land

MDPI

Article

Decoupling Urban Street Attractiveness: An Ensemble Learning Analysis of Color and Visual Element Contributions

Tao Wu [1,†], Zeyin Chen [1,†], Siying Li [2], Peixue Xing [3], Ruhang Wei [1], Xi Meng [4], Jingkai Zhao [5,*], Zhiqiang Wu [1,6,7,*] and Renlu Qiao [5]

[1] College of Architecture and Urban Planning, Tongji University, 1239 Siping Road, Shanghai 200092, China; taowu@tongji.edu.cn (T.W.); zet@tongji.edu.cn (Z.C.); weiruhang@tjad.cn (R.W.)
[2] School of Architecture, Tsinghua University, 30 Shuangqing Road, Beijing 100084, China; lisiying22@mails.tsinghua.edu.cn
[3] School of Environmental Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China; px_xing@sjtu.edu.cn
[4] Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing University of Technology, Beijing 100124, China; mengxi@bjut.edu.cn
[5] Shanghai Research Institute for Intelligent Autonomous Systems, Tongji University, 1239 Siping Road, Shanghai 200092, China; qiaorl@tju.edu.cn
[6] Department of Mathematics and Theories, Peng Cheng Laboratory, Shenzhen 518066, China
[7] Chinese Academy of Engineering, Beijing 100094, China
* Correspondence: jingkaizhao@tongji.edu.cn (J.Z.); wus@tongji.edu.cn (Z.W.)
† These authors contributed equally to this work.

Abstract: Constructing visually appealing public spaces has become an important issue in contemporary urban renewal and design. Existing studies mostly focus on single dimensions (e.g., vegetation ratio), lacking a large-scale integrated analysis of urban color and visual elements. To address this gap, this study employs semantic segmentation and color computation on a massive street-view image dataset encompassing 56 cities worldwide, comparing eight machine learning models in predicting Visual Aesthetic Perception Scores (VAPSs). The results indicate that LightGBM achieves the best overall performance. To unpack this "black-box" prediction, we adopt an interpretable ensemble approach by combining LightGBM with Shapley Additive Explanations (SHAPs). SHAP assigns each feature a quantitative contribution to the model's output, enabling transparent, post hoc explanations of how individual color metrics and visual elements drive VAPS. Our findings suggest that the vegetation ratio contributes the most to VAPS, but once greening surpasses a certain threshold, a "saturation effect" emerges and can no longer continuously enhance visual appeal. Excessive Sky Visibility Ratio can reduce VAPS. Moderate road visibility may increase spatial layering and vibrancy, whereas overly dense building significantly degrades overall aesthetic quality. While keeping the dominant color focused, moderate color saturation and complexity can increase the attractiveness of street views more effectively than overly uniform color schemes. Our research not only offers a comprehensve quantitative basis for urban visual aesthetics, but also underscores the importance of balancing color composition and visual elements, offering practical recommendations for public space planning, design, and color configuration.

Keywords: urban aesthetics; urban color; street-view visual elements; machine learning; Place Pulse 2.0

check for updates

Academic Editor: Scott Hawken

Received: 18 March 2025
Revised: 22 April 2025
Accepted: 28 April 2025
Published: 1 May 2025

Citation: Wu, T.; Chen, Z.; Li, S.; Xing, P.; Wei, R.; Meng, X.; Zhao, J.; Wu, Z.; Qiao, R. Decoupling Urban Street Attractiveness: An Ensemble Learning Analysis of Color and Visual Element Contributions. *Land* **2025**, *14*, 979. https://doi.org/10.3390/land14050979

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Over the past years, the demand for urban spaces has transitioned from merely pursuing quantitative expansion to prioritizing high standards and superior quality, thereby greatly elevating the importance of urban spatial environmental quality within sustainable urban development strategies [1,2]. Urban color and visual elements together constitute a crucial representation of the urban spatial environment. Among these, urban color plays a crucial role in shaping the overall quality of urban spaces [3], making it imperative to accurately assess the quality of the urban color environment [4]. However, despite this acknowledged importance, many studies have treated color chiefly as an adjunct to other visual features, thereby underestimating its standalone influence on residents' aesthetic and behavioral responses [5,6]. Such assessments not only provide guidance for urban color planning, shaping city identity and culture [7,8], but also help determine whether the quality of urban color satisfies the residents' psychological needs [4,9]. Moreover, the design and characteristics of visual elements in cities can stimulate positive emotions and create favorable sensory experiences [10], as well as directly influence how residents interact and behave in urban settings [11]. Consequently, these are the focus of current study. For instance, natural elements such as sky visibility ratio and vegetation ratio play a vital role in shaping urban spatial quality [12]. Studies grounded in environmental psychology have demonstrated their impact on human perception and physical well-being [13]. Similarly, research has shown that elements of the natural environment can elicit human aesthetic and emotional responses [14]. Positive perceptual feedback from natural surroundings can enhance individuals' inner well-being, which may have restorative effects for patients [15]. Furthermore, electroencephalogram (EEG) experiments have offered additional evidence of how visual elements influence internal human responses [16]. Therefore, comprehensively evaluating urban color and visual elements becomes especially important, and introducing a quantitative model and framework to assess how people perceive different colors and elements in cities and their communities is of considerable significance [17].

Street View Image (SVI) has become a novel tool for mapping and understanding urban environments, contributing significantly to urban research and spatial data analysis. With the expansion of map services provided by platforms like Google [18], Tencent [19], and Baidu [20], the application of SVI in urban studies has steadily increased. As a key element of the built environment, street views hold significant value in environmental perception, transportation planning, public health, and aesthetic design, thus attracting growing attention [21–25]. Consequently, numerous studies have been carried out to evaluate urban visual perception at the human scale in relation to SVI [26–28]. Technically, the integration of artificial intelligence and computing advancements, along with the increasing accessibility of large-scale SVI and machine learning techniques, has fostered new analytical approaches for understanding and interpreting color and visual elements. Various urban studies have employed diverse modeling techniques to effectively link urban color and visual element features with the psychological perception of the population [29]. For instance, machine learning including deep learning [30,31] has been utilized to mine large datasets and investigate the relationship between customer satisfaction and the quality of urban color and visual elements [32]. These methods excel in classifying urban color and visual elements, identifying image scene features, conducting street classification and mapping, and performing semantic segmentation and information extraction. Consequently, related research has attracted considerable attention and has been successfully implemented [2,26,33–36]. Against this background, many researchers have proposed concepts and indicator frameworks related to street-view perception, reorganizing and classifying key visual elements through semantic segmentation and information extraction. Current primary measures include openness [17,37], greenness [38], enclosure [39], and

walkability [40]. The indicators above play a positive role in deepening understanding of urban scenes and providing practical guidance.

The algorithms currently used in research exhibit significant differences in their ability to model nonlinear relationships, resist noise, and scale to large datasets, yet are rarely compared systematically, which hinders cross-study synthesis. In regression tasks, machine learning methods commonly fall into four categories [41]. Tree-based models (e.g., Decision Tree, Random Forest, XGBoost, LightGBM, CatBoost) capture feature interactions via hierarchical splits [42–45]; ensemble variants such as RF, XGBoost, LightGBM, and CatBoost improve fit and generalization through bagging or boosting, with XGBoost and LightGBM exploiting efficient gradient-boosting frameworks for rapid training and high accuracy on large data—LightGBM being especially adept at sparse feature handling—and CatBoost using ordered boosting to reduce categorical bias and overfitting. Distance—and kernel-based methods (KNN, SVM) address high-dimensional nonlinearity via nearest-neighbor assumptions and kernel transformations [46,47], respectively: KNN requires no parametric assumptions but is sensitive to noise, while SVM offers stable performance in high dimensions at the cost of greater computational expense. Neural networks (Multi-Layer Perceptron) are powerful for modeling complex nonlinear patterns but demand extensive hyperparameter tuning and longer training times [48]. Finally, the Decision Tree provides a simple, interpretable baseline [49]. All of these models have demonstrated excellent performance in regression applications such as urban environmental quality assessment and thus represent mainstream choices for predicting street-view aesthetic perception. A comprehensive comparison is therefore essential to identify the optimal algorithm for decoupling the contributions of color and visual elements. Additionally, these models can be used in conjunction with interpretable techniques such as Shapley Additive Explanations (SHAPs) to reveal the contribution of individual visual features to model predictions, improving the transparency and applicability of the analysis [50,51].

However, the current literature exhibits two main shortcomings. First, many studies have treated urban color merely as an adjunct to other visual features, thereby underestimating its standalone influence, and have seldom examined how color and visual elements interact synergistically [6,52]. Second, empirical efforts to quantify urban color perception have largely relied on small-scale, traditional survey methods, resulting in limited generalizability and challenges in standardization. In terms of urban color, current survey approaches are primarily traditional field investigations [53]. These methods commonly involve manual photography and computation [4], color card comparisons [54], and instrument-based color measurements. Although these approaches offer high accuracy, they come with significant costs, require substantial time, and are constrained to relatively small urban regions, posing challenges for large-scale implementation [28]. Moreover, manual photography is affected by equipment parameters and weather conditions, making it difficult to standardize color capture. With respect to urban color perception, most studies use only simple questionnaire data and lack systematic, standardized research on urban color quality [4]. Although some studies have begun to use large-scale street-view data for evaluation and computation [23,55], further improvements in data volume and comprehensiveness of analytical mechanisms are needed. Regarding visual elements, most research relies on surveys, in-person interviews, and field observations to examine how individuals perceive and interact with the built environment on both visual and sensory levels [56]. However, the recorded outcomes tend to be abstract rather than quantitative [17], introducing the possibility of substantial individual variability that can affect accuracy [17,57].

Based on this, the research integrates the color and urban visual environment elements, uniformly considering people's perception of them and conducting standardized judgments

to analyze what kind of color and visual elements have better visual aesthetics. Applying Place Pulse 2.0 SVI combined with machine learning models, we first perform pixel-level semantic segmentation to identify and classify color and visual elements from over 100,000 SVIs across 56 major cities worldwide. Simultaneously, the TrueSkill algorithm is employed to compute the Visual Aesthetic Perception Scores (VAPSs) for these images. Subsequently, we compare regression models using eight mainstream machine learning approaches and select the best-performing one. Through an analysis based on explainable machine learning, we clearly demonstrate the specific effects of different color and visual features on the prediction of the model. This technical method enables a comprehensive analysis of multiple perception indicators, a high-resolution interpretation, and a refined evaluation of street-view visual features. In doing so, this study not only reveals the intrinsic aesthetic mechanisms underlying street-view perception but also offers important guidance for urban planning and management in the context of street revitalization and urban renewal.

2. Data and Method

2.1. Study Area

The study area comprises multiple globally distributed urban environments, selected based on the availability of street-level imagery data. These cities span diverse geographic regions and cultural contexts, ensuring a broad spatial representation suitable for analyzing urban street views and their perceptual attributes. Given the extensive coverage of urban landscapes across different continents, the selected study area enhances the generalizability of the findings. Specifically, the study covers major cities across Asia, the Americas, Europe, Oceania, and Africa, including Hong Kong, New York, São Paulo, Paris, Melbourne, and Cape Town, among a total of 56 cities. Figure 1 shows the distribution of the study area in detail.

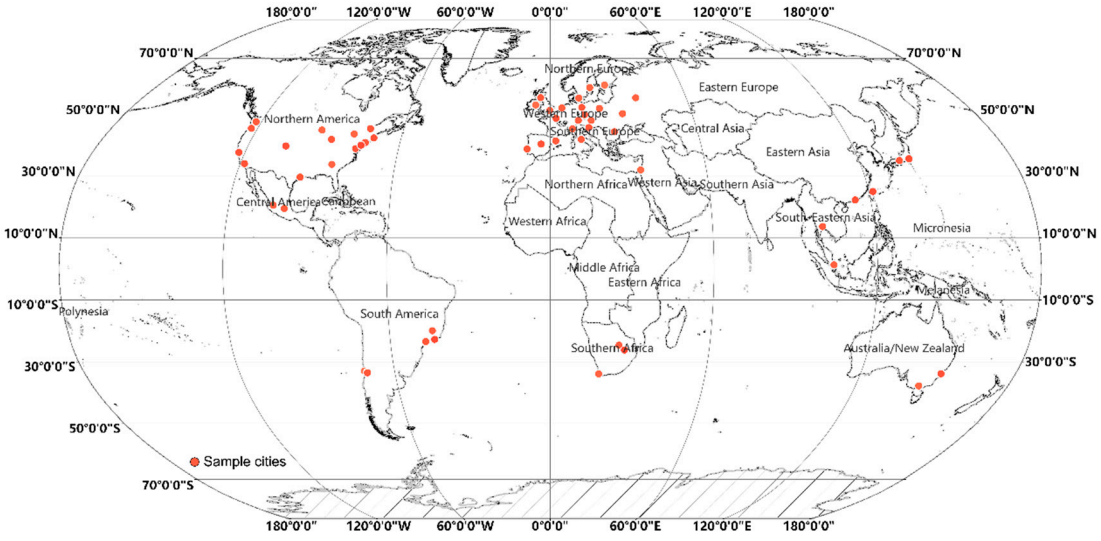

Figure 1. Distribution of the study area.

2.2. Description of Data Sources

This study utilizes data from the urban street-view perception dataset Place Pulse 2.0 [58]. The dataset contains a total of 110,988 Google SVIs captured between 2007 and 2012 in 56 cities on six continents, thus covering a wide range of geographic settings. Crucially, Place Pulse is built on a global, crowdsourced rating platform rather than locally confined surveys: participants from many countries were recruited through organic media outreach and targeted Facebook advertisements, and all images were evaluated on the same web

interface (https://centerforcollectivelearning.org/urbanperception; accessed on 4 August 2024). Because raters are not limited to the residents of the depicted cities, the resulting scores reflect a more universal aesthetic judgement and minimize city-specific cultural bias, even though the dataset itself does not include explicit sociocultural variables for each city.

Street-view frames were generated following the standard Place Pulse protocol. First, Google Street View panorama IDs were uniformly sampled along the OpenStreetMap road network at an adaptive spacing of roughly 50–100 m, ensuring coverage of both primary and secondary streets. For each panorama, two horizontal images (640 × 480 px, FOV ≈ 60°) were extracted with headings separated by 90° or 180°, while keeping the pitch at 0° (eye-level ≈ 1.6 m). In intersections or irregular street segments, up to four directions were captured to reflect the complete surrounding context [59]. This fixed sampling interval and limited set of viewing directions provide consistent spatial density and comparable visual perspectives across all 56 cities.

Place Pulse uses pairwise comparison—a method long employed to assess subjective attributes such as style or visual appeal in clothing [60], urban façades [22,61], animated GIFs [62], and artworks [63]. Pairwise ranking is widely regarded as more reliable and efficient than direct numerical scoring [64,65].

Participants were recruited via organic media sources and targeted Facebook advertisements, and were asked to answer subjective questions across six dimensions—for instance, "Which place looks safer?" or "Which place looks more beautiful?"—by selecting one of two images. This data collection process ran from May 2013 to February 2016. In this study, we mainly focus on responses to "Which place looks more beautiful?", for which 166,823 pairwise comparison responses were collected. Every individual image underwent an average of approximately 3.46 pairwise comparisons. Place Pulse 1.0 shows that ratings are largely independent of respondents' age, gender, or geographic location [59]; hence, the dataset offers a culturally diverse yet methodologically uniform benchmark.

Because evaluations are made on static images rather than on-site visits, extraneous local variables (e.g., transient noise, weather, or social activity) exert relatively little influence on the scores. Consequently, the derived VAPSs represent a consistent, image-based measure of perceived beauty that can be compared across all 56 cities without the confounding effects inherent in localized, in-person surveys.

2.3. Research Methods Process

Figure 2 illustrates the overall technical process of this study. First, we obtained SVIs from 56 major cities worldwide and their pairwise comparison results via the Place Pulse 2.0 platform. Based on the TrueSkill algorithm, we use VAPS to score these images, resulting in a large-scale and widely distributed research sample. Next, using semantic segmentation, we performed pixel-level recognition of roads, buildings, greenery, sky, and other primary visual elements, extracting both the presence and proportion of these elements. Meanwhile, the SVIs were transformed to HSV and RGB color spaces to calculate key indicators. Further, we extracted color composition metrics including Color Complexity (CC), Color Harmony (CH), and Dominant Color Ratio (DCR), thereby quantifying urban street views in terms of both visual composition and color features.

We integrated these visual element features and color features into a single dataset and employed various machine learning regression models to predict and analyze the VAPSs of street views. We assessed each model's predictive accuracy and generalization capability using performance metrics and training time, ultimately selecting the optimal model for prediction and evaluation.

Finally, to gain a clearer understanding of each feature's contribution to the model's predictions and their interrelationships, we incorporated an explainable machine learning

approach, SHAP. Building on the optimal model, we conducted a decoupling analysis to quantify how different visual elements and color indices affect VAPS, identifying both positive and negative effects as well as potential thresholds. Drawing on comprehensive and large-scale data and technical processes, this study examines the underlying mechanisms of visual aesthetics in public spaces, thereby offering refined empirical support for urban planning, street-view improvements, and public-space design.

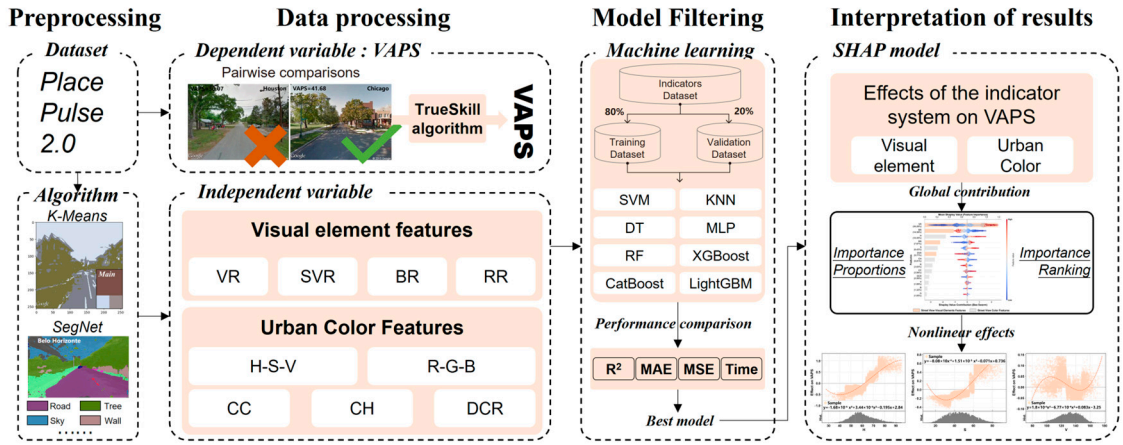

Figure 2. Research process.

2.3.1. Urban Color Features

Figure 3 summarizes the end-to-end workflow that converts each SVI into quantitative color descriptors and feeds them into the subsequent modeling pipeline. First, every image is transformed to HSV and RGB color spaces. A three-cluster K-means algorithm (k = 3, chosen as the minimum that captures foreground–middle-background variation while keeping computation light) groups pixels in HSV space; the cluster with the largest pixel share is defined as the dominant color [52,66]. Its relative size is recorded as the Dominant Color Ratio (DCR), while the cluster centroid provides the basic hue (H), saturation (S), value (V), and red–green–blue (R, G, B) channel values.

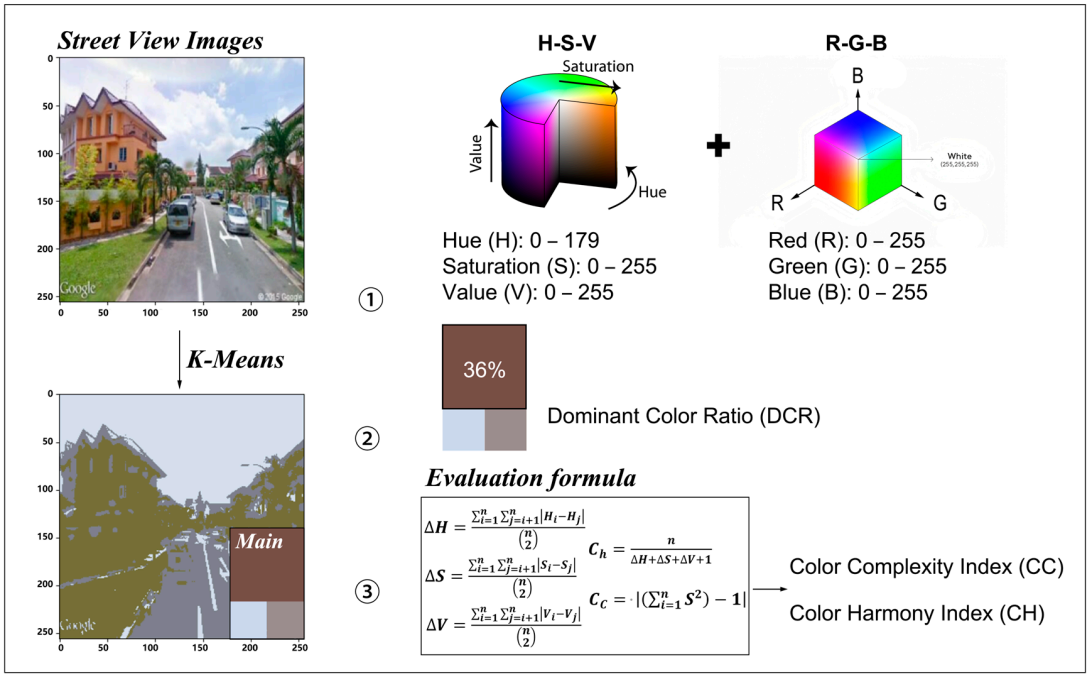

Figure 3. The process applied to extract the color features.

To better capture the richness and complexity of image colors, we have incorporated the concept of CC index into the Urban Color Features index [67]. This index measures the proportion of different colors in an image. The specific formula is as follows:

$$C_C = \left| \left(\sum_{i=1}^{n} S^2 \right) - 1 \right| \tag{1}$$

where C_C represents the CC score, while n and S represent the number of dominant colors in the image and the percentage of each color, respectively. A lower CC score implies a simpler overall color composition. To balance the accuracy of the results with computational efficiency, we divide the image into three clusters, representing the three dominant colors that make up the SVI [2], so the number of dominant colors (n) in the formula is set to 3.

CC analysis primarily quantifies both the count and proportional distribution of colors within an image, yet the aesthetic balance among these hues plays a crucial role in shaping human perception [68]. Therefore, we adopt a method that computes a CH score using color data extracted from images [2]. This method leverages the inherent attributes of the HSV color space to assess the consistency among colors. It quantifies the overall CH by measuring the relative distances between various pixels, effectively capturing the harmony in color distribution. Compared with existing research [67], this approach for measuring CH maintains adequate computational efficiency while preserving the essential color characteristics and achieving high research efficiency. The specific formula is as follows:

$$\Delta H = \frac{\sum_{i=1}^{n} \sum_{j=i+1}^{n} \left| H_i - H_j \right|}{\binom{n}{2}} \tag{2}$$

$$\Delta S = \frac{\sum_{i=1}^{n} \sum_{j=i+1}^{n} \left| S_i - S_j \right|}{\binom{n}{2}} \tag{3}$$

$$\Delta V = \frac{\sum_{i=1}^{n} \sum_{j=i+1}^{n} \left| V_i - V_j \right|}{\binom{n}{2}} \tag{4}$$

$$C_h = \frac{n}{\Delta H + \Delta S + \Delta V + 1} \tag{5}$$

C_h represents the CH score, n represents the number of dominant colors in the image—set to 3 for this study. The HSV values for these colors are represented by the variables H_i, S_i, V_i and H_j, S_j, V_j, respectively. Additionally, ΔH, ΔS, ΔV indicate the differences in hue, saturation, and value. Consequently, a higher CH score reflects a more harmonious color composition. Table 1 shows the urban color feature used in this paper.

Table 1. Urban color feature.

Variable Name	Description	Mean	Standard Deviation
Hue (H)	Hue value of the main clustered color in HSV color space.	60.730	11.412
Saturation (S)	Saturation value of the main clustered color in HSV color space.	46.798	14.729
Value (V)	Value (brightness) of the main clustered color in HSV color space.	129.705	17.971
Color Complexity Index (CC)	Index measuring the color complexity of the image (range 0–1).	0.625	0.041

Table 1. *Cont.*

Variable Name	Description	Mean	Standard Deviation
Color Harmony Index (CH)	Index assessing the harmony of the color composition in the image.	0.018	0.003
Dominant Color Ratio (DCR)	Proportion of the dominant (main) clustered color in the street view.	0.365	0.117
Red Value (R)	Red value of the main clustered color in RGB color space.	109.143	17.397
Green Value (G)	Green value of the main clustered color in RGB color space.	129.644	17.951
Blue Value (B)	Blue value of the main clustered color in RGB color space.	109.612	17.461

2.3.2. Urban Visual Elements Features

Google Street View images from Place Pulse 2.0 were semantically segmented to quantify the physical components of each streetscape. We adopted SegNet, an encoder–decoder convolutional network whose class-balanced training on the Cityscapes and CamVid benchmarks has proved both accurate and computationally efficient [34,69]. The pretrained weights were fine-tuned on a 6000-image subset of Place Pulse to accommodate color and perspective differences, using an 80:20 train–validation split and early stopping on mean Intersection-over-Union.

The final model predicts 14 pixel classes (road, building, vegetation, sky, traffic sign, etc.; see Figure 4). Because transient objects such as cars, buses, and trains add noise yet contribute little to long-term visual aesthetics, their masks were discarded [17]. For each retained class k, we computed the ratio, yielding four continuous indicators that the literature recognizes as pivotal to perceived environmental quality: Vegetation Ratio (VgR), Sky-Visibility Ratio (SkVR), Building Ratio (BR), and Road Ratio (RR) [17,37,70,71]. Table 2 summarizes the descriptive statistics of these visual element variables.



Figure 4. Example of semantic segmentation results.

Table 2. Visual element features.

Variable Name	Description	Mean	Standard Deviation
Vegetation Ratio (VgR)	Ratio of visible vegetation area to total viewable area.	0.226	0.195
Sky Visibility Ratio (SkVR)	Ratio of visible sky area to total viewable area.	0.130	0.114
Building Ratio (BR)	Ratio of visible building area to total viewable area.	0.360	0.237
Road Ratio (RR)	Ratio of visible road area to total viewable area.	0.132	0.109

This pixel-based workflow ensures that every SVI is translated into a consistent, image-derived vector of visual cues that can be directly linked—via the modeling steps detailed in Section 2.3.4—to human aesthetic judgements.

2.3.3. Visual Aesthetic Perception Score

To translate human pairwise judgments into a continuous aesthetic metric, we adopt the TrueSkill algorithm [72], a Bayesian ranking method originally developed for online gaming. This approach provides a robust and continually refined measure of aesthetic quality [58,61]. Specifically, the TrueSkill of each image is modeled as an $N(\mu, \sigma^2)$ random variable and updates after each comparison. When a user selects image x over image y in a pairwise comparison, the update equations are as follows:

$$\begin{aligned} \mu_x &\longleftarrow \mu_x + \frac{\sigma_x^2}{c} \cdot f\left(\frac{(\mu_x - \mu_y)}{c}, \frac{\varepsilon}{c}\right) \\ \mu_y &\longleftarrow \mu_y - \frac{\sigma_y^2}{c} \cdot f\left(\frac{(\mu_x - \mu_y)}{c}, \frac{\varepsilon}{c}\right) \\ \sigma_x^2 &\longleftarrow \sigma_x^2 \cdot \left[1 - \frac{\sigma_x^2}{c} \cdot g\left(\frac{(\mu_x - \mu_y)}{c}, \frac{\varepsilon}{c}\right)\right] \\ \sigma_y^2 &\longleftarrow \sigma_y^2 \cdot \left[1 - \frac{\sigma_y^2}{c} \cdot g\left(\frac{(\mu_x - \mu_y)}{c}, \frac{\varepsilon}{c}\right)\right] \end{aligned} \tag{6}$$

where $N(\mu_x, {\sigma_x}^2)$ and $N(\mu_y, {\sigma_y}^2)$ are TrueSkills of x and y. The pre-defined constant β represents the variance attributed to each comparison, and ε is the empirically determined probability of a tie. The functions $f(\theta) = \mathcal{N}(\theta)/\Phi(\theta)$ and $g(\theta) = f(\theta)(f(\theta) + \theta)$ are defined based on the Normal probability density function $\mathcal{N}(\theta)$ and the Normal cumulative density function $\Phi(\theta)$. Following Herbrich et al., we initialize rankings for all images with $\mu = 25$ and $\sigma = 25/3$, and set $\beta = 25/3$ and $\varepsilon = 0.1333$ [72].

This dynamic updating yields a robust, confidence-weighted VAPS for each image. Unlike simple win–loss tallies, TrueSkill accounts for the reliability of each comparison and handles ties explicitly, producing a stable ranking even when images receive different numbers of votes [58]. Figure 5 displays the final VAPS distribution across all SVIs. These scores serve as our dependent variable in the regression models of Section 2.3.4, providing a human-grounded benchmark that is independent of any image-derived features.

2.3.4. VAPS Decoupling Empirical Model

To identify the most suitable machine learning model for predicting the impact of urban color features and visual element features on VAPS, we evaluate the performance of eight widely used models across different algorithmic families. Specifically, we consider tree-based models, including Random Forest (RF) [43], XGBoost [44], CatBoost [45], and LightGBM [42]; distance-based and kernel methods, including k-Nearest Neighbors (KNN) [47] and Support Vector Machine (SVM) [46]; a neural network-based model, Multi-Layer Perceptron (MLP) [48]; and a traditional decision tree model, Decision Tree (DT) [49]. These models are chosen for their effectiveness in handling structured data and their widespread use in predictive modeling tasks.

All models were trained on the same feature matrix X (urban color and visual element variables) with VAPS as the response y. We employed five-fold cross-validation to obtain robust estimates of out-of-sample performance and guard against overfitting. To fine-tune their performance, we used Optuna's Bayesian optimization framework to select hyperparameters—such as learning rates, tree depths, and the number of estimators—by minimizing the validation loss in each fold [73,74].

For model comparison, we evaluated each algorithm's prediction accuracy—using Mean Absolute Error (MAE) to gauge average error magnitude, Mean Squared Error (MSE)

to place extra weight on large deviations, and R^2 to capture the proportion of VAPS variance explained—together with its training time as a measure of computational efficiency. By selecting the model that combined the lowest MAE and MSE, the highest R^2, and the shortest training time, we identified the optimal learner, which was then adopted for our final decoupling analysis and SHAP interpretation in Section 2.3.5.



Figure 5. VAPS of different SVIs. Note: Subfigures (**I**–**IV**) display representative street-view images at progressively higher VAPS levels: (**I**) Johannesburg (VAPS = 9.32); (**II**) Prague (VAPS = 20.59); (**III**) Houston (VAPS = 30.07); (**IV**) Chicago (VAPS = 41.68).

2.3.5. Interpretation of Driving Factor of VAPS

To uncover how each input feature contributes to the machine learning model's predictions of VAPS, we apply the SHAP framework [75]. Originating from cooperative game theory, SHAP assigns Shapley values to each predictor, quantifying the distinct impact of every feature on the model's output [76].

By summing each feature's Shapley value, SHAP provides an exhaustive depiction of how individual variables collectively shape predictions. The Shapley value for feature j, denoted $SHAP_j$, is derived by enumerating all possible subsets of the feature set V_p (with p being the total number of features). Mathematically, the value is computed as follows:

$$SHAP_j = \sum_{S \subseteq [V_1+V_2+\cdots+V_p] \setminus [V_j]} \frac{|S|!(p-|S|-1)!}{p!} \left(f_x\left(S \cup [V_j]\right) - f_x(S) \right)$$

Here, S is any subset of the full feature set excluding feature j, and $f_x(S)$ represents the model output restricted to the subset S. For each instance i, the total predicted value y_i is then

$$y_i = y_{base} + \sum_{j=1}^{k} SHAP(x_{i_j})$$

where y_{base} is the mean outcome across all samples, $SHAP(x_{i_j})$ is the contribution from feature j for observation i, and k is the total feature count.

In order to highlight each variable's discrete influence on the VAPS, we also consider scenarios in which feature interactions are minimized, ensuring a more straightforward view of each factor's individual role. By subtracting the multi-feature interaction components, the adjusted Shapley value for feature j at data point i can be expressed as

$$SHAP(X_{i_jj}) = SHAP(X_{i_j}) - \sum_{j \neq z}^{n} SHAP(X_{i_jz})$$

By integrating SHAP into our workflow, we obtain both global importance rankings (average $SHAP(X_{i_jj})$ across all images) and local explanations (the direction and magnitude of each feature's effect on a single prediction), thereby rendering the machine learning model's internal logic transparent and directly interpretable.

3. Results

3.1. Distribution of VAPS

This study first analyzed the distribution of the VAPSs across 56 major cities worldwide. Figure 6 displays the distribution of VAPSs across 56 major cities worldwide. A comparative analysis with the theoretical normal distribution indicates that the overall VAPS is approximately normally distributed, with a mean of 25.01 and a standard deviation of 5.60. This finding suggests that the evaluation data are relatively stable and of high quality, providing a solid foundation for subsequent empirical analyses on the factors that influence color perception tendencies.

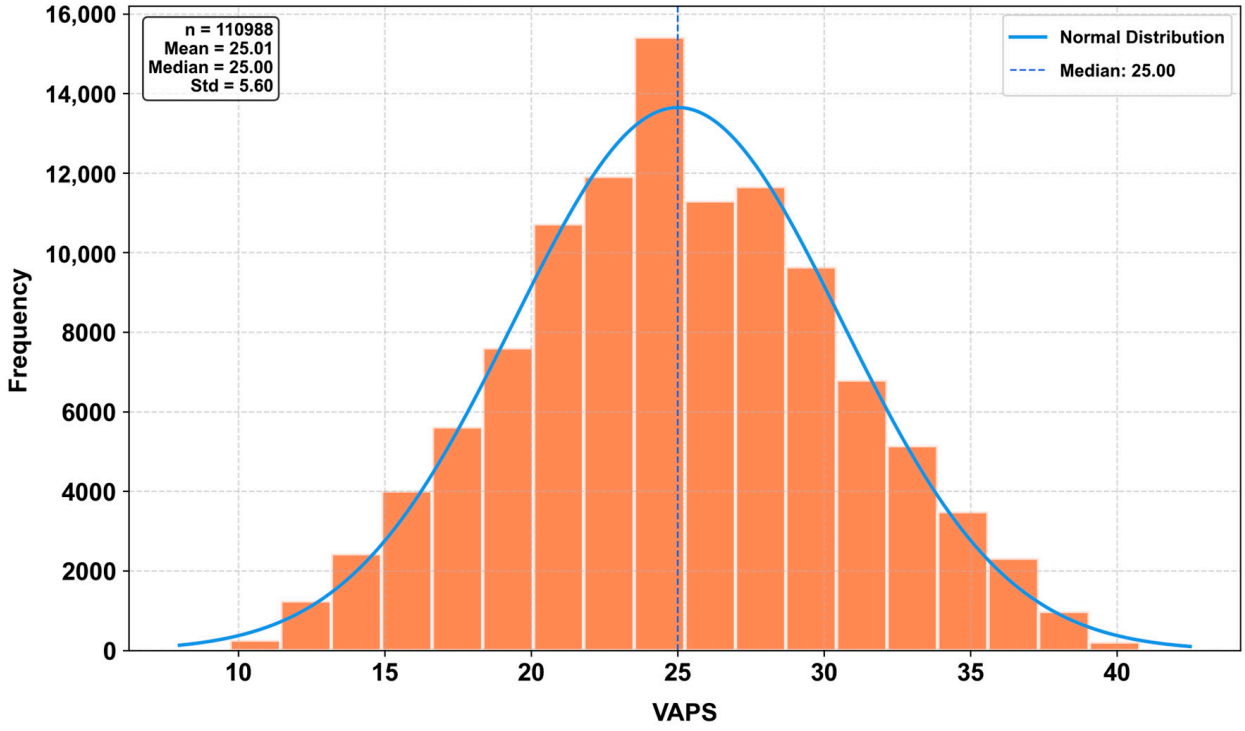

Figure 6. Distribution of VAPSs.

The Supplementary Material (Table S1) summarizes the main statistical indicators of street-view VAPSs for major global cities and Figure 7 illustrates the locations of the top four and bottom four cities by VAPS, along with the corresponding street-view points. Overall, the mean VAPS of different cities worldwide ranges from 21.889 to 28.031, indicating marked differences in street-view visual aesthetics among these cities. Meanwhile, the

standard deviations cluster between 5.0 and 6.0, suggesting consistent scoring dispersion within each city, as well as a stable level of spatial disparity. Higher-scoring cities are relatively concentrated in North America and Northern Europe, such as Washington DC and Atlanta in the United States, as well as Stockholm and Helsinki in Northern Europe. These cities typically have well-developed economies, higher levels of public-space design and urban greening, and strong competitiveness in color coordination and architectural aesthetics. In contrast, lower-scoring cities are mainly found in certain parts of Latin America and Africa, such as Belo Horizonte, Rio de Janeiro, São Paulo, and Gaborone, each with a mean VAPS below 22.5. These cities lag behind in areas such as urban infrastructure, color-environment management, and investment in public art.

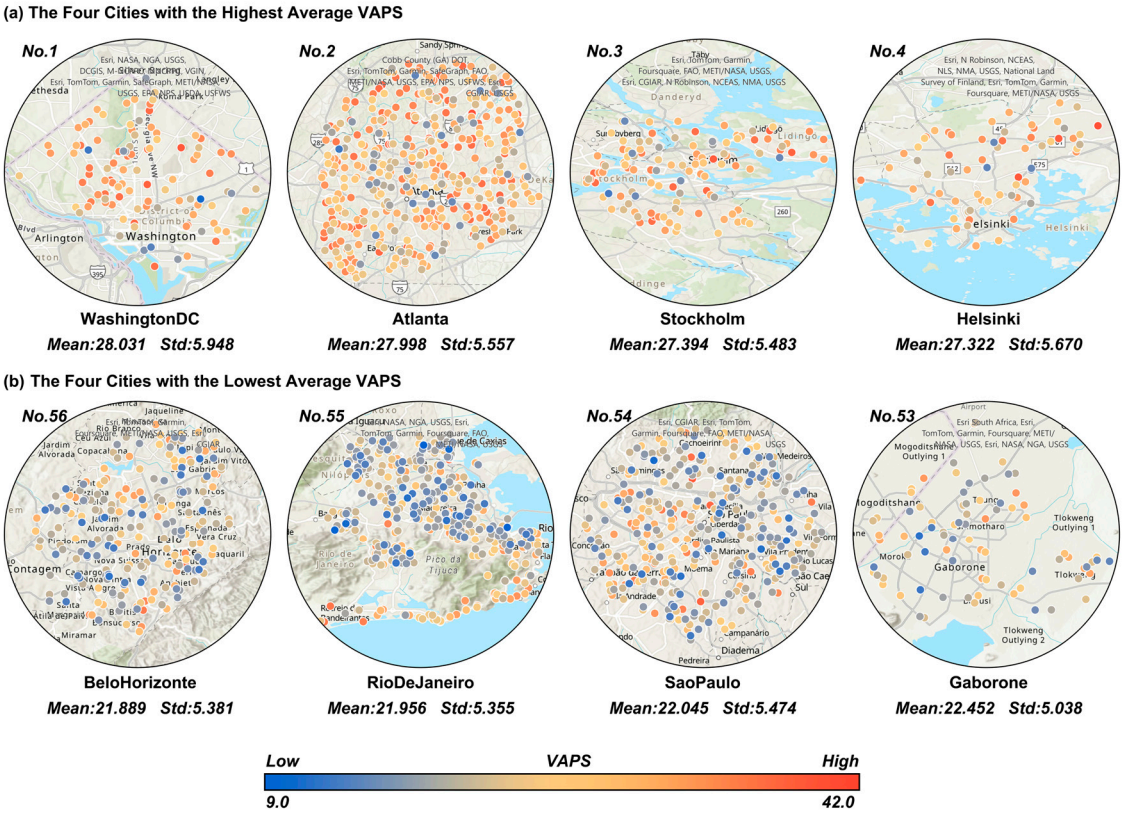

Figure 7. Distribution map of street-view points in the top four and bottom four cities by average VAPS. Note: (**a**) Distribution map of street-view points in the top four cities by average VAPS; (**b**) Distribution map of street-view points in the bottom four cities by average VAPS.

Overall, VAPS exhibits an approximately normal distribution worldwide, with stable data quality and pronounced variability. The spatial distribution patterns of high- and low-scoring cities also provide a regional backdrop for subsequent investigations into the effects of color and visual elements.

3.2. Distribution of Street-View Color and Visual Element Features

This section, in conjunction with representative street-view examples, visually illustrates the central tendency and skewness of each color and visual element feature metric in the street-view images. Figure 8 presents the kernel density distributions for different color and visual element feature indices, along with representative SVIs corresponding to the respective mean standards. From the perspective of visual elements, buildings occupy a relatively large proportion of global urban street views, with an average BR of about 0.36. In contrast, the averages for VgR and SkVR are 0.23 and 0.13, respectively, showing distinctly right-skewed unimodal distributions (with skewness around 0.854 and 0.735).

This indicates that most global urban street views exhibit relatively low greenery coverage and sky visibility.

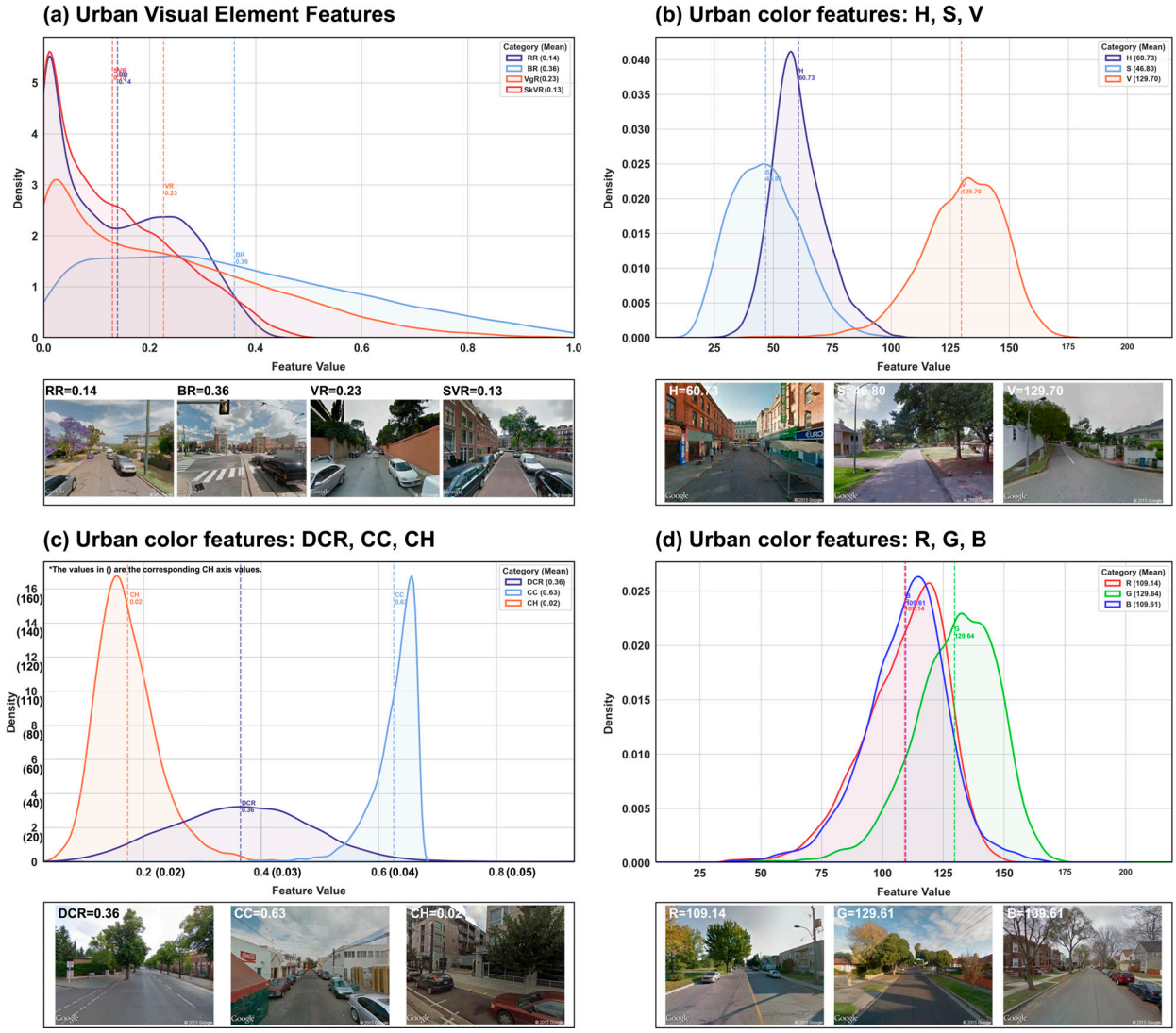

Figure 8. Kernel density distribution map and corresponding street-view examples at mean value standard. Note: the density distribution map uses the Gaussian Kernel Density Estimation method, and the street-view examples represent the mean value of the corresponding feature below the kernel density map.

From the color features of global urban street views, the mean H is 60.73, with a density peak of 57.43, indicating a leaning toward warm colors (green, yellow, orange), while cooler Hs (blue, purple) are relatively uncommon in urban architecture or street views. This trend is also evident in the RGB distributions. Additionally, low S and high V are more prevalent. The mean DCR is approximately 0.365 with a rough symmetrical distribution, suggesting that no matter how the city changes, the proportion of the dominant color in the street scene changes relatively little. Finally, in terms of color composition, the mean CC is 0.625, exhibiting left skew, and the mean CH is 0.018, exhibiting right skew; this indicates relatively high CC but lower CH in global street views. Such results may stem from the interplay of various color elements—including buildings, billboards, and natural features.

Global street views are characterized by high building dominance, limited natural elements, and warm-leaning but moderately bright colors. The combination of substantial color complexity and low harmony suggests visually rich yet unevenly coordinated urban palettes.

3.3. Performance Comparison of Different Machine Learning Decoupling Models

This section evaluates eight representative regression models in the street-view aesthetic perception regression task, aiming to identify the algorithmic framework best suited to disentangle the contributions of color and visual elements. Figure 9 presents the performance metrics of eight machine learning models in the urban color perception regression task. Overall, the DT model achieves a test R^2 of 0.751, with test MAE and test MSE of 2.106 and 7.302, respectively—noticeably underperforming compared to the ensemble models. Meanwhile, the traditional KNN and SVM models attain test R^2 values of 0.680 and 0.647, respectively, and exhibit relatively high error metrics, indicating certain limitations in handling nonlinear problems.

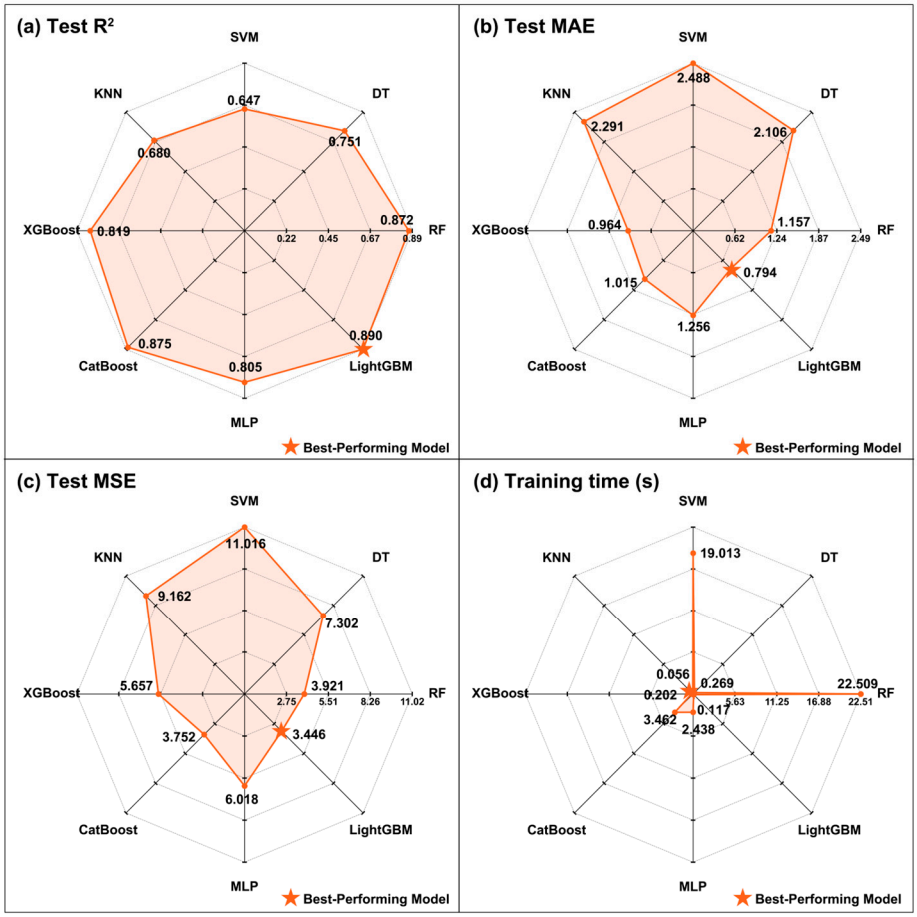

Figure 9. Decoupling performance of different machine learning models.

In contrast, ensemble learning-based models (RF, XGBoost, CatBoost, LightGBM) all exhibit high fitting and generalization capabilities, with test R^2 values exceeding 0.80. Among them, RF, XGBoost, and CatBoost achieve test R^2 values of 0.872, 0.819, and 0.875, respectively, indicating strong predictive accuracy. However, when both predictive accuracy and training efficiency are considered, LightGBM stands out: it achieves a test R^2 of 0.890, with Test MAE and Test MSE of 0.794 and 3.446, respectively, while requiring only 0.117 s of training time—significantly less than RF (22.509 s) and SVM (19.013 s). The MLP attains a test R^2 of 0.805, coming close to some ensemble models but still slightly underperforming overall. Based on these comparative analyses, this study selects the ensemble model LightGBM as the decoupling model for urban color perception to enable efficient and precise analysis of urban color and visual features.

Meanwhile, using the LightGBM model, we compared the performance of only urban color features, only visual elements features, and both sets of features (Table 3). Whether

in the training set or the test set, the model incorporating both urban color and visual feature indicators significantly outperforms the models using a single set of indicators. Specifically, the combined model achieves a very high degree of fit (R^2 close to 0.96) in the training phase and maintains strong generalization (R^2 up to 0.895) during testing, while its prediction errors (MAE and MSE) are notably lower than those of the single-dimension models. This finding further confirms that the visual aesthetic perception of urban public spaces is influenced jointly by color and visual factors.

Table 3. Model performance metrics for different dimension features.

Model	Training Dataset			Test Dataset		
	R^2	MAE	MSE	R^2	MAE	MSE
Single Urban Color Features	0.872	1.897	3.992	0.852	1.512	2.592
Single Urban Visual Element Features	0.782	2.507	4.709	0.759	2.111	3.312
Integrated Dimension	0.958	0.911	1.310	0.895	0.964	3.446

Based on the foregoing analyses and across all evaluated metrics, LightGBM not only surpasses traditional models and other ensemble algorithms in predictive accuracy but also exhibits clear advantages in training efficiency, and is thus chosen as the decoupling model for this study. Furthermore, the performance comparison explicitly shows that a single-dimensional feature set cannot comprehensively explain street-view aesthetics, whereas the integrated model combining color and visual element features markedly improves prediction accuracy, thereby further validating the methodological soundness of our multidimensional feature decoupling approach.

3.4. Analysis of Overall Feature Contribution

Building on our selection of LightGBM as the optimal decoupling model, we next examined how each feature drives the prediction of street-view aesthetic perception. Using the SHAP framework, we quantified each variable's relative contribution to the model output. SHAP values not only rank features by importance but also indicate the direction and magnitude of their effects (Figure 10).

Figure 10 shows that the VgR contributes most substantially—36.28%—underscoring the dominant role of greenery in shaping perceived urban beauty. BR and H follow, with contributions of 14.38% and 10.45%, respectively. A higher BR tends to convey urban density and structural intensity, which can suppress visual appeal, while H captures the influence of color tone on aesthetic perception.

Across all features, natural elements (VgR) emerge as the primary driver of VAPS, reinforcing the importance of green spaces in urban design. In contrast, structural elements (BR) and key color attributes (H) also exert strong but secondary influences, demonstrating that both form and color jointly shape street-view aesthetics.

3.5. Nonlinear Effects

Based on the feature-contribution analysis above, this section further elucidates the complex nonlinear relationships between street-view visual elements and color features and the VAPS, in order to explore the heterogeneous effects of each indicator across distinct threshold intervals.

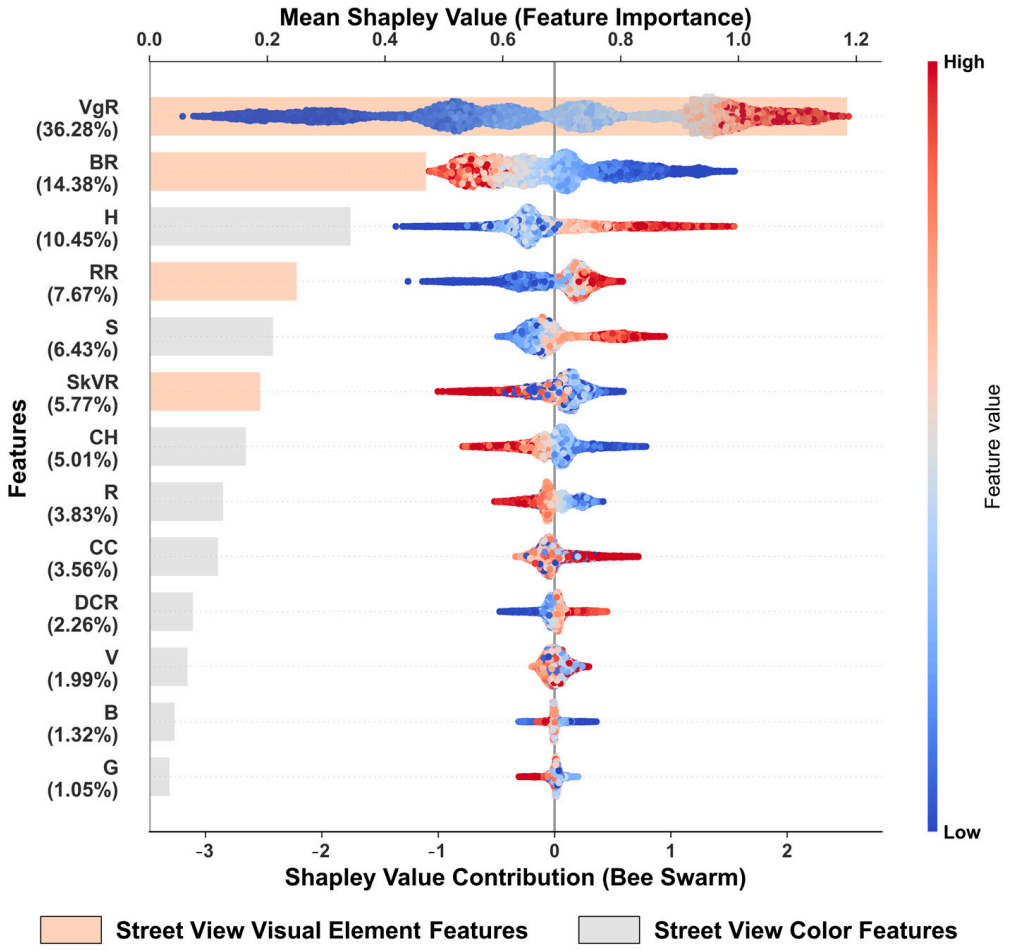

Figure 10. Feature importance proportions and ranking.

3.5.1. Nonlinear Effects of Street-View Visual Element Features on VAPS

Figure 11 illustrates the complex nonlinear relationships between different visual element features and VAPS, clearly indicating significant threshold effects and nonlinear variations in how these features influence street-view aesthetic perception. The impact of VgR on VAPS shows a distinctly nonlinear pattern. As the VgR increases, VAPS rises markedly, especially when the VgR is low, where incremental increases have a pronounced effect. However, once the VgR surpasses the 0.4 threshold, the rate of increase in VAPS begins to slow, approaching a saturation point, and shows a slight decline around 0.5. This trend suggests that while enhancing greenery exerts a clear positive effect on street-view aesthetic perception, after reaching the 0.4 level, excessive greenery no longer significantly boosts visual perception and may instead lead to visual monotony or the adverse effects of over-greening. Meanwhile, the relationship between SkVR and VAPS exhibits a more pronounced inverted U-shaped pattern. When sky visibility is low (between 0 and 0.15), increases in the sky region elicit a strong positive response in VAPS, noticeably improving visual aesthetic perception. However, once the SkVR exceeds 0.15, the effect on VAPS becomes negative, indicating that although more open sky areas can initially enhance aesthetic perception, an overly high proportion of sky may result in a visually empty or less layered scene, thereby substantially reducing VAPS.

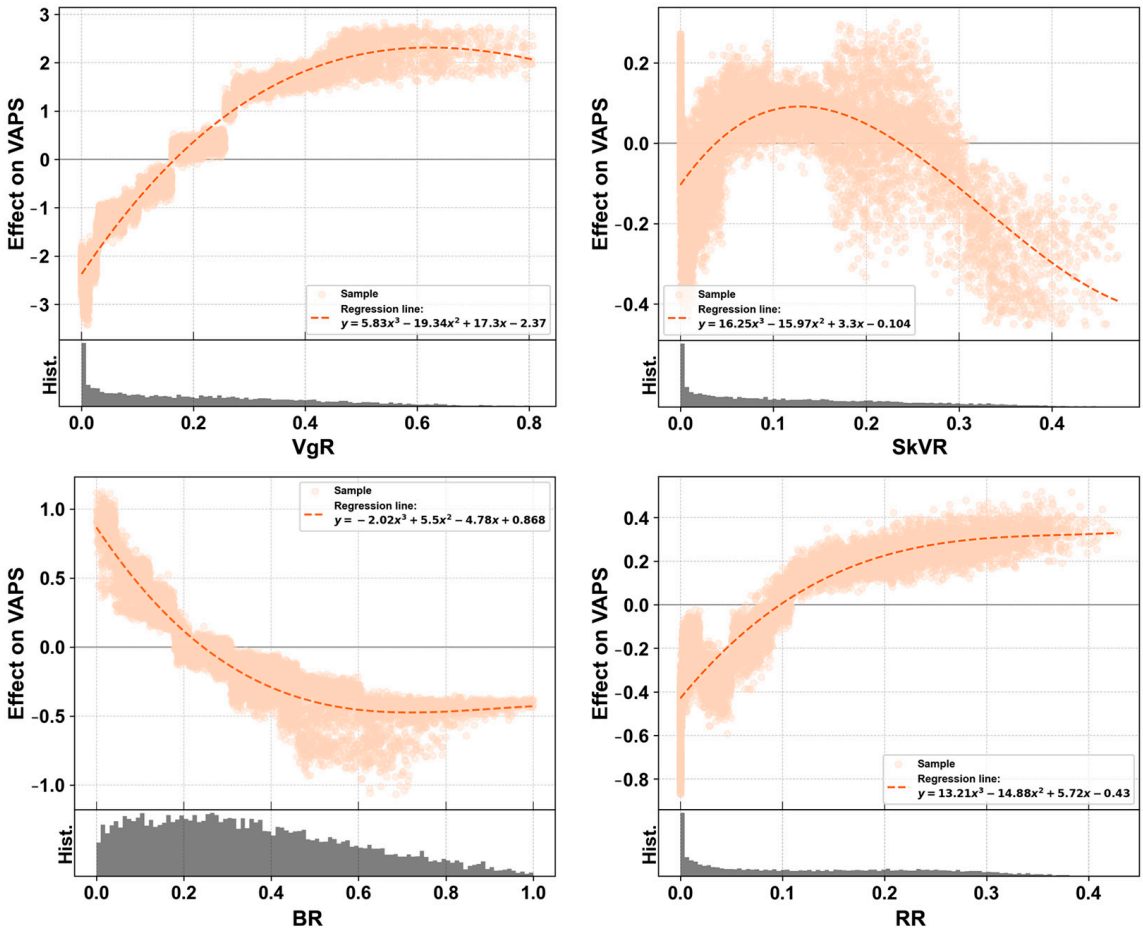

Figure 11. Nonlinear relationship between visual element features and VAPS.

The effect of BR on VAPS exhibits a clearly negative relationship. As the BR increases, VAPSs gradually decrease, and once it exceeds 0.4, the downward effect begins to level off. This indicates that an excessive concentration of buildings in urban street views may lead to a sense of spatial oppression, diminishing visual appeal. A moderately low proportion of street-facing buildings can alleviate this sense of compression and enhance the overall aesthetic perception. A certain degree of RR (between 0 and 0.2) significantly boosts the score; however, once the RR surpasses 0.3, the impact of further increases in road area gradually weakens.

In summary, the nonlinear analysis of visual elements indicates that VgR and SkVR each have critical thresholds at 0.4 and 0.15, respectively, and that BR and RR exhibit optimal value ranges, with excessively high or low levels impairing aesthetic perception. This finding underscores the importance of maintaining all elements within appropriate bounds.

3.5.2. Nonlinear Effects of Street-View Color Features on VAPS

Figure 12 further reveals the nonlinear influence patterns between color features (H, S, V, CC, CH, DCR, and RGB) and VAPS. As H increases, VAPS rises, suggesting that public spaces with higher H values—particularly those in the green-blue range (approximately 70–90)—are more visually appealing to observers, likely due to their close association with natural settings. S exhibits a similar nonlinear relationship with VAPS: within the lower S range (0–55), VAPS tends to be low; however, once S surpasses the threshold of 55, it significantly enhances VAPS, indicating that more vibrant color palettes substantially boost the visual allure of street views. The relationship between V and VAPS is more complex.

Lower V levels (below 80) typically render public spaces dim, but VAPSs rise in tandem with V, peaking around the 100–130 range. Beyond that point, further increases in V exert a negative impact on VAPS.

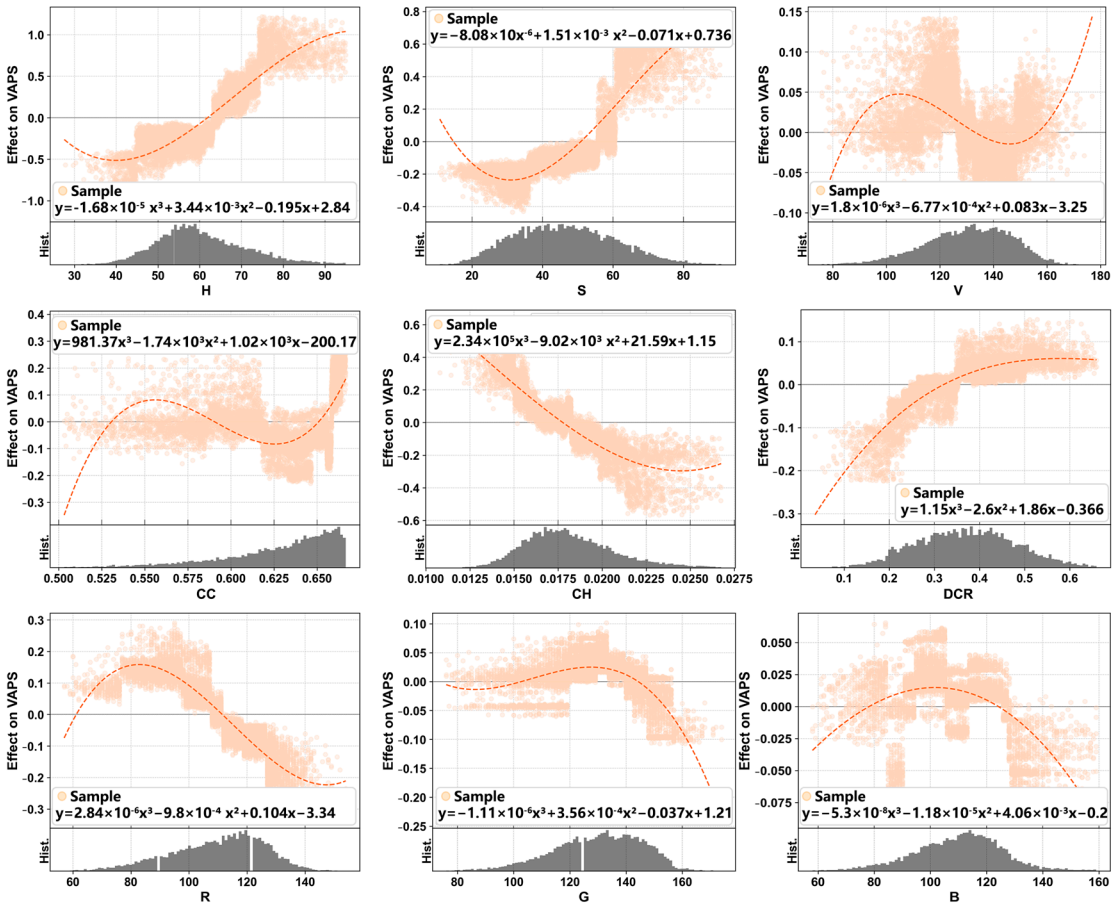

Figure 12. Nonlinear relationship between color features and VAPS. Note: the equation shown in the legend represents the fitted regression line for the scatter-plot data.

The performances of the two color composition evaluation indicators, CC and CH, corroborate each other. In the low-complexity range, increasing CC does not markedly enhance VAPS. However, when the CC index reaches a higher level, the visual effect of the street view becomes significantly richer, and the VAPS increases accordingly. This suggests that a moderate degree of CC can bolster visual impact, making street views appear more vibrant and engaging. In contrast, as CH rises, VAPSs gradually decline. Although a coordinated and unified color scheme can improve the sense of visual harmony, excessively uniform color coordination can render street views monotonous and lacking in visual impact. Street views with rich visual variety are generally more attractive than those that are monotonous, and moderately increasing color layering can effectively improve VAPS.

DCR displays a different trend in its relationship with VAPS. As the DCR increases, VAPSs gradually rise, most notably at higher levels of dominant color. A greater DCR can help create a unified visual focal point in the street view, making the overall visual perception more focused and pronounced. However, once the DCR surpasses 50%, the street view may become overly uniform and oppressive, diminishing its visual layering. All three RGB channels exhibit inverted U-shaped relationships with VAPS. At moderate levels, these color channels can effectively enhance the visual appeal and aesthetic perception of street views. However, when color values are too high or too low, the effect begins to plateau or may even have negative consequences for visual impact. Each channel has a

distinct optimal threshold range for improving VAPS. In general, these optimal ranges increase in the order of R, B, then G: 80–90 for R, 100–110 for B, and 130–140 for G.

In conclusion, the nonlinear effects of color features indicate that moderate levels of hue, saturation, brightness, and complexity are most conducive to aesthetic perception, whereas excessive uniformity or extreme values are detrimental. Street-view color schemes should strike a balance between "richness" and "harmony": moderately increasing saturation and complexity, maintaining the dominant hue proportion at 40–50%, and prioritizing blue-green tones can significantly enhance VAPS; by contrast, overly high brightness or excessive coordination may weaken visual impact.

4. Discussion

4.1. The Importance of Street Vegetation Visibility Ratio

This study demonstrates that the street VgR occupies a central role in influencing visual aesthetic perception in urban street views, contributing far more to VAPSs than other visual or color features. This finding not only aligns with conventional views in urban planning but also complements the significant role of greenery in the aesthetic perception of urban public spaces. Previous research has long recognized that urban greenery can improve residents' quality of life, enhance air quality, and mitigate urban heat island effects [77]. However, relatively few studies have investigated how greenery specifically affects people's aesthetic perception of street views [78,79]. Accordingly, this study offers a new perspective: the street VgR not only influences the ecological functions of the environment but also directly affects individuals' visual experiences of urban spaces.

Several factors help explain VgR's high contribution to street-view visual aesthetic perception. First, vegetation provides soft, natural colors that strongly contrast with the hard-edged, concrete buildings, reducing the "visual hardness" of the urban landscape and adding both depth and vibrancy [80]. The introduction of such green elements not only enriches visual perception but also makes street views more appealing and approachable. Moreover, greenery offers greater biodiversity and spaces for psychological respite within urban settings, giving people a sense of mental release from busy city life [81]. Consequently, improving the VgR directly enhances the visual attractiveness and comfort of a street view, thereby elevating its visual aesthetic perception.

However, it is important to note that the influence of the VgR is not unconditional. Excessive vegetation coverage may lead to visual monotony and an overload of "green", resulting in a lack of variety in the street-view experience. Once the VgR reaches a certain threshold, its positive impact on aesthetic perception gradually diminishes, potentially giving rise to a "saturation effect" [82,83]. This finding aligns with the "moderation" theory of urban greening, which posits that too much greenery can result in a homogenized visual perception, thereby reducing spatial diversity and complexity [83]. While a high VgR contributes to environmental sustainability, it can compromise neighborhood vibrancy and functionality if it excessively obscures commercial, cultural, or social spaces, potentially also reducing the sense of security [84,85]. Therefore, an optimal VgR of around 20% to 40% is recommended to maintain the psychological comfort offered by greenery while preventing any adverse effects on urban distinctiveness, spatial layering, and functionality.

4.2. Preferences in the Composition of Urban Street-View Visual Elements

In the realm of urban street-view visual aesthetic perception, VgR clearly plays a pivotal role. However, aside from vegetation, the proportions and configurations of other visual elements within the street view also significantly affect people's aesthetic perceptions [86,87]. In particular, the roles of BR and RR in influencing visual preferences reveal a more intricate understanding of street-view aesthetics.

First, the negative impact of BR on street-view aesthetic perception is noteworthy. The findings indicate that an overly high proportion of buildings leads to a significant drop in VAPSs. This aligns with previous scholarly views, suggesting that overly dense clusters of buildings can evoke a sense of oppression and reduce openness and perspective [70]. While urban architecture must exhibit a certain scale and functionality, it should also be distributed in a balanced manner to avoid a "wall effect". A moderate BR can contribute to order and structure in street views; however, an excessively high BR may make streets appear narrow and cramped, thereby reducing visual comfort [17].

In contrast, the effect of the RR on street-view aesthetic perception follows a complex positive nonlinear trend. Under conditions of low RR, increasing road space significantly enhances aesthetic perception—particularly in cities with insufficient transportation infrastructure or compact street layouts—by boosting the sense of flow and openness [57]. Related studies suggest that heightened road visibility can improve visual accessibility [88]. However, as the RR grows, particularly beyond a critical threshold, an excessive road presence can upset the balance of the street view, resulting in an overly urbanized environment. A disproportionately high RR not only reduces greenery but may also render the street view overly functional, lacking a sense of vitality, thereby diminishing aesthetic perception [57].

Additionally, the role of SkVR in street-view visual preferences suggests that a higher sky ratio and openness are not necessarily a good thing. Although sky visibility is relatively low in most cities, a moderate sky area imparts a sense of expansiveness and clarity, thereby enhancing visual comfort [89]. Once the SkVR surpasses 0.15, however, the overly vacant sky can evoke feelings of dullness and insecurity. This finding aligns with Whyte's observations of public spaces in New York—that people favor open spaces that still retain some sense of boundary, rather than boundless vistas [90]. Consequently, urban design should strive to improve street permeability without creating excessively open environments, preserving a degree of structural definition and enclosure in the street view.

4.3. Preferences in Urban Color for Different Populations

From the perspective of color preference tendencies, the optimal threshold ranges for H and the RGB color channels reflect a pronounced public preference for particular visual attributes in public spaces. The relationship between H and VAPS exhibits a nonlinear inverted-U pattern, indicating that the most favorable H lies primarily in the greenish-blue range (70–90). Moreover, the RGB channels suggest increasing optimal threshold intervals for red, green, and blue. People's preference for bluish-green hues may be linked to the human affinity for natural environments—blue skies and green vegetation are often associated with tranquility and harmony, offering visual comfort and a soothing effect [91]. The vividness of the space's colors plays a critical role in the visual appeal of street views, and a moderate increase in color saturation can significantly enhance the sense of vitality and uniqueness in urban landscapes.

From a color composition standpoint, visual aesthetic perception in public spaces relies on two key elements: the complexity of colors and the uniformity of the dominant color palette. Street-view design must incorporate an appropriate level of CC and richness to heighten visual impact, while the main color ensures a focused and unified overall appearance. Overly simplistic color combinations risk making a space appear uninteresting, whereas excessive complexity can render visual perception chaotic [92]. Achieving a balanced color strategy involves providing sufficient visual variety while using a dominant color theme to guide viewers' attention and avoid visual fatigue [93].

This study challenges traditional urban design theories, particularly regarding the influence of CH index on street-view aesthetic perception. The conventional view posits that a high level of CH enhances environmental beauty and visual comfort [94]. However,

our findings indicate that when CH becomes excessively strong, VAPSs actually decline, suggesting that people do not always favor overly coordinated color schemes. Instead, color combinations that provide a certain level of visual impact—featuring relatively high contrast and moderate CC—can significantly increase the attractiveness of street views. This discovery offers a new perspective for urban color planning, emphasizing that cityscapes should maintain a certain degree of layering rather than adopt monochromatic or overly complex palettes. It also aligns with the Optimal Color Richness Hypothesis [95], which proposes that people's color preferences have an optimal complexity range—neither overly simple nor overly complex [92]. In other words, urban color design should strike a balance between richness and harmony, ensuring sufficient visual stimulation without becoming cluttered.

The role of the DCR is likewise significant. An appropriate dominant color scheme can offer a clear visual focal point, directing viewers' attention to the core areas of a space. However, an overly high DCR may result in visual monotony and a lack of variety [96]. Consequently, color design in urban public spaces should revolve around a clear dominant color palette (DCR of 40–60%), while introducing moderate color variations to enrich spatial layering.

4.4. Limitations and Future Prospects

Despite systematically examining the factors influencing street-view aesthetic perception from the two primary dimensions of urban color and visual elements, this study still has several limitations. First, due to data and research scope constraints, the study focuses only on representative indicators such as dominant color, color composition, and VgR. This may overlook other dimensions within urban space—such as sociocultural factors, historical district characteristics, microclimatic conditions, or qualitative aspects like urban cultural background and artistic design intent—that could potentially affect aesthetic perception. Second, although the data scale and sources used in this study are relatively extensive, they largely consist of static street-view information, making it difficult to capture the dynamic changes in urban spaces across different seasons, time periods, or activity contexts. Third, the global scope of 56 cities makes it challenging to fully control for local contextual variables and to delve into region-specific dynamics. To overcome this, future research will include subgroup analyses by clustering cities according to geographic region, economic development level, or cultural heritage, and will also undertake focused case studies on selected city groups to yield deeper, more precise insights.

In future research, we will further expand the model's features, incorporating multiple dimensions such as urban cultural context, functional zoning of neighborhoods, and behavioral patterns, to more comprehensively reveal the complex mechanisms underlying urban visual aesthetic perception. In particular, we plan to integrate qualitative methods—such as in-depth interviews, expert reviews, and artistic design case studies—with our quantitative framework to capture the influence of cultural and design intentions on public-space aesthetics. In addition, by leveraging real-time or periodically updated street-view data and employing neural networks and large-scale image recognition technologies, we aim to achieve automated bulk collection and feature extraction of urban street-view samples and to provide dynamic monitoring and prediction of urban aesthetic perception.

5. Conclusions

This study leveraged over 100,000 street-view images from 56 major cities worldwide to develop a VAPS framework based on the TrueSkill algorithm. We systematically compared eight leading machine learning regression models—Decision Tree, KNN, SVM, MLP, RF, XGBoost, CatBoost, LightGBM—and identified LightGBM as the optimal decoupling

tool due to its superior accuracy and computational efficiency. By integrating SHAP, we further quantified each color and visual element feature's contribution to VAPS and uncovered their nonlinear threshold effects. Methodologically, our work fills a critical gap in large-scale, joint quantitative analysis of urban color and visual elements, establishing a reproducible, scalable pipeline for urban aesthetic evaluation. Theoretically, it deepens understanding of the mechanisms driving urban street-view beauty; practically, it offers actionable guidance for urban renewal and public-space planning.

Our key findings are as follows. (1) VgR is the single most influential predictor, accounting for 36.28% of the model's explanatory power, yet exhibits diminishing marginal returns beyond a 40% coverage threshold. (2) SkVR follows an inverted-U pattern, enhancing VAPS up to 15% but reducing perceived appeal when exceeding this level. (3) BR and RR each display an "overuse penalty": excessive BR ($\geq$60%) saturates compressive effects, while RR peaks at around 20% before further increases diminish vibrancy. (4) In the color domain, optimal hue values lie in the 70–90 range (bluish-green), with saturation above 55 and brightness between 100 and 130 significantly boosting aesthetic perception. (5) Moderate CC enriches visual impact, whereas overly high CH undermines layering. (6) DCR of 40–50% best balances visual focus without inducing monotony.

Based on these insights, we recommend the following design and policy measures. Urban green coverage should be maintained at 20–40%, and sky visibility at no more than 15%, to balance biophilic comfort with visual legibility, multifunctionality, and safety. Building height zoning and street-section optimization should ensure BR $\leq$ 60% and RR $\leq$ 30%, preserving spatial hierarchy while avoiding oppressive or fragmented streetscapes. In color planning, a bluish-green dominant palette at 40–50% DCR—complemented by accent colors—should be adopted, with saturation controlled above 55, brightness between 100 and 130, and color complexity elevated moderately without overharmonization, thereby achieving both visual impact and overall coherence. Implementing these guidelines can effectively enhance the aesthetic quality of urban street views and inform evidence-based public-space interventions.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/land14050979/s1, Table S1: VAPS in 56 Different Cities Worldwide.

Author Contributions: Conceptualization, T.W. and Z.W.; methodology, Z.C. and P.X.; software, Z.C., J.Z. and R.W.; validation, P.X., J.Z. and R.W.; investigation, T.W. and X.M.; resources, P.X. and J.Z.; data curation, T.W. and Z.C.; writing—original draft preparation, T.W. and Z.C.; writing—review and editing, S.L., R.Q. and J.Z.; visualization, Z.C.; supervision, T.W. and Z.W.; project administration, Z.W.; funding acquisition, R.Q. All authors have read and agreed to the published version of the manuscript.

Funding: This study is financed by The Project Supported by the National Natural Science Foundation of China No. 524B2110 and Shanghai Intelligent Science and Technology IV Summit Discipline Cross-Innovation Science and Education Integration Fund.

Data Availability Statement: The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

Conflicts of 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.

References

1. Chen, Z.; Qiao, R.; Li, S.; Zhou, S.; Zhang, X.; Wu, Z.; Wu, T. Heat and Mobility: Machine Learning Perspectives on Bike-Sharing Resilience in Shanghai. *Transp. Res. Part D Transp. Environ.* **2025**, *142*, 104692. [CrossRef]
2. Li, X.; Qin, J.; Long, Y. Urban Architectural Color Evaluation: A Cognitive Framework Combining Machine Learning and Human Perception. *Buildings* **2024**, *14*, 3901. [CrossRef]

3. Tosca, T.F. Environmental Colour Design for the Third Millennium: An Evolutionary Standpoint. *Color Res. Appl.* **2002**, *27*, 441–454. [CrossRef]
4. Zhang, C.; Tan, G. Quantifying Architectural Color Quality: A Machine Learning Integrated Framework Driven by Quantitative Color Metrics. *Ecol. Indic.* **2023**, *157*, 111237. [CrossRef]
5. McLellan, G.; Guaralda, M. Exploring Environmental Colour Design in Urban Contexts. *J. Public Space* **2018**, *3*, 93–102. [CrossRef]
6. Zhang, J.; Aziz, F.A.; Hasna, M.F. Exploring the Influence of Color Preference on Urban Vitality: A Review of the Role of Color in Regulating Pedestrian Streets. *Tuijin Jishu/J. Propuls. Technol.* **2023**, *44*, 4033–4043. [CrossRef]
7. Jeong, J.S.; Montero-Parejo, M.J.; García-Moruno, L.; Hernández-Blanco, J. The Visual Evaluation of Rural Areas: A Methodological Approach for the Spatial Planning and Color Design of Scattered Second Homes with an Example in Hervás, Western Spain. *Land Use Policy* **2015**, *46*, 330–340. [CrossRef]
8. Mehanna, W.A.E.-H.; Mehanna, W.A.E.-H. Urban Renewal for Traditional Commercial Streets at the Historical Centers of Cities. *Alex. Eng. J.* **2019**, *58*, 1127–1143. [CrossRef]
9. Bian, W. Study on the Characteristics and the Causes of Urban Color Evolution Based on New Contextualism. *IOP Conf. Ser. Earth Environ. Sci.* **2018**, *189*, 062071. [CrossRef]
10. Park, K.; Ewing, R.; Sabouri, S.; Larsen, J. Street Life and the Built Environment in an Auto-Oriented US Region. *Cities* **2019**, *88*, 243–251. [CrossRef]
11. Cavalcante, A.; Mansouri, A.; Kacha, L.; Barros, A.K.; Takeuchi, Y.; Matsumoto, N.; Ohnishi, N. Measuring Streetscape Complexity Based on the Statistics of Local Contrast and Spatial Frequency. *PLoS ONE* **2014**, *9*, e87097. [CrossRef] [PubMed]
12. Chen, Q.; Cheng, Q.; Chen, Y.; Li, K.; Wang, D.; Cao, S. The Influence of Sky View Factor on Daytime and Nighttime Urban Land Surface Temperature in Different Spatial-Temporal Scales: A Case Study of Beijing. *Remote Sens.* **2021**, *13*, 4117. [CrossRef]
13. Kaplan, R.; Kaplan, S. *The Experience of Nature: A Psychological Perspective*; CUP Archive: New York, NY, USA, 1989; ISBN 978-0-521-34939-0.
14. Wohlwill, J.F. Environmental Aesthetics: The Environment as a Source of Affect. In *Human Behavior and Environment: Advances in Theory and Research. Volume 1*; Altman, I., Wohlwill, J.F., Eds.; Springer: Boston, MA, USA, 1976; pp. 37–86. ISBN 978-1-4684-2550-5.
15. Ulrich, R.S. Aesthetic and Affective Response to Natural Environment. In *Behavior and the Natural Environment*; Altman, I., Wohlwill, J.F., Eds.; Springer: Boston, MA, USA, 1983; pp. 85–125. ISBN 978-1-4613-3539-9.
16. Kacha, L.; Matsumoto, N.; Mansouri, A. Electrophysiological Evaluation of Perceived Complexity in Streetscapes. *J. Asian Archit. Build. Eng.* **2015**, *14*, 585–592. [CrossRef]
17. Ma, X.; Ma, C.; Wu, C.; Xi, Y.; Yang, R.; Peng, N.; Zhang, C.; Ren, F. Measuring Human Perceptions of Streetscapes to Better Inform Urban Renewal: A Perspective of Scene Semantic Parsing. *Cities* **2021**, *110*, 103086. [CrossRef]
18. Han, X.; Wang, L.; Seo, S.H.; He, J.; Jung, T. Measuring Perceived Psychological Stress in Urban Built Environments Using Google Street View and Deep Learning. *Front. Public Health* **2022**, *10*, 891736. [CrossRef]
19. Dong, R.; Zhang, Y.; Zhao, J. How Green Are the Streets Within the Sixth Ring Road of Beijing? An Analysis Based on Tencent Street View Pictures and the Green View Index. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1367. [CrossRef] [PubMed]
20. Chen, X.; Meng, Q.; Hu, D.; Zhang, L.; Yang, J. Evaluating Greenery around Streets Using Baidu Panoramic Street View Images and the Panoramic Green View Index. *Forests* **2019**, *10*, 1109. [CrossRef]
21. Beil, K.; Hanes, D. The Influence of Urban Natural and Built Environments on Physiological and Psychological Measures of Stress—A Pilot Study. *Int. J. Environ. Res. Public Health* **2013**, *10*, 1250–1267. [CrossRef] [PubMed]
22. Chen, Z.; Li, S.; Liu, C. Challenges and Potentials: Environmental Assessment of Particulate Matter in Spaces Under Highway Viaducts. *Atmosphere* **2024**, *15*, 1325. [CrossRef]
23. Gao, M.; Fang, C. Pedaling through the Cityscape: Unveiling the Association of Urban Environment and Cycling Volume through Street View Imagery Analysis. *Cities* **2025**, *156*, 105573. [CrossRef]
24. Jiang, B.; Larsen, L.; Deal, B.; Sullivan, W.C. A Dose–Response Curve Describing the Relationship between Tree Cover Density and Landscape Preference. *Landsc. Urban Plan.* **2015**, *139*, 16–25. [CrossRef]
25. Mahmoudi, M.; Ahmad, F.; Abbasi, B. Livable Streets: The Effects of Physical Problems on the Quality and Livability of Kuala Lumpur Streets. *Cities* **2015**, *43*, 104–114. [CrossRef]
26. Ito, K.; Biljecki, F. Assessing Bikeability with Street View Imagery and Computer Vision. *Transp. Res. Part C Emerg. Technol.* **2021**, *132*, 103371. [CrossRef]
27. Tang, J.; Long, Y. Measuring Visual Quality of Street Space and Its Temporal Variation: Methodology and Its Application in the Hutong Area in Beijing. *Landsc. Urban Plan.* **2019**, *191*, 103436. [CrossRef]
28. Wang, R.; Lu, Y.; Zhang, J.; Liu, P.; Yao, Y.; Liu, Y. The Relationship between Visual Enclosure for Neighbourhood Street Walkability and Elders' Mental Health in China: Using Street View Images. *J. Transp. Health* **2019**, *13*, 90–102. [CrossRef]
29. Wan, J.; Zhou, Y.; Li, Y.; Su, Y.; Cao, Y.; Zhang, L.; Ying, L.; Deng, W. Research on Color Space Perceptions and Restorative Effects of Blue Space Based on Color Psychology: Examination of the Yijie District of Dujiangyan City as an Example. *Int. J. Environ. Res. Public Health* **2020**, *17*, 3137. [CrossRef]

30. Grossman, I.; Bandara, K.; Wilson, T.; Kirley, M. Can Machine Learning Improve Small Area Population Forecasts? A Forecast Combination Approach. *Comput. Environ. Urban Syst.* **2022**, *95*, 101806. [CrossRef]
31. Pang, B.; Nijkamp, E.; Wu, Y.N. Deep Learning With TensorFlow: A Review. *J. Educ. Behav. Stat.* **2020**, *45*, 227–248. [CrossRef]
32. Rui, J.; Cai, C.; Wu, Y. Towards Equal Neighborhood Evolution? A Longitudinal Study of Soundscape and Visual Evolution and Housing Value Fluctuations in Shenzhen. *J. Environ. Manag.* **2024**, *370*, 122829. [CrossRef]
33. Goel, R.; Garcia, L.M.T.; Goodman, A.; Johnson, R.; Aldred, R.; Murugesan, M.; Brage, S.; Bhalla, K.; Woodcock, J. Estimating City-Level Travel Patterns Using Street Imagery: A Case Study of Using Google Street View in Britain. *PLoS ONE* **2018**, *13*, e0196521. [CrossRef]
34. Hu, C.-B.; Zhang, F.; Gong, F.-Y.; Ratti, C.; Li, X. Classification and Mapping of Urban Canyon Geometry Using Google Street View Images and Deep Multitask Learning. *Build. Environ.* **2020**, *167*, 106424. [CrossRef]
35. Lu, Y. Using Google Street View to Investigate the Association between Street Greenery and Physical Activity. *Landsc. Urban Plan.* **2019**, *191*, 103435. [CrossRef]
36. Wu, C.; Peng, N.; Ma, X.; Li, S.; Rao, J. Assessing Multiscale Visual Appearance Characteristics of Neighbourhoods Using Geographically Weighted Principal Component Analysis in Shenzhen, China. *Comput. Environ. Urban Syst.* **2020**, *84*, 101547. [CrossRef]
37. Li, X.; Ratti, C.; Seiferling, I. Mapping Urban Landscapes Along Streets Using Google Street View. In *Advances in Cartography and GIScience*; Peterson, M.P., Ed.; Springer International Publishing: Cham, Switzerland, 2017; pp. 341–356.
38. Ye, Y.; Richards, D.; Lu, Y.; Song, X.; Zhuang, Y.; Zeng, W.; Zhong, T. Measuring Daily Accessed Street Greenery: A Human-Scale Approach for Informing Better Urban Planning Practices. *Landsc. Urban Plan.* **2019**, *191*, 103434. [CrossRef]
39. Wang, R.; Liu, Y.; Lu, Y.; Zhang, J.; Liu, P.; Yao, Y.; Grekousis, G. Perceptions of Built Environment and Health Outcomes for Older Chinese in Beijing: A Big Data Approach with Street View Images and Deep Learning Technique. *Comput. Environ. Urban Syst.* **2019**, *78*, 101386. [CrossRef]
40. Su, S.; Zhou, H.; Xu, M.; Ru, H.; Wang, W.; Weng, M. Auditing Street Walkability and Associated Social Inequalities for Planning Implications. *J. Transp. Geogr.* **2019**, *74*, 62–76. [CrossRef]
41. Faouzi, J.; Colliot, O. Classic Machine Learning Methods. In *Machine Learning for Brain Disorders*; Colliot, O., Ed.; Springer: New York, NY, USA, 2023; pp. 25–75, ISBN 978-1-0716-3195-9.
42. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.-Y. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: New York, NY, USA, 2017; Volume 30.
43. Rigatti, S.J. Random Forest. *J. Insur. Med.* **2017**, *47*, 31–39. [CrossRef]
44. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; Association for Computing Machinery: New York, NY, USA, 13 August 2016; pp. 785–794.
45. Hancock, J.T.; Khoshgoftaar, T.M. CatBoost for Big Data: An Interdisciplinary Review. *J Big Data* **2020**, *7*, 94. [CrossRef]
46. Noble, W.S. What Is a Support Vector Machine? *Nat. Biotechnol.* **2006**, *24*, 1565–1567. [CrossRef]
47. Kramer, O. (Ed.) K-Nearest Neighbors. In *Dimensionality Reduction with Unsupervised Nearest Neighbors*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 13–23, ISBN 978-3-642-38652-7.
48. Taud, H.; Mas, J.F. Multilayer Perceptron (MLP). In *Geomatic Approaches for Modeling Land Change Scenarios*; Camacho Olmedo, M.T., Paegelow, M., Mas, J.-F., Escobar, F., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 451–455, ISBN 978-3-319-60801-3.
49. Charbuty, B.; Abdulazeez, A. Classification Based on Decision Tree Algorithm for Machine Learning. *J. Appl. Sci. Technol. Trends* **2021**, *2*, 20–28. [CrossRef]
50. Zhang, S.; Liu, N.; Ma, B.; Yan, S. The Effects of Street Environment Features on Road Running: An Analysis Using Crowdsourced Fitness Tracker Data and Machine Learning. *Environ. Plan. B* **2024**, *51*, 529–545. [CrossRef]
51. Gao, M.; Fang, C. Deciphering Urban Cycling: Analyzing the Nonlinear Impact of Street Environments on Cycling Volume Using Crowdsourced Tracker Data and Machine Learning. *J. Transp. Geogr.* **2025**, *124*, 104179. [CrossRef]
52. Nguyen, L.; Teller, J. Color in the Urban Environment: A User-Oriented Protocol for Chromatic Characterization and the Development of a Parametric Typology. *Color Res. Appl.* **2017**, *42*, 131–142. [CrossRef]
53. Chen, L.; Kong, F. Quantitative Method of Regional Color Planning—Field Investigation on Renewal Design of Jiangchuan Street. In *Advances in Creativity, Innovation, Entrepreneurship and Communication of Design*; Markopoulos, E., Goonetilleke, R.S., Ho, A.G., Luximon, Y., Eds.; Springer International Publishing: Cham, Switzerland, 2021; pp. 608–617.
54. Wang, J.; Zhang, L.; Gou, A. Study of the Color Characteristics of Residential Buildings in Shanghai. *Color Res. Appl.* **2021**, *46*, 240–257. [CrossRef]
55. Zhong, T.; Ye, C.; Wang, Z.; Tang, G.; Zhang, W.; Ye, Y. City-Scale Mapping of Urban Façade Color Using Street-View Imagery. *Remote Sens.* **2021**, *13*, 1591. [CrossRef]

56. McGinn, A.P.; Evenson, K.R.; Herring, A.H.; Huston, S.L.; Rodriguez, D.A. Exploring Associations between Physical Activity and Perceived and Objective Measures of the Built Environment. *J. Urban Health* **2007**, *84*, 162–184. [CrossRef] [PubMed]
57. Zhou, H.; He, S.; Cai, Y.; Wang, M.; Su, S. Social Inequalities in Neighborhood Visual Walkability: Using Street View Imagery and Deep Learning Technologies to Facilitate Healthy City Planning. *Sustain. Cities Soc.* **2019**, *50*, 101605. [CrossRef]
58. Dubey, A.; Naik, N.; Parikh, D.; Raskar, R.; Hidalgo, C.A. Deep Learning the City: Quantifying Urban Perception At A Global Scale. In Proceedings of the Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016.
59. Salesses, P.; Schechtner, K.; Hidalgo, C.A. The Collaborative Image of The City: Mapping the Inequality of Urban Perception. *PLoS ONE* **2013**, *8*, e68400. [CrossRef]
60. Kiapour, M.H.; Yamaguchi, K.; Berg, A.C.; Berg, T.L. Hipster Wars: Discovering Elements of Fashion Styles. In *Computer Vision—ECCV 2014*; Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2014; Volume 8689, pp. 472–488, ISBN 978-3-319-10589-5.
61. Naik, N.; Philipoom, J.; Raskar, R.; Hidalgo, C. Streetscore—Predicting the Perceived Safety of One Million Streetscapes. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops, Columbus, OH, USA, 23–28 June 2014; pp. 793–799.
62. Jou, B.; Bhattacharya, S.; Chang, S.-F. Predicting Viewer Perceived Emotions in Animated GIFs. In Proceedings of the 22nd ACM International Conference on Multimedia, Orlando, FL, USA, 3 November 2014; ACM: Orlando, FL, USA, 2014; pp. 213–216.
63. Sartori, A.; Yanulevskaya, V.; Salah, A.A.; Uijlings, J.; Bruni, E.; Sebe, N. Affective Analysis of Professional and Amateur Abstract Paintings Using Statistical Analysis and Art Theory. *ACM Trans. Interact. Intell. Syst. (TiiS)* **2015**, *5*, 1–27. [CrossRef]
64. Bijmolt, T.H.; Wedel, M. The Effects of Alternative Methods of Collecting Similarity Data for Multidimensional Scaling. *Int. J. Res. Mark.* **1995**, *12*, 363–371. [CrossRef]
65. Stewart, N.; Brown, G.D.; Chater, N. Absolute Identification by Relative Judgment. *Psychol. Rev.* **2005**, *112*, 881. [CrossRef] [PubMed]
66. Gunduz, A.B.; Taskin, B.; Yavuz, A.G.; Karsligil, M.E. A Better Way of Extracting Dominant Colors Using Salient Objects with Semantic Segmentation. *Eng. Appl. Artif. Intell.* **2021**, *100*, 104204. [CrossRef]
67. Qi, Z.; Li, J.; He, Z. The Influence of Urban Streetscape Color on Tourists' Emotional Perception Based on Streetscape Images. *J. Geo Inf. Sci.* **2024**, *26*, 514–529.
68. Wei, S.-T.; Ou, L.-C.; Luo, M.R.; Hutchings, J.B. Package Design: Colour Harmony and Consumer Expectations. *Int. J. Des.* **2014**, *8*, 109–126.
69. Badrinarayanan, V.; Kendall, A.; Cipolla, R. Segnet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [CrossRef]
70. Yin, L.; Wang, Z. Measuring Visual Enclosure for Street Walkability: Using Machine Learning Algorithms and Google Street View Imagery. *Appl. Geogr.* **2016**, *76*, 147–153. [CrossRef]
71. Li, X.; Zhang, C.; Li, W.; Ricard, R.; Meng, Q.; Zhang, W. Assessing Street-Level Urban Greenery Using Google Street View and a Modified Green View Index. *Urban For. Urban Green.* **2015**, *14*, 675–685. [CrossRef]
72. Herbrich, R.; Minka, T.; Graepel, T. TrueSkill^{TM}: A Bayesian Skill Rating System. In *Advances in Neural Information Processing Systems 19, Proceedings of the 2006 Conference, Vancouver, BC, Canada, 4–7 December 2006*; MIT Press: Cambridge, MA, USA, 2006; Volume 19. [CrossRef]
73. Shekhar, S.; Bansode, A.; Salim, A. A Comparative Study of Hyper-Parameter Optimization Tools. In Proceedings of the 2021 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE), Brisbane, Australia, 8–10 December 2021; pp. 1–6.
74. Rastogi, D.; Johri, P.; Tiwari, V.; Elngar, A.A. Multi-Class Classification of Brain Tumour Magnetic Resonance Images Using Multi-Branch Network with Inception Block and Five-Fold Cross Validation Deep Learning Framework. *Biomed. Signal Process. Control* **2024**, *88*, 105602. [CrossRef]
75. Aggarwal, A.; Lohia, P.; Nagar, S.; Dey, K.; Saha, D. Black Box Fairness Testing of Machine Learning Models. In Proceedings of the 2019 27th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering, Tallinn, Estonia, 26–30 August 2019; Association for Computing Machinery: New York, NY, USA, 2019; pp. 625–635.
76. Lundberg, S.M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J.M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S.-I. From Local Explanations to Global Understanding with Explainable AI for Trees. *Nat. Mach. Intell.* **2020**, *2*, 56–67. [CrossRef]
77. Rajagopal, P.; Priya, R.S.; Senthil, R. A Review of Recent Developments in the Impact of Environmental Measures on Urban Heat Island. *Sustain. Cities Soc.* **2023**, *88*, 104279. [CrossRef]
78. Biljecki, F.; Ito, K. Street View Imagery in Urban Analytics and GIS: A Review. *Landsc. Urban Plan.* **2021**, *215*, 104217. [CrossRef]
79. Helbich, M.; Poppe, R.; Oberski, D.; Zeylmans Van Emmichoven, M.; Schram, R. Can't See the Wood for the Trees? An Assessment of Street View- and Satellite-Derived Greenness Measures in Relation to Mental Health. *Landsc. Urban Plan.* **2021**, *214*, 104181. [CrossRef]

80. Xie, X.; Jiang, Q.; Wang, R.; Gou, Z. Correlation between Vegetation Landscape and Subjective Human Perception: A Systematic Review. *Buildings* **2024**, *14*, 1734. [CrossRef]
81. Zhang, L.; Tan, P.Y.; Richards, D. Relative Importance of Quantitative and Qualitative Aspects of Urban Green Spaces in Promoting Health. *Landsc. Urban Plan.* **2021**, *213*, 104131. [CrossRef]
82. Lin, W.; Zeng, C.; Bao, Z.; Jin, H. The Therapeutic Look up: Stress Reduction and Attention Restoration Vary According to the Sky-Leaf-Trunk (SLT) Ratio in Canopy Landscapes. *Landsc. Urban Plan.* **2023**, *234*, 104730. [CrossRef]
83. Zhao, M.; Zhang, Y. Exploring the Dose-Response of Landscape Preference: A Case Study in Singapore. *Appl. Geogr.* **2024**, *170*, 103357. [CrossRef]
84. Jorgensen, A.; Hitchmough, J.; Calvert, T. Woodland Spaces and Edges: Their Impact on Perception of Safety and Preference. *Landsc. Urban Plan.* **2002**, *60*, 135–150. [CrossRef]
85. Lis, A.; Pardela, Ł.; Iwankowski, P. Impact of Vegetation on Perceived Safety and Preference in City Parks. *Sustainability* **2019**, *11*, 6324. [CrossRef]
86. Ogawa, Y.; Oki, T.; Zhao, C.; Sekimoto, Y.; Shimizu, C. Evaluating the Subjective Perceptions of Streetscapes Using Street-View Images. *Landsc. Urban Plan.* **2024**, *247*, 105073. [CrossRef]
87. Tang, F.; Zeng, P.; Wang, L.; Zhang, L.; Xu, W. Urban Perception Evaluation and Street Refinement Governance Supported by Street View Visual Elements Analysis. *Remote Sens.* **2024**, *16*, 3661. [CrossRef]
88. Liu, C.; Yu, Y.; Yang, X. Perceptual Evaluation of Street Quality in Underdeveloped Ethnic Areas: A Random Forest Method Combined with Human–Machine Confrontation Framework Provides Insights for Improved Urban Planning—A Case Study of Lhasa City. *Buildings* **2024**, *14*, 1698. [CrossRef]
89. Zhang, J.; Heng, C.K.; Malone-Lee, L.C.; Hii, D.J.C.; Janssen, P.; Leung, K.S.; Tan, B.K. Evaluating Environmental Implications of Density: A Comparative Case Study on the Relationship between Density, Urban Block Typology and Sky Exposure. *Autom. Constr.* **2012**, *22*, 90–101. [CrossRef]
90. Fitzpatrick, M. Bridging Theories, William H. Whyte and the Sorcery of Cities. *Archit. Cult.* **2016**, *4*, 381–393. [CrossRef]
91. Valdez, P.; Mehrabian, A. Effects of Color on Emotions. *J. Exp. Psychol. Gen.* **1994**, *123*, 394–409. [CrossRef] [PubMed]
92. Yu, M.; Zheng, X.; Qin, P.; Cui, W.; Ji, Q. Urban Color Perception and Sentiment Analysis Based on Deep Learning and Street View Big Data. *Appl. Sci.* **2024**, *14*, 9521. [CrossRef]
93. Zhang, S.; Liu, Y.; Nie, H. Geographical Feature Based Research on Urban Color Environment—Taking Wuhan as an Example. *IERI Procedia* **2014**, *9*, 190–195. [CrossRef]
94. Naige, C.; Xiaofan, X.; Minghong, T.; Xianming, W. A Spatial Analysis of Urban Color Harmony in Five Global Metropolises. *J. Resour. Ecol.* **2022**, *13*, 238–246. [CrossRef]
95. Bellizzi, J.A.; Crowley, A.E.; Hasty, R.W. The Effects of Color in Store Design. *J. Retail.* **1983**, *59*, 21–45.
96. Qi, Z.; Li, J.; Yang, X.; He, Z. How the Characteristics of Street Color Affect Visitor Emotional Experience. *Comput. Urban Sci.* **2025**, *5*, 7. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.