allows us to keep the survey and study context consistent and control for potential demographic biases, it may limit the generalizability of results to broader urban populations and varying age groups. To address this limitation, we conducted a parallel indoor public survey at the Singapore City Gallery as part of the "Well-being in the City" national exhibition, involving a diverse age demographic. The survey was set up as an interactive exhibit (Fig. 11), with the same set of locational images shown randomly. Participants were instructed to imagine themselves in each scene and then respond to the same TCV question. A total of 755 responses across 50 images were collected over two months of the exhibition's duration.

A comparative analysis (Fig. 11) revealed notable variations between these two groups. Group B (indoor survey) consistently exhibited greater variance in TCV and generally reported higher comfort levels, likely due to the controlled (air-conditioned) indoor environment. Interestingly, certain images rated as highly comfortable by Group A (the outdoor participants) were perceived as less comfortable by Group B, particularly those featuring shade, reduced *sky(SEG)%*, lower *bStdDev* and *HotZone%* values; as well as increased *magenta* and *building(SEG)%*. These features, beneficial for outdoor participants, lost their perceptual effectiveness for the indoor participants. Other reasons for this result may include redundancy of shading, age-related variations, and indoor versus outdoor environmental settings.

Class-wise classification performance for TSV and TCV prediction

TSV	Precision	Recall	F1-score	Support sample
Cool	0.73	0.65	0.69	17
Neutral	0.65	0.69	0.67	16
Hot	0.72	0.76	0.74	17
macro avg	0.70	0.70	**0.70**	50
accuracy			**0.70**	50

TCV	Precision	Recall	F1-score	Support sample
Uncomfortable	0.71	0.71	0.71	17
Neutral	0.71	0.59	0.65	17
Comfortable	0.63	0.75	0.69	16
macro avg	0.68	0.68	**0.68**	50
accuracy			**0.68**	50

Fig. 7. Model selection and evaluation process.

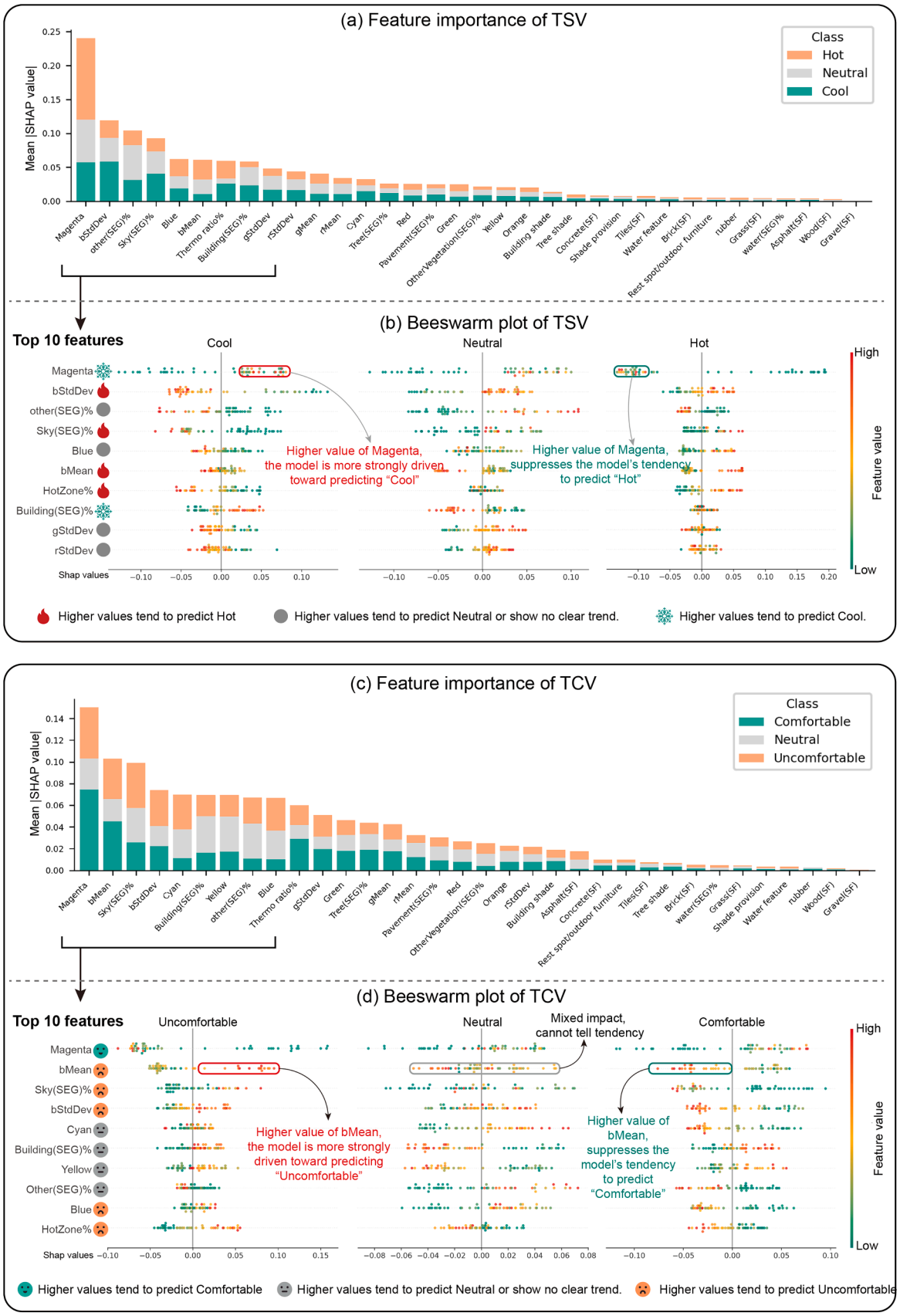


Fig. 8. SHAP analyses of visual features affecting thermal perception.

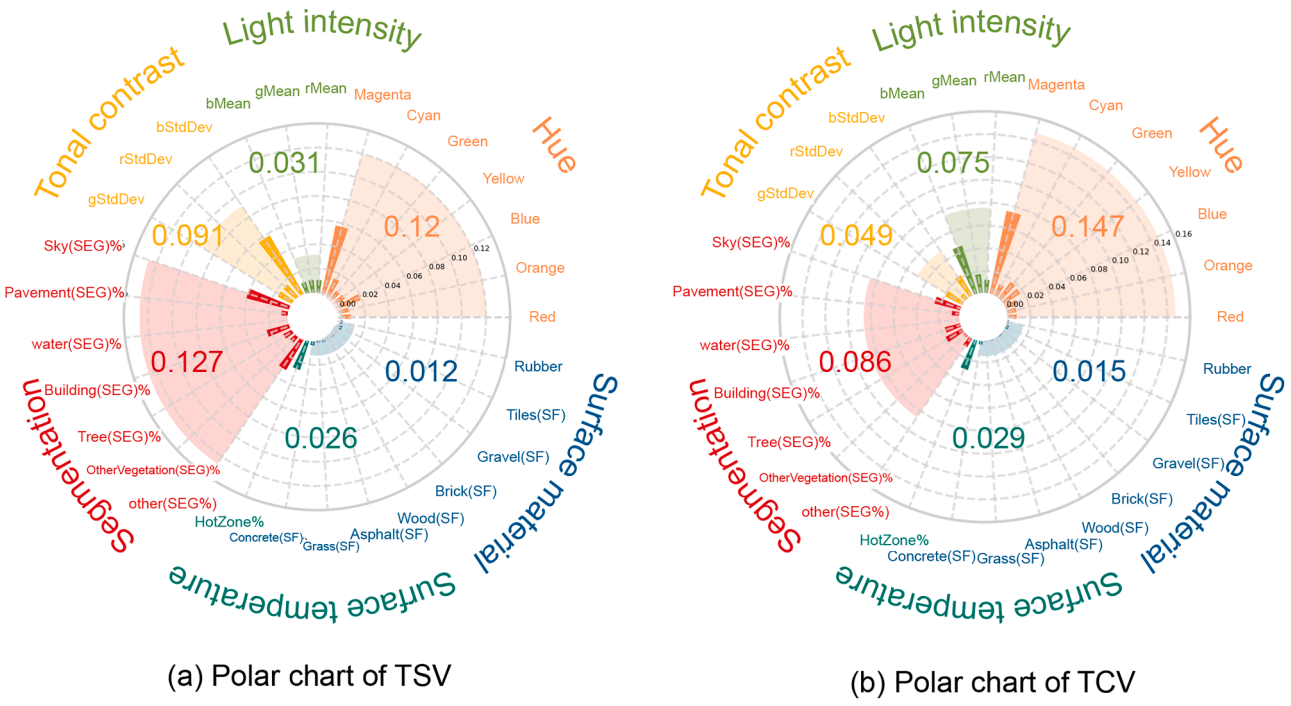


Fig. 9. Polar charts of TSV/TCV summarizing feature category influence.

Another notable finding is that, for 39 of the 50 images, the mean comfort vote in Group B was higher than in Group A. This is most likely explained by the different survey environments. Respondents in Group B who completed the survey in an air-conditioned indoor setting generally reported greater imagined comfort in the viewed scenes than those in Group A's outdoor setting.

However, 11 of the 50 images (highlighted in Fig. 11) were rated as less comfortable in Group B than in Group A. Taking a closer look at these images revealed some shared characteristics: most featured shade provision, lower sky-exposure ratios, lower *bStdDev* values, lower *Hot-Zone%* values, and higher magenta and building coverage. Interestingly, these features were previously identified as effective in enhancing thermal comfort in outdoor environments. However, this effect was not replicated in the indoor survey results. A possible explanation is that, within an air-conditioned indoor setting – where participants are already thermally comfortable and protected from direct environmental exposure – the perceived relevance of such features is diminished. For example, shading becomes functionally redundant indoors, and participants may instead shift their attention to other environmental cues that are more salient or meaningful in an indoor context.

Overall, the contrasting results between the two survey groups indicate that the survey setting (air-conditioned indoor vs. outdoor environment) influences participants' TCV. In most cases, TCV results collected indoors tend to reflect a higher level of perceived comfort compared to those collected outdoors. However, certain features identified as enhancing thermal comfort in outdoor voting results were associated with lower TCV ratings in the indoor setting. This suggests that although the study focuses solely on the influence of visual elements on thermal perception and all surveys were conducted based on image-based questions, the environmental conditions under which participants complete the survey can also affect the outcomes. Therefore, in this study, we chose to use the survey results collected in the outdoor setting as the training data, as the outdoor thermal conditions are more representative of the real environmental context depicted in the selected images. In most cases, TCV results collected indoors tend to reflect a higher level of perceived comfort compared to those collected outdoors. However, certain features identified as enhancing thermal comfort in outdoor voting results were associated with lower TCV ratings in the indoor setting. This suggests that although the study focuses solely on the influence of visual elements on thermal perception and all surveys were conducted based on image-based questions, the environmental conditions under which participants complete the survey can also affect the outcomes. Therefore, in this study, we chose to use the survey results collected in the outdoor setting as the training data, as the outdoor thermal conditions are more representative of the real environmental context depicted in the selected images.

It is important to note that this comparative analysis represents an opportunistic study conducted by the research team, made possible by the opportunity afforded through a public exhibition facilitated by the planning agency in Singapore. As a result, the survey environment and respondent demographics vary simultaneously. While this prevents the isolation of independent effects, the comparison nevertheless provides a valuable validity check by contrasting a controlled academic sample under some potential heat stress with a diverse public sample evaluated under thermally comfortable conditions.

4. Discussion

4.1. Influential visual features

This study utilized SHAP analysis to identify visual parameters significantly influencing TSV and TCV. *Magenta* consistently emerged as a prominent feature, strongly associated with perceptions of coolness and comfort. In contrast, blue-related visual features, including *Blue, bMean*, and *bStdDev*, were linked to warmer sensations and decreased comfort. Increased sky exposure (*Sky(SEG)%*) and *Hot Zone%* also demonstrated clear associations with hotter and more uncomfortable predictions. Conversely, higher values of *Other(SEG)%* predominantly corresponded to neutral perceptions, indicating minimal biased effects.

Notably, certain visual features exhibited differing (and divergent) impacts between TSV and TCV predictions. For example, *Building (SEG)%* showed context-dependent effects in the TCV model, associated with both comfort and discomfort depending on the context. In contrast, for TSV, this feature primarily predicted cooler sensations. Despite the strong correlation previously established between TSV and TCV through Pearson analysis, individual visual features did not consistently exert

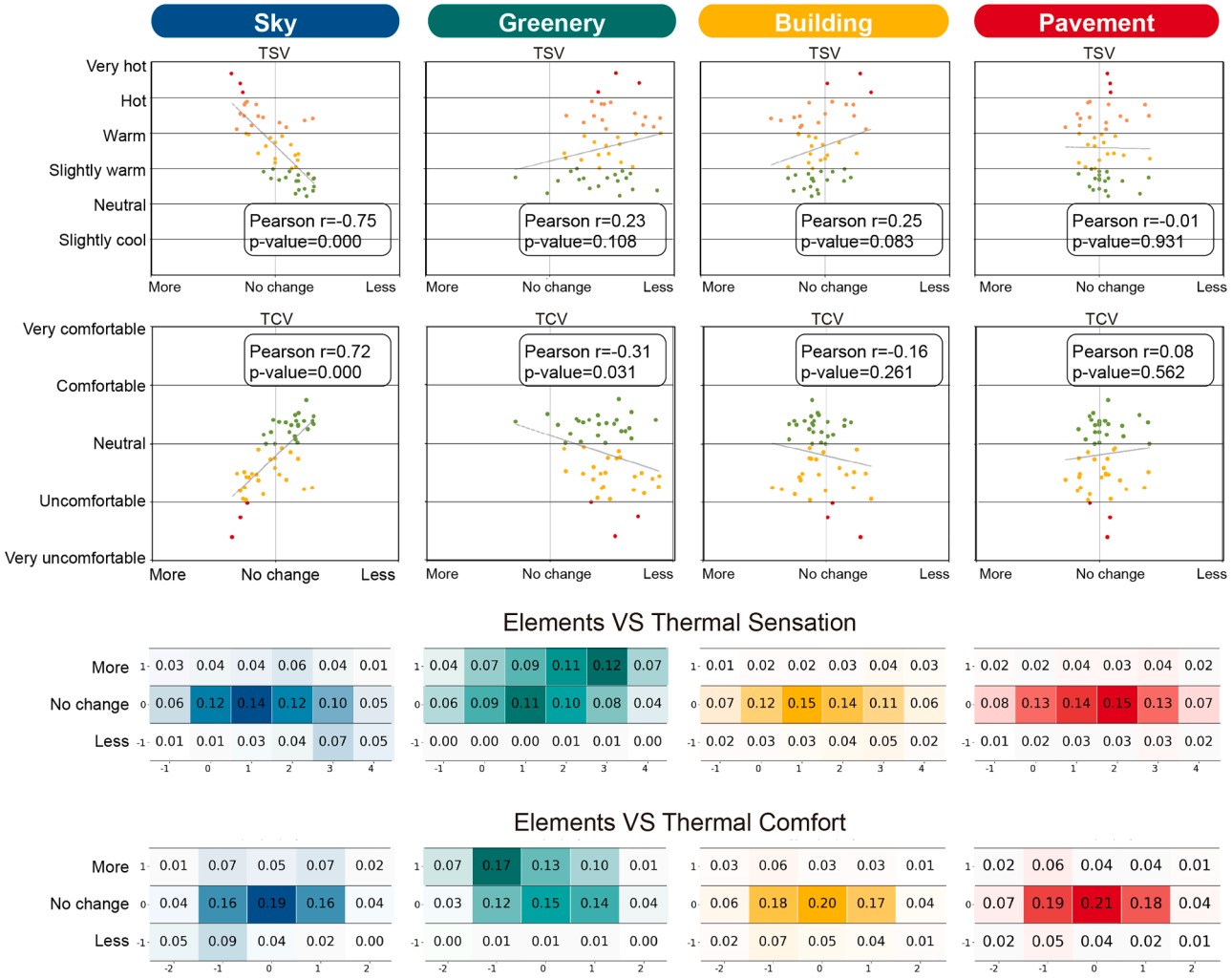


Fig. 10. Pearson correlation analyses and Heatmaps of Element perference.

	Original (Group A)	Exhibit (Group B)
Survey	Online survey	Interactive screen
Environment	Outdoor environment (warm or hot, varies per individual)	Indoor environment (cooled)
Respondent	Campus students (age 18 to 30)	Visitors (Diverse age group)

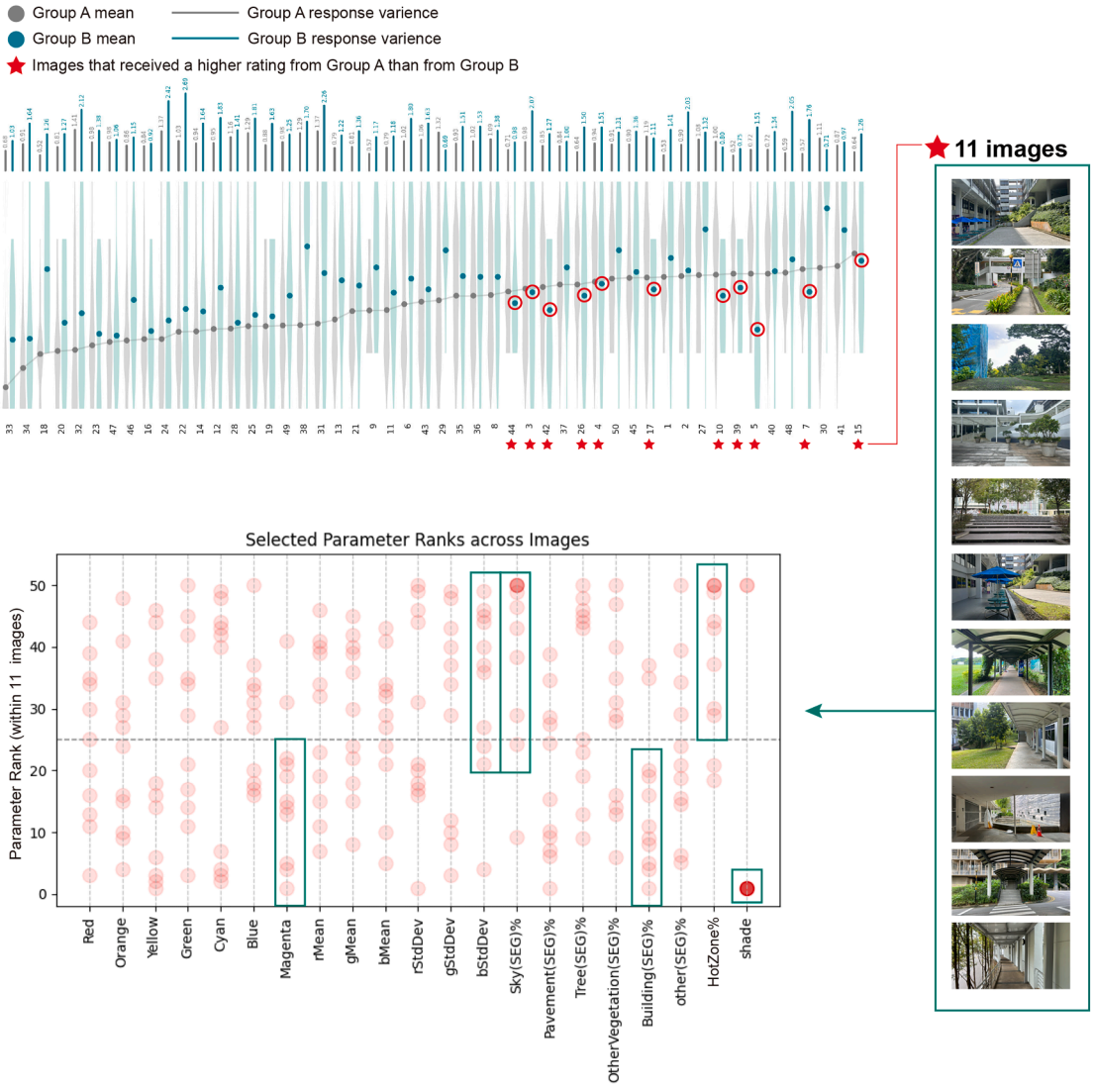


Fig. 11. TCV response comparison between Group A and Group B.

uniform directional influences across these two measures.

Contrary to common findings in indoor environments, where blue typically signals cooler conditions and comfort [12,46,47], our study found the opposite trend outdoors. Increased presence of blue was associated with warmer and less comfortable perceptions. This divergence may stem from contextual differences – particularly that in indoor settings, high blue values often correlate with skylight or vertical window exposure, which may contribute to a perceived cooler atmosphere. In contrast, in outdoor environments, increased blue may indicate direct sky exposure *(Sky(SEG)%)*, which is often linked to higher solar radiation and perceived heat.

Similarly, while magenta often has minimal influence (and is often considered a neutral or cool colour) indoors [46], our findings highlighted its significant role outdoors, suggesting complexity in how magenta influences thermal perception that merits further exploration. This unexpected prominence of magenta suggests that its role in shaping outdoor thermal perception is more complex and warrants further investigation.

4.2. Limitations

One limitation of the web-based survey modality is that participants' immediate local microclimatic conditions during survey completion could not be fully controlled. Although participants were instructed to complete the survey in an outdoor, sunny environment prior to participation, individual variations in environmental conditions may still exist.

It is important to note, however, that the primary objective of this study is not to estimate absolute thermal sensation under tightly controlled microclimatic conditions, but to investigate the relative influence of visual characteristics on perceived thermal sensation and comfort. Participants were therefore explicitly instructed to focus on the visual content of the images when providing TSV and TCV responses. Moreover, each image was evaluated by approximately 50 participants on average, and aggregating responses across a large sample helps reduce the influence of individual-level environmental noise, yielding more robust image-level perceptual estimates.

The survey was intentionally conducted in an outdoor context to enhance ecological validity. Supplementary experiments performed under a fully controlled indoor environment (at Singapore city gallery) indicate that visual thermal perception assessed indoors differs systematically from outdoor evaluations, further supporting the benefits of an outdoor survey setting despite the associated variability.

In this study, the experimental environment where each respondent conducted their web-based survey purposely remained constant across all images. Therefore, the physical environment in which the respondents answered the questionnaire was different from the physical environment in which the images were collected, yet Singapore's afternoon climate is generally stable and consistently warm. Unlike other studies that incorporated in-field data for validation [18], we deliberately chose not to adopt this approach for two main reasons. First, the primary objective of this study was to focus specifically on visual parameters as the core research subject. Therefore, we employed a web-based survey format to minimize the influence of external environmental factors. Second, collecting on-site data poses considerable challenges. In outdoor settings, participants' sensory perceptions are influenced by various physical conditions such as ambient noise, wind speed, and temperature fluctuations across different image collection locations. If not held constant, these environmental features would also interfere with our objective of isolating and independently analysing the effects of visual elements.

In subsequent research, we may consider integrating environmental factors to comprehensively examine the combined influence of visual and physical parameters on thermal perception. In our study, we found that the *HotZone%* indicator we collected, although not as influential as colour in shaping thermal perception, still exerted a measurable effect. Moreover, the limited size and visual diversity of the current image dataset – collected exclusively within campus and evaluated by a campus-based participant population – may constrain the generalizability of our findings beyond the campus context. To address this limitation, future studies could adopt a similar methodology but incorporate larger and more diverse datasets and validate the predictive models across different urban settings, climatic conditions, and population groups.

4.3. Insights for urban design

Analysis of element preference votes revealed that sky exposure*(Sky (SEG)%)* exhibited the strongest statistical correlation with thermal perceptions, but participants expressed relatively low preference for it as a cooling strategy. In contrast, greenery, despite weaker statistical correlations with TSV and TCV, received overwhelming public support, reflecting widespread belief in its physical cooling benefits [45]. Furthermore, in a recent study [20], which also employed an image-based method to investigate the restorative potential of urban landscapes, the findings critically questioned the intuitive notion of "the greener, the better" in relation to urban health benefits. Their findings suggest that more trees do not necessarily enhance comfort; instead, openness and the diversity of vegetation are also important factors influencing perceived comfort. This indicates that perceived outdoor thermal comfort relies not solely on physical cooling features but also on visual and psychological cues embedded in colour composition and visual scene complexity.

Urban design thus needs to integrate physical environmental performance with perceptual (subjective) and psychological considerations. Our findings suggest that perceived thermal comfort is not solely governed by elements that have been empirically and scientifically proven to provide physical cooling, such as greenery, but is also shaped by visual and psychological cues embedded in colour composition and image perception. In Singapore, several established cooling strategies exist, including measures related to material reflectance and vegetation planting [48]. For instance, the Singapore Government has started heat-reflective paint initiatives, applying cool coatings to the exterior of HDB buildings (public housing developed by the Housing and Development Board in Singapore) buildings to reduce surface temperatures [49], However, such cooling strategies, which primarily focus on modifying the physical environment, have yet to account for other dimensions such as visual influences. For instance, while cool paint coatings on roads can reduce surface temperatures [50], their high reflectance may also create glare under strong sunlight – visual perception shown to negatively affect perceived thermal comfort [8,9]. Consequently, urban thermal design must account for the distinction between physical environmental performance and perceptual outcomes. Other design-related theories also emphasize this point. For example, the Contemplative Landscape Model [51] underscores the pivotal role of visual perception in shaping psychological experiences in outdoor landscapes, with direct implications for human well-being and mental health. Accordingly, designing for outdoor comfort should consider how colour tones, visual complexity, and scene composition influence how people perceive and respond to thermal environments.

5. Conclusion

Our study advances the understanding of outdoor thermal comfort by integrating computer vision and (explainable) machine learning with image-based visual analysis and human perception data. The first objective was to extract a comprehensive set of visual parameters from images, encompassing six key dimensions*: hue, light intensity, tonal contrast, segmentation, surface temperature, and surface material,* then to identify the visual parameters most strongly associated with thermal perception. SHAP analysis identified magenta as the most influential visual feature predicting both thermal sensation and comfort, while

blue-related features and sky exposure (*Sky(SEG)%*) were linked to higher perceived heat and discomfort. These findings challenge traditional indoor-based colour assumptions and underscore the unique role of outdoor visual contexts. The second objective was to reveal how different visual elements relate to thermal perception. While sky exposure (*Sky(SEG)%*) showed the strongest statistical association with TSV and TCV, participants expressed a stronger preference for more greenery. Interestingly, greenery had less predictive power than colour features, suggesting a possible gap between perceived and actual thermal impact. This highlights the need to consider both psychological and physiological cues in urban comfort design. For the third objective, we successfully developed a framework that tightly integrates computer vision, standardized perceptual measures (TSV/TCV), and interpretable machine learning approaches. Specifically, we extracted 34 visual parameters using computer vision techniques and collected subjective thermal perception data through web-based TSV/TCV surveys. In terms of model training, we confirmed that thermal votes are not equidistant and that classification models – especially Random Forests with equal-frequency binning – outperform regression approaches. This supports the use of ordinal-aware methods in perceptual modelling. This framework offers potential for future larger-scale urban analysis using crowdsourced data, such as street-view imagery, which is increasingly available for pedestrian paths [16,19,26].

Finally, our study contributes a comprehensive framework for analysing outdoor thermal perception using visual parameters, introduces a novel image selection methodology leveraging computer vision and explainable machine learning, and provides new urban design insights that go beyond physical simulation by incorporating perceptual and visual cues. Our method is scalable and adaptable to different geographic or cultural contexts, requiring only that the model be retrained with images and survey data representative of the target environment and participant population. Such flexibility provides localized insights into thermal perception and enables the framework's application across diverse urban climates.

CRediT authorship contribution statement

Lujia Zhu: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Holly W. Samuelson:** Writing – review & editing, Supervision, Conceptualization. **Filip Biljecki:** Writing – review & editing, Supervision. **Chun Liang Tan:** Writing – review & editing, Supervision. **Nyuk Hien Wong:** Writing – review & editing, Supervision, Conceptualization. **Yu Qian Ang:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

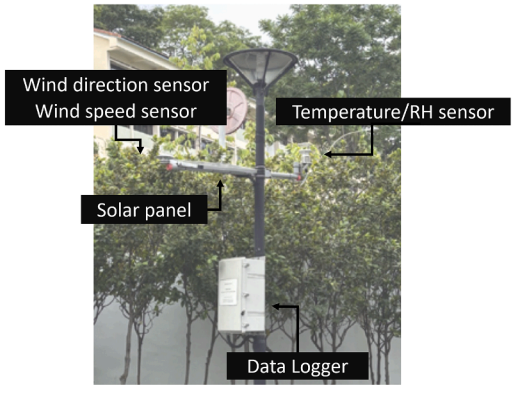
The material presented is based partly on work supported by the NUS Startup Grant and the CoolNUS-BEAM initiative. Any opinions, findings, conclusions, or recommendations expressed in the material are those of the author(s) and do not reflect the views of the grantors. The authors thank Alex Yeung, Kai Wen Wong, Zhongyu Chiam, and Yixuan Tan from the Singapore Urban Redevelopment Authority (URA) for their assistance in setting up the exhibition, and Lester Ong for developing the application for the interactive exhibit.

Appendix A

(Image from FLIR product website)

Camera set

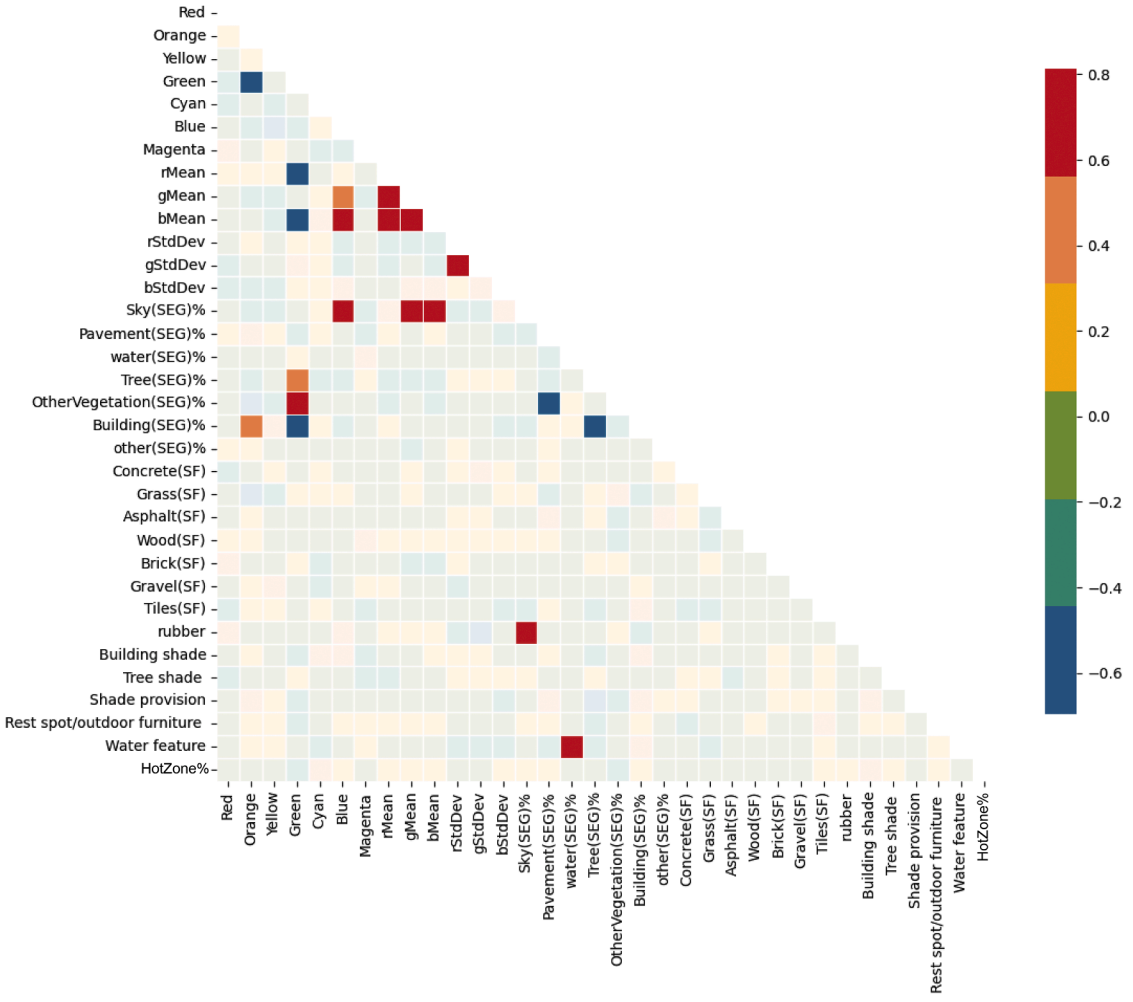
Model	Measurement
Flir One® Pro	Thermal image
Smart phone	Photographic image



Weather station

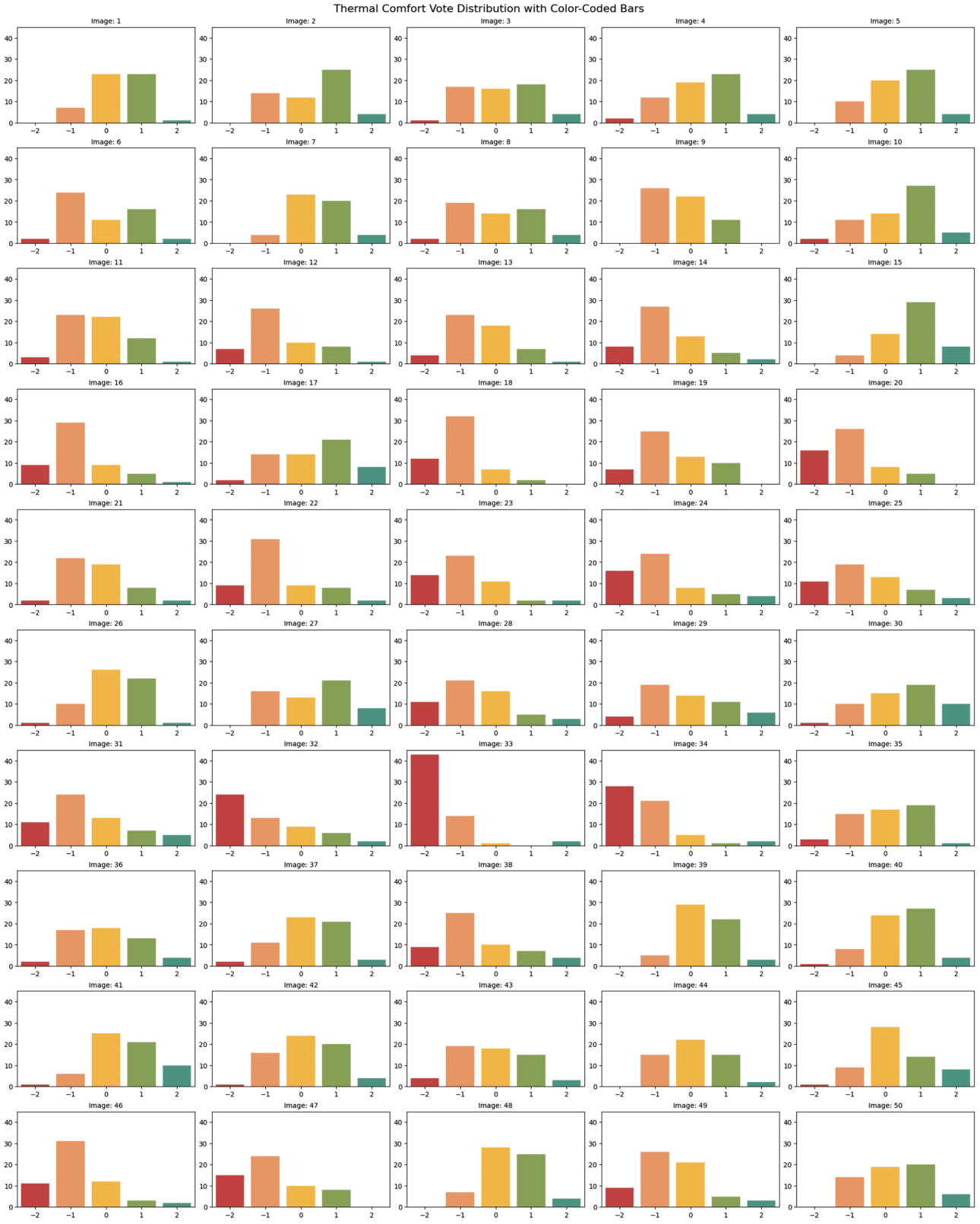
Model	Measurement
Wind speed/direction sensor	Wind speed/direction
Temperature/RH sensor	temperature/ relative humidity

Appendix A1. Details of measurement instruments.



Appendix A2. Pairwise correlation matrix of the selected 34 visual features.

Appendix A3. Distribution of TSV vote per image.



Appendix A4. Distribution of TCV vote per image.

Data availability

Data will be made available on request.

References

[1] S. Jia, Y. Wang, N.H. Wong, Q. Weng, A hybrid framework for assessing outdoor thermal comfort in large-scale urban environments, Landsc. Urban. Plan. 256 (2025) 105281, https://doi.org/10.1016/j.landurbplan.2024.105281.

[2] S. Chen, N.H. Wong, W. Zhang, M. Ignatius, The impact of urban morphology on the spatiotemporal dimension of estate-level air temperature: a case study in the tropics, Build. Environ. 228 (2023) 109843, https://doi.org/10.1016/j.buildenv.2022.109843.

[3] L. Chen, E. Ng, Outdoor thermal comfort and outdoor activities: a review of research in the past decade, Cities 29 (2012) 118–125, https://doi.org/10.1016/j.cities.2011.08.006.

[4] K. Kavee, K.A. Flanigan, Encoding experience: quantifying multisensory perception of urban form through a systematic review, Comput. Environ. Urban. Syst. 122 (2025) 102349, https://doi.org/10.1016/j.compenvurbsys.2025.102349.

[5] M. Schweiker, G.M. Huebner, B.R.M. Kingma, Kramer, H. Rick, Pallubinsky, Drivers of diversity in human thermal perception – a review for holistic comfort models, Temperature 5 (2018) 308–342, https://doi.org/10.1080/23328940.2018.1534490.

[6] 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, Build. Environ. 186 (2020) 107297, https://doi.org/10.1016/j.buildenv.2020.107297.

[7] M. Du, B. Hong, C. Gu, Y. Li, Y. Wang, Multiple effects of visual-acoustic-thermal perceptions on the overall comfort of elderly adults in residential outdoor environments, Energy Build. 283 (2023) 112813, https://doi.org/10.1016/j.enbuild.2023.112813.

[8] F. Rosso, A.L. Pisello, F. Cotana, M. Ferrero, On the thermal and visual pedestrians' perception about cool natural stones for urban paving: a field survey in summer conditions, Build. Environ. 107 (2016) 198–214, https://doi.org/10.1016/j.buildenv.2016.07.028.

[9] L. Samira, S. Abdou, S. Reiter, Effect of vegetation cover on thermal and visual comfort of pedestrians in urban spaces in hot and dry climate, (2017).

[10] J. Yang, Z.-H. Wang, K.E. Kaloush, H. Dylla, Effect of pavement thermal properties on mitigating urban heat islands: a multi-scale modeling case study in Phoenix, Build. Environ. 108 (2016) 110–121, https://doi.org/10.1016/j.buildenv.2016.08.021.

[11] J. Huang, X. Tang, P. Jones, T. Hao, R. Tundokova, C. Walmsley, S. Lannon, P. Frost, J. Jackson, Mapping pedestrian heat stress in current and future heatwaves in Cardiff, Newport, and Wrexham in Wales, UK, Build. Environ. 251 (2024) 111168, https://doi.org/10.1016/j.buildenv.2024.111168.

[12] L. Zhang, X. Li, C. Li, T. Zhang, Research on visual comfort of color environment based on the eye-tracking method in subway space, J. Build. Eng. 59 (2022) 105138, https://doi.org/10.1016/j.jobe.2022.105138.

[13] A.B. Avci, G.A. Balci, T. Basaran, Optimizing thermal comfort in physical exercise spaces: a study of spatial and thermal factors, Energy Build. 303 (2024) 113782, https://doi.org/10.1016/j.enbuild.2023.113782.

[14] S.Y. Chan, C.K. Chau, T.M. Leung, On the study of thermal comfort and perceptions of environmental features in urban parks: a structural equation modeling approach, Build. Environ. 122 (2017) 171–183, https://doi.org/10.1016/j.buildenv.2017.06.014.

[15] J. Wen, D.A. Abuhani, M. Mazzarello, F. Duarte, L. Norford, R. Xu, N.H. Wong, C. Ratti, Walking smart in the heat: a dynamic shade-oriented pathfinding approach to enhance pedestrian comfort in arid cities, Comput. Environ. Urban. Syst. 122 (2025) 102337, https://doi.org/10.1016/j.compenvurbsys.2025.102337.

[16] 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, Int. J. Appl. Earth Observ. Geoinf. 122 (2023) 103385, https://doi.org/10.1016/j.jag.2023.103385.

[17] H. Zhu, F. Yang, Z. Bao, X. Nan, A study on the impact of visible Green Index and vegetation structures on brain wave change in residential landscape, Urban. For. Urban. Green. 64 (2021) 127299, https://doi.org/10.1016/j.ufug.2021.127299.

[18] S. Yang, A. Chong, P. Liu, F. Biljecki, Thermal comfort in sight: thermal affordance and its visual assessment for sustainable streetscape design, Build. Environ. 271 (2025) 112569, https://doi.org/10.1016/j.buildenv.2025.112569.

[19] Z. Wang, K. Ito, F. Biljecki, Assessing the equity and evolution of urban visual perceptual quality with time series street view imagery, Cities 145 (2024) 104704, https://doi.org/10.1016/j.cities.2023.104704.

[20] X. Zhang, E.S. Lin, P.Y. Tan, J. Qi, R. Ho, A. Sia, R. Waykool, X.P. Song, A. Olszewska-Guizzo, L. Meng, Y. Cao, Beyond just green: explaining and predicting restorative potential of urban landscapes using panorama-based metrics, Landsc. Urban. Plan. 247 (2024) 105044, https://doi.org/10.1016/j.landurbplan.2024.105044.

[21] M. Vellei, R. de Dear, C. Inard, O. Jay, Dynamic thermal perception: a review and agenda for future experimental research, Build. Environ. 205 (2021) 108269, https://doi.org/10.1016/j.buildenv.2021.108269.

[22] C. Cen, E. Tan, S. Valliappan, E. Ang, Z. Chen, N.H. Wong, Students' thermal comfort and cognitive performance in tropical climates: a comparative study, Energy Build. 341 (2025) 115817, https://doi.org/10.1016/j.enbuild.2025.115817.

[23] V. Gnecco, I. Pigliautile, A. Pisello, Long-term thermal comfort monitoring using wearable devices, in: 2022 IEEE International Workshop on Metrology for Living Environment (MetroLivEn), 2022, pp. 13–17, https://doi.org/10.1109/metrolivenv54405.2022.9826920.

[24] M.T. Ribeiro, S. Singh, C. Guestrin, "Why should I trust you?": explaining the predictions of any classifier, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, San Francisco California USA, 2016, pp. 1135–1144, https://doi.org/10.1145/2939672.2939778.

[25] K. Fujiwara, M. Khomiakov, W. Yap, M. Ignatius, F. Biljecki, Microclimate Vision: multimodal prediction of climatic parameters using street-level and satellite imagery, Sustain. Cities. Soc. 114 (2024) 105733, https://doi.org/10.1016/j.scs.2024.105733.

[26] K. Ito, Y. Zhu, M. Abdelrahman, X. Liang, Z. Fan, Y. Hou, T. Zhao, R. Ma, K. Fujiwara, J. Ouyang, M. Quintana, F. Biljecki, ZenSVI: an open-source software for the integrated acquisition, processing and analysis of street view imagery towards scalable urban science, Comput. Environ. Urban. Syst. 119 (2025) 102283, https://doi.org/10.1016/j.compenvurbsys.2025.102283.

[27] M.J. Hilz, S. Glorius, A. Berić, Thermal perception thresholds: influence of determination paradigm and reference temperature, J. Neurol. Sci. 129 (1995) 135–140, https://doi.org/10.1016/0022-510x(94)00262-m.

[28] S. Sharma, A. Singh, R. Kumar, A. Mehta, Analysis of land surface temperature at 10-meter pixel size (Spatial Resolution) for Ahmedabad City, in: 2023 International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES), 2023, pp. 134–138, https://doi.org/10.1109/CISES58720.2023.10183518.

[29] K. Chen, J. Wang, J. Pang, Y. Cao, Y. Xiong, X. Li, S. Sun, W. Feng, Z. Liu, J. Xu, Z. Zhang, D. Cheng, C. Zhu, T. Cheng, Q. Zhao, B. Li, X. Lu, R. Zhu, Y. Wu, J. Dai, J. Wang, J. Shi, W. Ouyang, C.C. Loy, D. Lin, MMDetection: open MMLab detection toolbox and benchmark, (2019). https://doi.org/10.48550/arXiv.1906.07155.

[30] D. Jonauskaite, C. Mohr, J.-P. Antonietti, P.M. Spiers, B. Althaus, S. Anil, N. Dael, Most and least preferred colours differ according to object context: new insights from an unrestricted colour range, PLoS ONE 11 (2016) e0152194, https://doi.org/10.1371/journal.pone.0152194.

[31] Y. Gu, M. Quintana, X. Liang, K. Ito, W. Yap, F. Biljecki, Designing effective image-based surveys for urban visual perception, Landsc. Urban. Plan. 260 (2025) 105368, https://doi.org/10.1016/j.landurbplan.2025.105368.

[32] iTeh Inc, ISO 10551:2019 - Ergonomics subjective scales for physical environments, iTeh standards (n.d.). https://standards.iteh.ai/catalog/standards/iso/306c74da-8dce-44f8-9bcc-7230eeaf26e0/iso-10551-2019 (Accessed 9 January 2026).

[33] ANSI/ASHRAE Addendum d to ANSI/ASHRAE Standard 55-2017, (n.d.).

[34] W. Yang, N.H. Wong, S.K. Jusuf, Thermal comfort in outdoor urban spaces in Singapore, Build. Environ. 59 (2013) 426–435, https://doi.org/10.1016/j.buildenv.2012.09.008.

[35] W. Yang, N.H. Wong, G. Zhang, A comparative analysis of human thermal conditions in outdoor urban spaces in the summer season in Singapore and Changsha, China, Int. J. Biometeorol. 57 (2013) 895–907, https://doi.org/10.1007/s00484-012-0616-9.

[36] X. Zhan, D. Liang, X. Lin, L. Li, C. Wei, C. Dingle, Y. Liu, Background noise but not urbanization level impacted song frequencies in an urban songbird in the Pearl River Delta, Southern China, Glob. Ecol. Conserv. 28 (2021) e01695, https://doi.org/10.1016/j.gecco.2021.e01695.

[37] E. Isaac, K.S. Easwarakumar, J. Isaac, Urban landcover classification from multispectral image data using optimized AdaBoosted random forests, Remote Sens. Lett. 8 (2017) 350–359, https://doi.org/10.1080/2150704X.2016.1274443.

[38] L. Xiong, Y. Yao, Study on an adaptive thermal comfort model with K-nearest-neighbors (KNN) algorithm, Build. Environ. 202 (2021) 108026, https://doi.org/10.1016/j.buildenv.2021.108026.

[39] L. Lin, Y. Zhao, J. Zhao, Optimizing urban green space spatial patterns for thermal environment improvement: a multi-objective approach in the context of urban renewal, Comput. Environ. Urban. Syst. 121 (2025) 102320, https://doi.org/10.1016/j.compenvurbsys.2025.102320.

[40] S. Yi, X. Li, D. Li, X. Dong, R. Wang, Q. Xu, Hyperlocal heat stress around bus stops in Philadelphia: insights from spatio-temporal microclimate modeling and explainable AI, Comput. Environ. Urban. Syst. 122 (2025) 102341, https://doi.org/10.1016/j.compenvurbsys.2025.102341.

[41] N. Chen, L. Wang, T. Xu, M. Wang, Perception of urban street visual color environment based on the CEP-KASS framework, Landsc. Urban. Plan. 259 (2025) 105359, https://doi.org/10.1016/j.landurbplan.2025.105359.

[42] X. Xue, Z. Tian, Y. Yang, J. Wang, S.-J. Cao, Sustaining the local color of a global city, Nat. Cities 2 (2025) 400–412, https://doi.org/10.1038/s44284-025-00225-x.

[43] M. Ding, Quantitative contrast of urban agglomeration colors based on image clustering algorithm: case study of the Xia-Zhang-Quan metropolitan area, Front. Architect. Res. 10 (2021) 692–700, https://doi.org/10.1016/j.foar.2021.05.003.

[44] Y. Fan, Understanding Thermal Comfort and Thermal Adaptation in Singapore, 2021. https://scholarbank.nus.edu.sg/handle/10635/201661 (Accessed 9 January 2026).

[45] N.H. Wong, C.L. Tan, D.D. Kolokotsa, H. Takebayashi, Greenery as a mitigation and adaptation strategy to urban heat, Nat. Rev. Earth. Environ. 2 (2021) 166–181, https://doi.org/10.1038/s43017-020-00129-5.

[46] S. Hu, C. Liu, Experimental investigation about thermal effect of colour on thermal sensation and comfort, Energy Build. 173 (2018) 710–718, https://doi.org/10.1016/j.enbuild.2018.06.008.

[47] H. Qin, J. Chen, J. Niu, J. Huo, X. Wei, J. Yan, G. Han, The effects of brightness and prominent colors on outdoor thermal perception in Chongqing, China, Int. J. Biometeorol. 68 (2024) 1143–1154, https://doi.org/10.1007/s00484-024-02654-0.

[48] L. Ruefenacht, J.A. Acero, Strategies for Cooling Singapore: A Catalogue of 80+ Measures to Mitigate Urban Heat Island and Improve Outdoor Thermal Comfort, 2017, https://doi.org/10.3929/ethz-b-000258216.

[49] Heat-Reflective Paint Initiative to be Rolled Out to All HDB Estates by 2030, The Straits Times, 2025. https://www.straitstimes.com/singapore/housing/heat-reflective-paint-initiative-to-be-rolled-out-to-all-hdb-estates-by-2030 (Accessed 15 July 2025).

[50] Cool Paint Coatings Help City Dwellers Feel Up to 1.5 degrees Celsius Cooler, Study Finds, Corporate NTU (n.d.). https://www.ntu.edu.sg/news/detail/cool-paint-coatings-help-city-dwellers-feel-up-to-1.5-degrees-celsius-cooler-study-finds (Accessed 15 July 2025).

[51] A. Olszewska-Guizzo, A. Sia, N. Escoffier, Revised Contemplative Landscape Model (CLM): a reliable and valid evaluation tool for mental health-promoting urban green spaces, Urban. For. Urban. Green. 86 (2023) 128016, https://doi.org/10.1016/j.ufug.2023.128016.