applied sciences

MDPI

Article

Urban Color Perception and Sentiment Analysis Based on Deep Learning and Street View Big Data

Mingyang Yu [1], Xiangyu Zheng [1], Pinrui Qin [2,*], Weikang Cui [1,3] and Qingrui Ji [1]

[1] School of Surveying and Geo-Informatics, Shandong Jianzhu University, Jinan 250101, China; ymy@sdjzu.edu.cn (M.Y.); 18263085938@163.com (X.Z.); 2023166102@stu.sdjzu.edu.cn (W.C.); 2023160106@stu.sdjzu.edu.cn (Q.J.)
[2] No. 1 Geological Team of Shandong Provincial Bureau of Geology and Mineral Resources, Jinan 250014, China
[3] Inspur Smart City Technology Co., Ltd., Jinan 250014, China
* Correspondence: qinpinrui1982@126.com

Abstract: The acceleration of urbanization has resulted in a heightened awareness of the impacts of urban environments on residents' emotional states. This present study focuses on the Lixia District of Jinan City. By using urban street view big data and deep learning methods, we undertook a detailed analysis of the impacts of urban color features on residents' emotional perceptions. In particular, a substantial corpus of street scene image data was extracted and processed. This was performed using a deep convolutional neural network (DCNN) and semantic segmentation technology (PSPNet), which enabled the simulation and prediction of the subjective perception of the urban environment by humans. Furthermore, the color complexity and coordination in the street scene were quantified and combined with residents' emotional feedback to carry out a multi-dimensional analysis. The findings revealed that color complexity and coordination were significant elements influencing residents' emotional perceptions. A high color complexity is visually appealing, but can lead to fatigue, discomfort, and boredom; a moderate complexity stimulates vitality and pleasure; high levels of regional harmony and aesthetics can increase perceptions of beauty and security; and low levels of coordination can increase feelings of depression. The environmental characteristics of different areas and differences in the daily activities of residents resulted in regional differences regarding the impacts of color features on emotional perception. This study corroborates the assertion that environmental color coordination has the capacity to enhance residents' emotions, thereby providing an important reference point for urban planning. Planning should be based on the functional characteristics of the region, and color complexity and coordination should be reasonably regulated to optimize the emotional experiences of residents. Differentiated color management enhances urban aesthetics, livability, and residents' happiness and promotes sustainable development. In the future, the influences of color and environmental factors on emotions can be explored in depth, with a view to assist in the formulation of fine urban design.

Keywords: urban streetscape; deep convolutional neural network; emotion perception; color features; random forest

check for updates

Citation: 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. https://doi.org/10.3390/app14209521

Academic Editor: Andrea Prati

Received: 25 August 2024
Revised: 7 October 2024
Accepted: 16 October 2024
Published: 18 October 2024

Copyright: © 2024 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

The phenomenon of urbanization has become one of the most significant global occurrences, as evidenced by the World Urbanization Prospects report, which indicates that over two-thirds of the world's population is projected to reside in urban areas by 2050 [1]. It has been demonstrated that the strategy of urban densification is important for accommodating the rapid growth of urban populations. However, emerging evidence suggests that urban densification may pose serious risks to the mental health of individuals [2]. In recent decades, China has undergone the most rapid urbanization in the world, with the government investing considerable resources in the construction of a network of urban streets to facilitate this expansion. The urban street and its surrounding neighborhoods

are complex entities that intertwine the physical components of the built environment. According to the principles of urban design, streets have a natural 'place' function, which provides residents with the opportunity to conduct activities, socialize, and congregate [3].

The term 'urban environment' is used to describe the landscape of a city, encompassing the physical characteristics and cultural accumulation of the built environment. Color design in urban environmental planning is the study of how color can be applied to enhance the visual appeal of urban environments. In the field of urban environment design, the use of color is a crucial element in determining the overall aesthetic impact of a space [4]. As an indispensable component of the urban environment, urban color can be used to intuitively convey the image of a city, to shape the characteristics of a city, and to affect the psychological feelings of residents and tourists. In the field of urban research, there has been a notable increase in interest surrounding the utilization of deep learning and big data technologies, largely due to their distinctive capabilities and substantial potential. The advent of the big data era has seen the widespread use of web images with spatial location attributes as open data for urban built environment research. The combination of web images and machine learning allows for large-scale quantitative measurements of urban color [5]. It is, therefore, imperative that the spatial quality of urban streets is quantified using techniques such as urban streetscape imagery and deep learning. This is a crucial issue for researchers and planners alike, as it enables improvements in the quality of life of urban residents and the achievement of sustainable urban development [6].

The application of deep learning and big data analytics has been a significant factor in the success of numerous industrial sectors over the past decade [7]. It is anticipated that data resources will constitute a significant asset in the future [8]. The processing and analysis of substantial quantities of street-scene data enable the discovery of hitherto unidentified patterns and trends, thereby providing a foundation for the formulation of urban management and policy. To illustrate this, the implementation of big data analytics and artificial intelligence (AI) in environmental governance can facilitate the provision of data and technical assistance for environmental public governance [8]. For an extended period, researchers sought to extract insights from data for decision making and forecasting. This led to the advent of data mining and numerous machine learning algorithms. From the standpoint of contemporary internet development, big data and artificial intelligence represent pivotal areas of advancement. The internet relies on big data to fulfill its potential, while artificial intelligence will further extend the scope of internet applications [8]. These two disciplines are mutually reinforcing, offering novel methodologies and instruments for addressing intricate urban challenges.

In recent years, the relationship between urban color and residents' emotional perception has become a hot research topic in the fields of urban planning and environmental psychology. Color not only shapes visual aesthetics, but also profoundly affects the emotional and psychological states of individuals [9]. Most studies have shown that urban color directly or indirectly affects people's emotional experiences by influencing residents' daily activities, enhancing social interactivity, and increasing exposure to green space [10]. Although studies have highlighted the role of street vegetation (e.g., streets with a high tree and plant density) in alleviating depression and facilitating attention recovery, there is a lack of research on how other streetscape features (e.g., color) are perceived emotionally [1].

At present, deep learning is making considerable progress in several fields, including image classification, image segmentation, target detection, speech recognition, natural language processing, and automated driving [11]. It has also become increasingly proficient in performing rapid, precise, and expansive image segmentation or target detection [12]. The Mask R-CNN (Mask Region-based Convolutional Neural Network) algorithm proposed by He et al. [13] and the Pyramid Pooling Model (PSPNet) proposed by Zhao et al. [14] have achieved significant results in the semantic segmentation of urban images. Tang et al. [15] used the machine learning segmentation method SegNet (Segmentation Network) to automatically achieve the physical visual quality scoring of street space. This was conducted by assessing the visual perceptual quality in five dimensions and rating

quality changes based on the identified physical spatial changes. Deepank et al. [16] used deep learning modeling to extract high-level and low-level features from collected audio and video. Zhang et al. [17] proposed a data-driven perceptual scoring training of a large number of street-level images for a deep learning method to simulate and predict the emotional perception of a place by humans. Yang et al. [18] constructed a new method for evaluating the color harmony of historical buildings through street view technology, a semantic segmentation algorithm, a quantification of color harmony method based on image attribute detection and classification, and questionnaire validation. Hang Chen et al. [19] designed a street view segmentation algorithm with a joint attention mechanism and multi-level feature fusion, which uses a compression and excitation network (SENet) to improve the Panpotic-DeepLab network, aggregates the channel context information, enhances the network to learn the important features, and suppresses the useless features, which effectively improves the accuracy of panoramic segmentation. Gong et al. [20], using Google Street View (GSV) data and combining deep learning algorithms, proposed a method to accurately evaluate the sky visibility index (SVF), tree cover index (TVF), and building shading index (BVF) in street valleys. Zang et al. [21] demonstrated the application of deep learning in urban land use classification, which provides background support for the use of the PSPNet model for street view image processing in this study. Liang et al. [22] devised a network designated as the Feature-Driven Spatial Attention Network (FDSANet), which ingeniously incorporates deep separable convolution and attention mechanisms to achieve the optimal balance between the segmentation accuracy and processing speed of the network model, significantly improving the efficiency and performance of semantic segmentation tasks. Yunjie Zhao [23] proposed a semantic segmentation model improved by DeepLabv3+ for the semantic segmentation task of city streetscapes and conducted experimental attempts on a city-level surveillance quality evaluation system platform, which improved the accuracy of semantic segmentation. Yang Heng [24] used the semantic segmentation technique to accurately extract the target buildings in his study and cleverly introduced the attention mechanism to enhance the discriminative ability of the model, which significantly improved the classification effect and provided a high-quality database for the design of a subsequent building classification model.

Ding et al. [5] applied color extraction, computer vision processing techniques, and clustering algorithms for image recognition to establish a city color database and quantify color attributes. Hongnan [25] used the interpolation and regression algorithm of the MATLAB R2023a software platform in urban color to achieve the automatic filling of the dominant city color map and obtain the ideal color map of the dominant city color. Ziyin Qi et al. [26] integrated the full convolutional neural network (FCN) and random forest (RF) algorithm, successfully extracted the color characteristics of streetscapes, and constructed a corresponding color quantification index. Chang Liu et al. [27] transformed the pixel information of their detection target into HSV color space, introduced the DBSCAN clustering algorithm in the hue–saturation (H-S) plane, and fitted the precise distribution range of the main color from the scattered data by the Quickhull (fast hull) convex packet algorithm. Jiafeng Liu et al. [28] designed a new image segmentation method by combining Lab (CIELAB) color space and the improved adaptive K-means clustering method in order to achieve the adaptive segmentation of different images.

In light of the aforementioned findings, the present study represents an extension of previous research, employing deep learning and street view big data to investigate the influences of urban color features on residents' subjective emotional perceptions.

2. Data and Methods

2.1. Overview of the Study Area

Jinan is a prefecture-level city, provincial capital, and sub-provincial city under the jurisdiction of Shandong Province. It is located in east China and enjoys a warm, temperate continental monsoon climate. It is the center of the southern wing of the Bohai Rim, as approved by the State Council. As of the end of 2023, the city is subdivided into 10 districts

and 2 counties, with a total area of 10,244.45 Km^2 [29], a resident population of 9,437,000, and an urbanization rate of 75.3% [30]. Lixia District, situated in the eastern portion of Jinan City, constitutes a pivotal area in the nascent development of Jinan City and serves as a significant locale for the daily activities of local residents (Figure 1). Jinan is known as the "Spring City" due to its abundant spring water resources, with Lixia District being particularly rich in such resources. The residents of this district live in close proximity to these resources, and their livelihoods and economic activities are inextricably linked to them. Consequently, Lixia District was selected as the research site to investigate the relationship between Jinan's urban colors and residents' emotional perceptions.

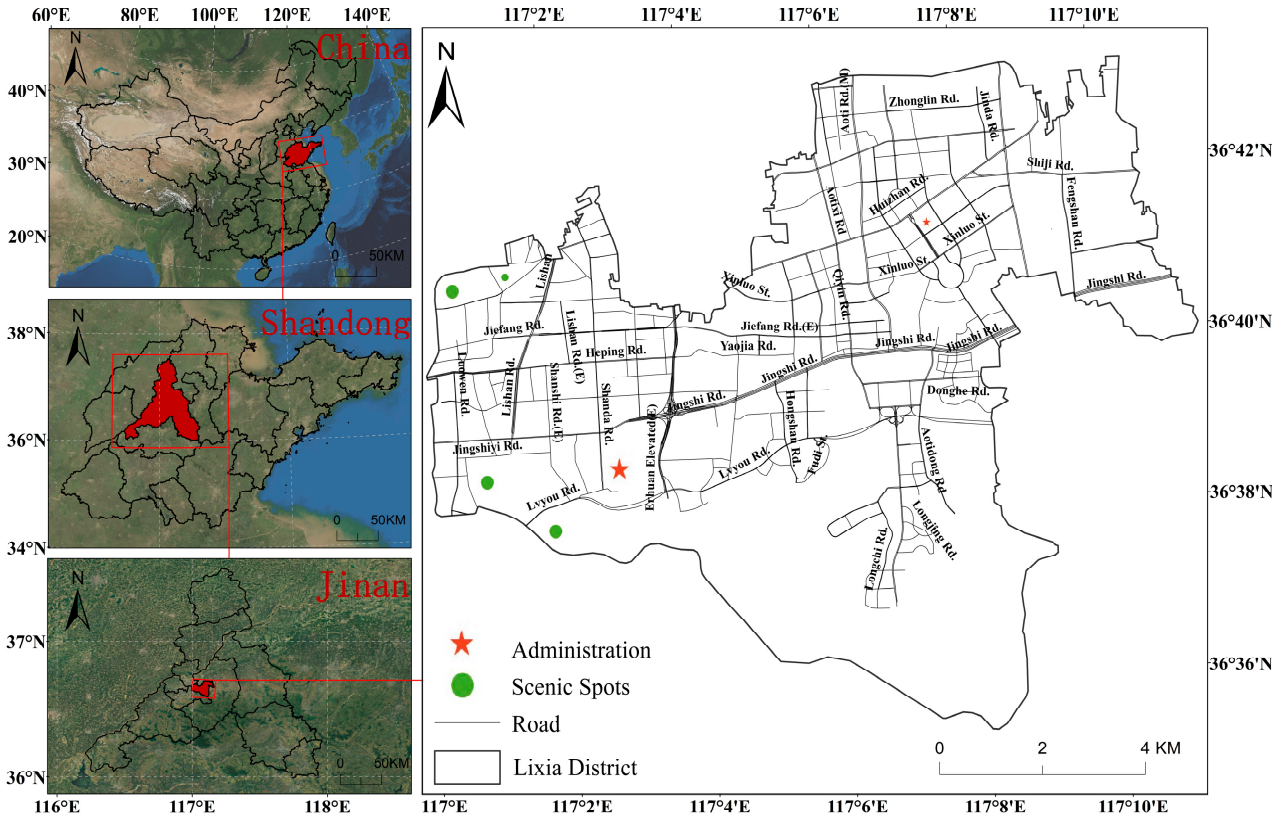


Figure 1. Overview of Lixia District, Jinan.

2.2. Data Sources

2.2.1. Road Network Collection and Processing

In this study, the Overpass API index, which is suitable for use at the urban scale, was employed to download the road data pertaining to Jinan City in 2023 (Figure 2). Initially, the ArcGIS 10.7 intersection tool was employed to extract the road data from Lixia District. The corresponding buffers were then delineated in accordance with the specific characteristics of the roads, ensuring that these buffers could fully encompass the road network of the designated class. Secondly, the ArcScan tool was employed to transform the two-lane road network into a single-lane road network. Ultimately, along the simplified single-line road network, the sampling points of streets were generated at 50 m intervals, resulting in a total of 5048 sampling points after the exclusion of invalid sampling points, including those lacking streetscape data [31] (Figure 3).

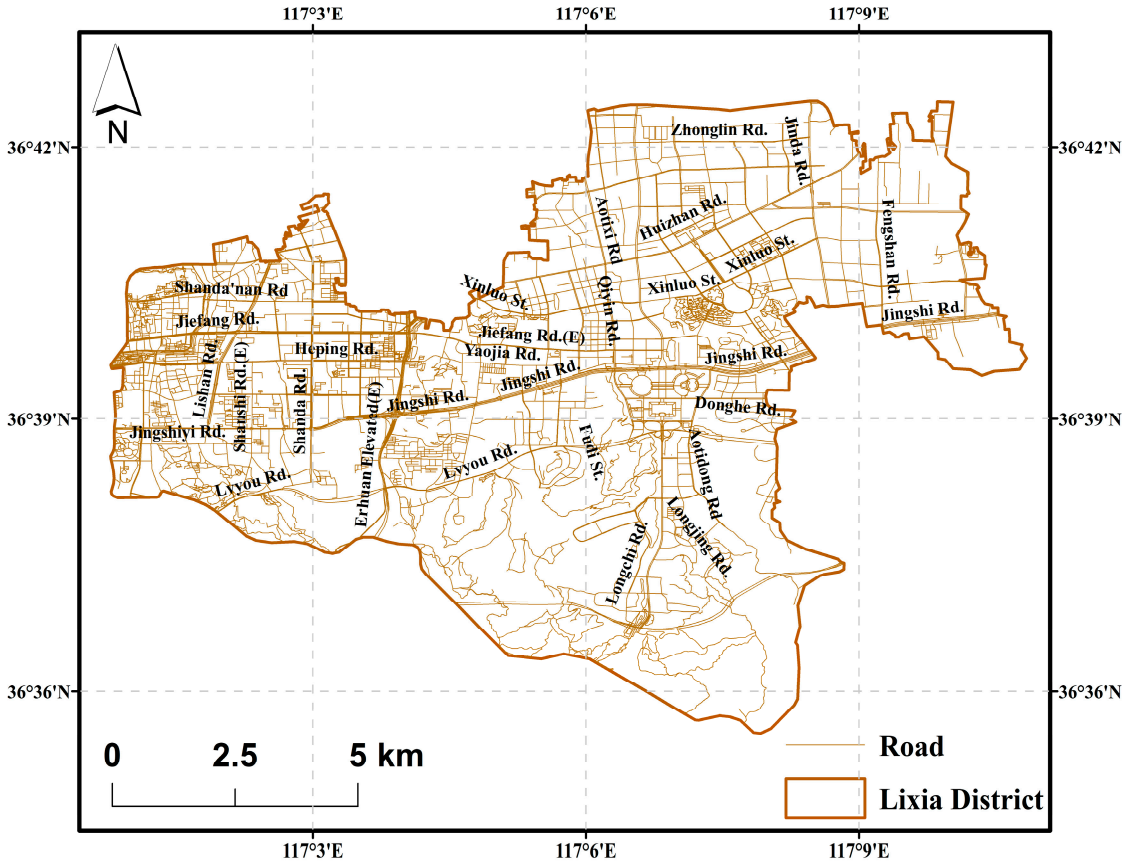


Figure 2. Overview of the road network in Lixia District.

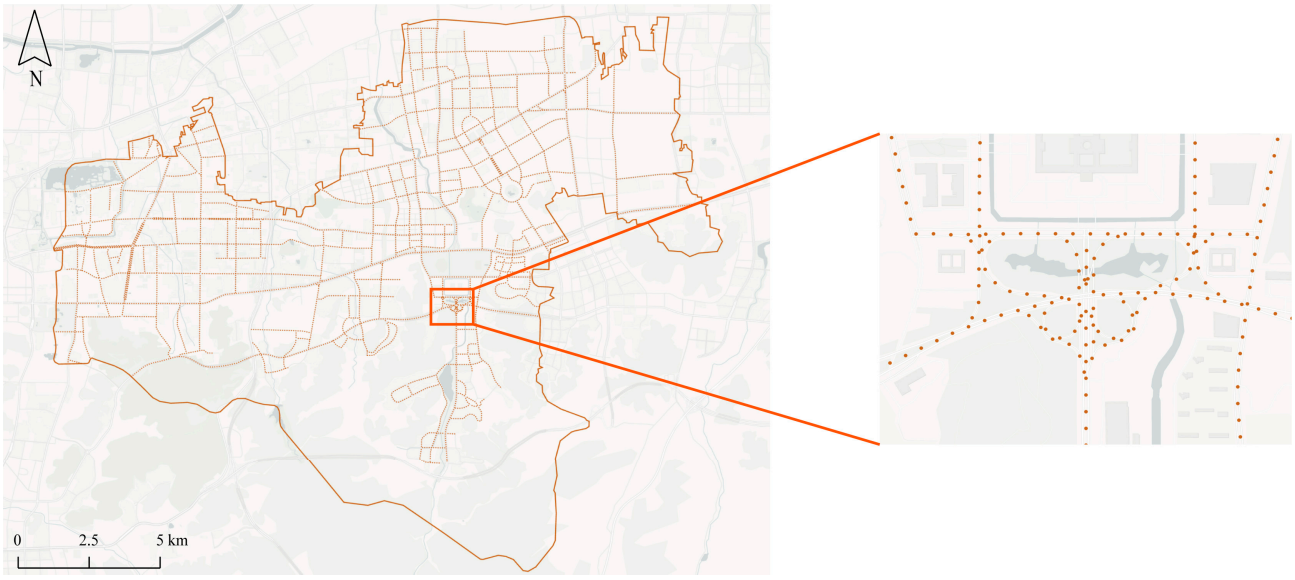


Figure 3. Schematic diagram of street sampling points in Lixia District.

2.2.2. Collection of Street View Image Data

The growing popularity of street view data, as exemplified by the advent of Baidu Street View and Google Street View in recent years, has opened up new avenues for the rapid acquisition of high-precision street view data [31]. The street view images utilized in this study were sourced from the Baidu Map Street View service (Figure 4). A crawler program written in the Python language was employed to retrieve street view images at

sampling points distributed at 50 m intervals, in accordance with the coordinates provided by the Baidu Map. The Baidu Street View service API was used to capture four street view images (Figure 5) parallel to the road (in front and behind) and perpendicular to the road (on the left and on the right), with a viewing angle of 90° in each direction of the line of sight [32]. Following calibration and screening, a total of 4797 street view images of the study area were acquired.

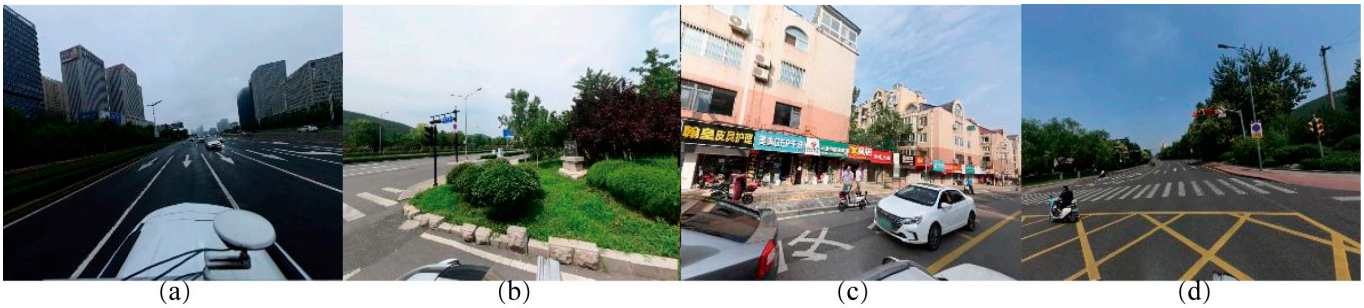


Figure 4. Example of a street scene picture. (**a**) This image shows a main urban road with modern skyscrapers on the left, demonstrating the density and spatial structure of the city. (**b**) This image shows a green space, demonstrating the presence of green belts in the city and emphasising the importance of natural elements in the urban environment. (**c**) This image shows a commercial street lined with a number of shops and pedestrians, reflecting the vibrancy and economic activity of the city. (**d**) This image shows a quieter street, signalling the traffic mobility and residential environment that characterise the city.



Figure 5. Schematic diagram of street view image crawling.

2.2.3. MIT Places Pulse Dataset

This paper makes use of the Place Pulse 2.0 project, an online data collection platform from the MIT Media Lab, which contains 110,988 street view images from 56 cities around the world. These images are used to ascertain people's perceptions of how their cities look. Participants are presented with two randomly selected images of the same city and are asked to indicate which one appears to possess greater characteristics associated with a specific attribute. In response to this question, they provide ratings on six dimensions, including perceptions of safety, beauty, liveliness, richness, depressing, and boring.

A total of 162 countries were represented by the volunteers who participated in the experiment, thereby illustrating the extensive geographical diversity of the dataset. In order to ascertain whether there were any potential biases in the data collection process due to the demographic characteristics of the volunteers, a correlation test was conducted. The results demonstrated that there was no significant bias among the groups with different demographic characteristics [33]. Furthermore, the internal consistency of the ratings within the dataset was also evaluated by examining the reproducibility and transferability between users, with both found to be high [34].

2.3. Image-Aware Score Calculation and Classification

This study employs a quantitative approach based on a deep learning framework initially proposed by Dr. Fan Zhang [17], with the objective of modeling and predicting people's subjective perceptions of the physical environment of a specific location. In particular, this study utilizes a Deep Convolutional Neural Network (DCNN) model based on a Residual Network (ResNet) to parse high-level semantic information from images by employing street-level image representations of locations in conjunction with deep learning techniques. This approach enables the automatic and efficient measurement of subjective human perceptions of large urban environments, thereby offering novel perspectives and tools for urban environmental assessment and planning.

In order to evaluate each image sample, a single-sample rating calculation method was employed. This method entailed comparing each image sample with as many other images as possible, with the perceived intensity of a particular dimension serving as the scoring criterion. The specific scoring process was calculated based on the percentage of selected images and their weighted sum. The specific calculation method is as follows:

Firstly, the positive rate (P) and negative rate (N) of image i on a specific perceptual metric are defined as follows [17]:

$$P_i = \frac{p_i}{p_i + e_i + n_i} \tag{1}$$

$$N_i = \frac{n_i}{p_i + e_i + n_i} \tag{2}$$

In this context, the variable p_i represents the number of times image i has been selected, while n_i denotes the number of times it has not been selected. The variable e_i, on the other hand, signifies the number of times image i has been considered as equal to another image.

Secondly, based on the aforementioned data, the Q-score of image i on a specific perceptual metric is further delineated as follows [17]:

$$\mathrm{Q}_i = \frac{10}{3}\left(P_i + \frac{1}{p_i}\sum_{k_1=1}^{p_i} P_{k_1} - \frac{1}{n_i}\sum_{k_2=1}^{n_{k_1}} N_{k_2} + 1\right) \tag{3}$$

where k_1 denotes the number of times image i is selected and k_2 denotes the number of times image i is not selected. Although the final Q-score is actually calculated using the positive rate P_i of image i, it needs to be corrected based on the selected and unselected frequencies $P_{i'}$ and $N_{i'}$ of the other images with which image i is compared. Finally, in order to normalize the scores to a range from 0 to 10, a multiplying factor of $10/3$ is introduced into the equation and a constant term is added.

The human perception prediction problem is approached as a binary classification task, whereby the dataset is divided into positive and negative groups based on the Q-scores of the samples. This approach allows for the straightforward categorization of the predicted images as either positive or negative. In order to minimize the introduction of noise and error, we initially calculated the mean μ_v and standard deviation σ_v of the dataset, and introduced a ratio variable δ to determine the threshold for classification. In particular, for image i, whose Q-score is denoted as Q_i^v and its label as y_i^v, this can be expressed as [17]:

$$y_i^v = \begin{cases} -1 & if \quad Q_i^v < \mu_v - \delta\sigma_v \\ 1 & if \quad Q_i^v > \mu_v + \delta\sigma_v \end{cases} \tag{4}$$

where two thresholds $-\mu_v - \delta\sigma_v$ and $\mu_v + \delta\sigma_v$ define the boundaries of the positive and negative samples, with any data that fall between these two thresholds being deemed as "noisy" and subsequently being removed. The data are labeled as "-1" for negative samples and "1" for positive samples [17]. The variable δ is employed to adjust the interval (i.e., bandwidth) between these two thresholds. By modifying the value of δ, we can observe and analyze changes in model performance.

2.4. PSPNet

The field of image segmentation constitutes an important branch of computer vision. The application of semantic segmentation algorithms enables the categorization of pixels within an image, whereby distinct semantic categories are assigned to objects based on their intrinsic characteristics and spatial positions. This approach ultimately facilitates pixel-level segmentation [35]. In this study, the PSPNet (Pyramid Scene Parsing Network) semantic segmentation model is employed with the objective of predicting emotional states in large urban areas and investigating the specific influences of street scene elements on emotional states in depth. The PSPNet system utilizes a multi-level feature pyramid module on a high-level feature map, enabling the network to adapt effectively to a range of target scales and integrate global and local features, thus reducing the loss of contextual information [36]. This module is capable of extracting and fusing feature maps of four distinct scales derived from convolutional neural network (CNN) features, thereby capturing and integrating the contextual information of disparate regions within images [37]. Furthermore, feature images are restored to their original dimensions through an up-sampling method, which markedly enhances the precision and efficacy of semantic segmentation (Figure 6).

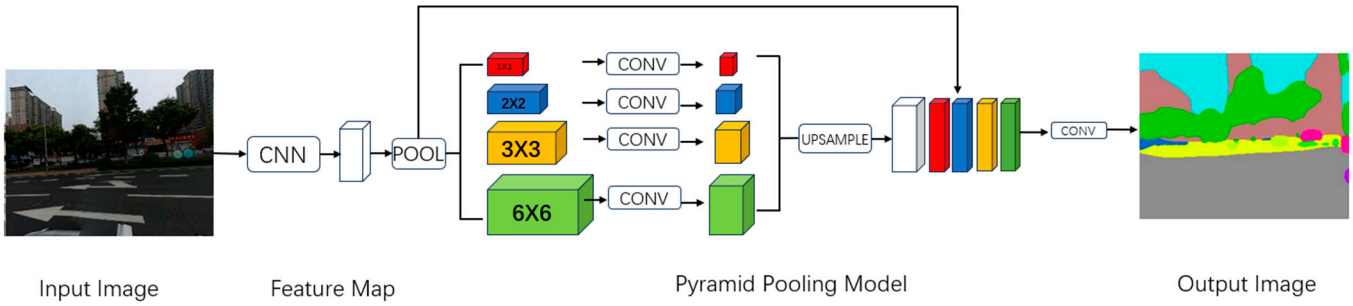


Figure 6. PSPNet architecture.

2.5. Modeling the Effects of Street Environments on Emotional States in Metropolitan Areas

As demonstrated by prior research [17], it is viable to conceptualize the process of modeling emotional states in metropolitan regions as a binary classification task when examining subjective visual perceptions. A binary classifier, specifically a Support Vector Machine (SVM), is employed in the supervised learning technique to train high-dimensional deep features (Figure 7). The kernel function, at the core of the SVM, identifies linear boundaries in high-dimensional space, enabling distinctions between different classes. For a typical Support Vector Classifier (SVC), the decision function can be expressed as follows:

$$\mathrm{sgn}\left(\sum_{i=1}^{N} \alpha_i y_i K(x_i, x) + b\right) \quad (5)$$

where $x_i \in R^p (i = 1, 2 \ldots c)$ is the input feature variable and $y_i \in \{1, -1\}^n$ is the class label. The Gaussian kernel function, $K(x_i, x)$, is used to calculate the similarity between samples, while the Lagrange multiplier, α_i, represents the contribution of the support vector to the decision boundary. In this instance, we utilize one of the most prevalent basis functions, namely the Gaussian Radial Basis Function (RBF), which is defined as follows:

$$K(x_i, x) = \exp\left(\frac{-\|x_i - x\|}{2\sigma^2}\right) \quad (6)$$

2.6. Assessment of the Impact of Streetscape Elements on Emotional States

In order to identify the street scene visual elements that may cause a place to be perceived as beautiful, depressing, boring, etc., this study employed image-based perceptual scores with PSPNet semantic segmentation to conduct a multiple regression analysis on six perceptual metrics (Figure 8). For each perceptual indicator, the corresponding perceptual scores were initially obtained from the streetscape dataset. Subsequently, the number of

pixels in the semantic segmentation mask was employed to calculate the proportion of area occupied by each visual element in the image [17].

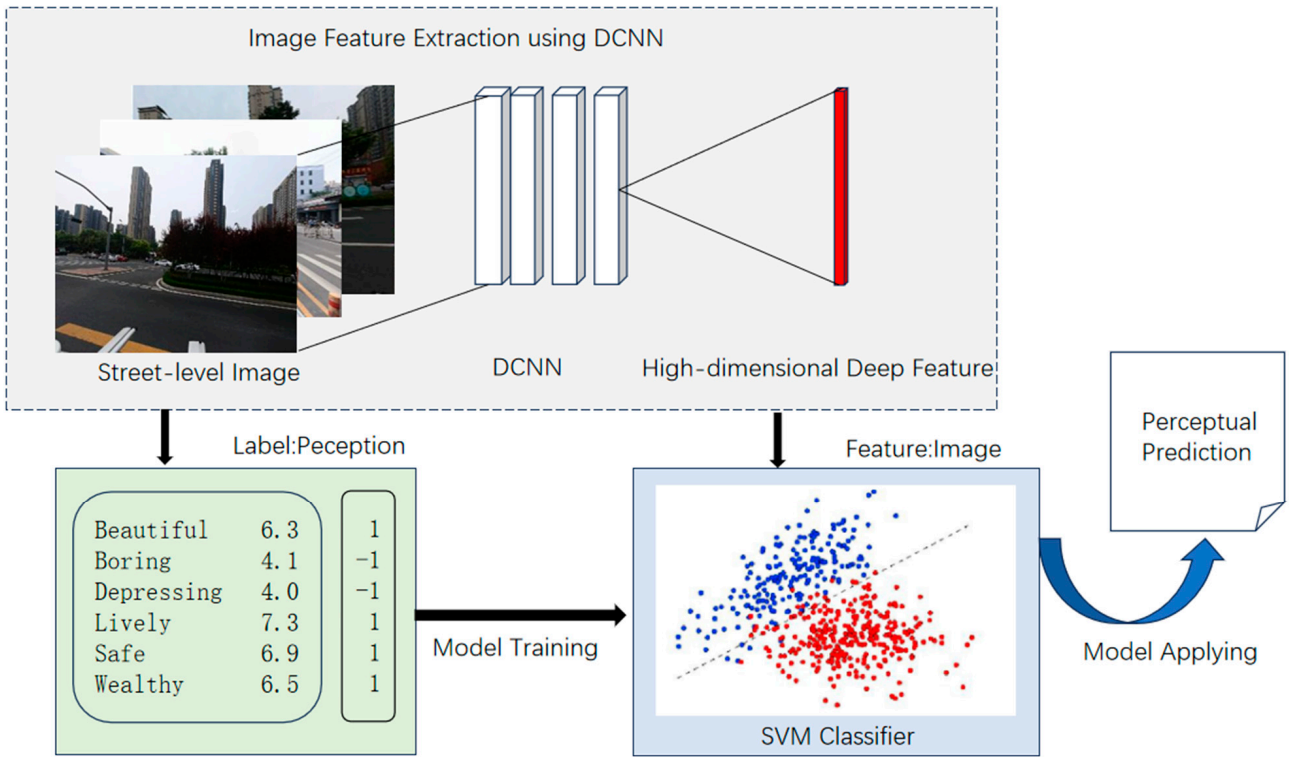


Figure 7. Deep learning model training and prediction.

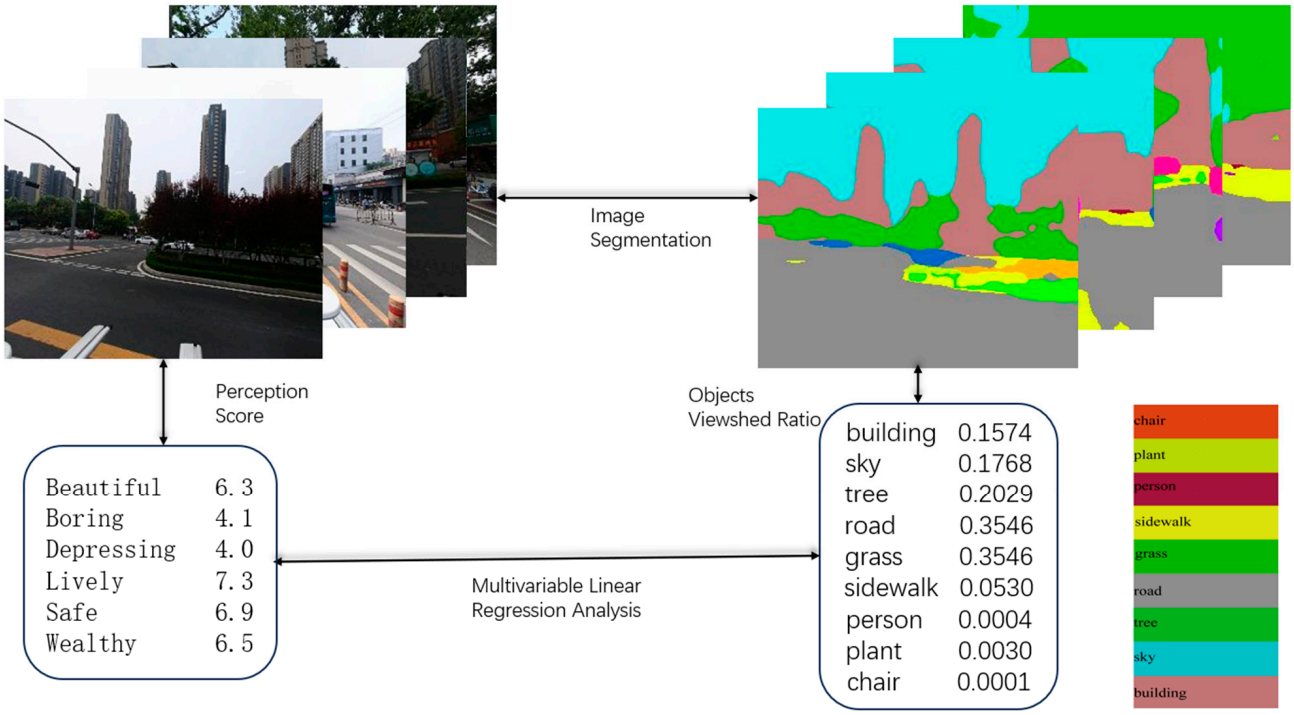


Figure 8. Overview of multiple regression analysis.

2.7. Urban Color Spatial Transformation and Visualization Technology

In practical applications, images are predominantly represented in the RGB color space. However, if an image in the Lab color space is to be acquired, a relative conversion process is necessary. In this study, the process of converting the RGB color space to the Lab color space primarily comprises the following steps: (1) The street scene image is loaded and its RGB vector is calculated. (2) The K-means clustering algorithm is then applied to cluster the RGB vectors. (3) The RGB color vectors of the street scene image colors are then converted into the Lab color space. (4) The results of the color analysis are output to a CSV file. This file should contain the following information: image path, color name, Lab value, and pixel percentage.

Color is the subjective experience of the human eye in response to different frequencies of light, which exist objectively and can be perceived subjectively by people [28]. The visualization of color spatial distribution is conducted in three principal stages. Initially, color identification and classification are performed. The street spots are clustered into 20 different color categories based on K-means clustering, which extracts color information. Subsequently, the study area is delineated by the creation of fishing nets. Scale is a pivotal consideration in landscape ecology [38]. This study employs a fishing net tool to assess the scale effects of landscape patterns, with fishing net block sizes spanning from 100 m × 100 m to 10 km × 10 km. A 160 m × 160 m net is selected as the unit of analysis following a comparison of the experimental results obtained with different net sizes. This is performed in order to ensure that the net contains as much color feature information as possible and is able to cover the entire area of Lixia District. The final stage of the process involves marking, calculation, and visualization. The marking tool is employed to align the street spots with the nets, and the resulting data are exported. The highest percentage of colors within each net is then calculated and visualized on the map with the assistance of the symbolization tool.

2.8. Color Metrics Calculation

2.8.1. Color Complexity Metrics

The color complexity of an image is typically quantified by the number of colors present. However, several studies have not considered the number of specific colors in an image (aspect ratio) and not measured the different colors in the whole image, resulting in some degree of inaccuracy. To more accurately assess the overall color complexity of a street scene image, this study draws on the HHI index in economics to extract the n colors from the image and calculate their complexity, as shown in the following equation [26].

$$C_c = \left| \left(\sum_{i=0}^{n} S^2 \right) - 1 \right| \quad (7)$$

The variable C_c denotes color complexity, while n represents the number of primary colors extracted from the streetscape image. The variable S denotes the proportion of the image occupied by each color, with color complexity taking a value between 0 and 1. A lower value indicates a more simplified color composition, whereas a higher value indicates a more complex one.

2.8.2. Color Coordination Indicators

While color coordination is typically founded upon the principles of the color wheel and color theory, it can be defined as a measure of the visual coherence of a given color combination. For example, color combinations that are similar, complementary, or separated complements are typically regarded as harmonious. To quantitatively assess the consistency of a color distribution, the mean and standard deviation of the colors in the Lab

color space can be calculated. This method is specifically concerned with determining the degree of concordance in the Lab color space to evaluate the degree of color coordination.

$$C_h = \frac{\sigma_L + \sigma_a + \sigma_b}{\frac{1}{3}\left(\overline{L} + \overline{a} + \overline{b}\right)} \tag{8}$$

In this context, C_h represents the degree of coordination of the color, while ($\overline{\mathrm{L}}$, $\overline{\mathrm{a}}$, $\overline{\mathrm{b}}$) denotes the mean value of the color data, which encapsulate the overall color characteristics of the data set in the Lab space. Furthermore, (σ_{L}, σ_{a}, σ_{b}) signifies the degree of dispersion of the color data, which is calculated by comparing the total fluctuation of the color (the sum of the standard deviations) with the average color level. A lower degree of harmonization indicates a smaller difference between colors, resulting in a more harmonious color combination. Conversely, a higher degree of harmonization corresponds to a greater difference between colors, creating a less harmonious color combination.

2.9. Spatial Analysis Methods

2.9.1. Spatial Autocorrelation Analysis

The term "spatial autocorrelation" is used to describe the potential correlations between specific variables and observed data within the same geographical area. This analysis is primarily concerned with the distribution status of a spatial attribute feature and its change trend, with the objective of assessing the overall distribution pattern of the attribute on a regional scale. This approach allows for a more accurate understanding and description of the spatial characteristics of geographic elements and their interactions. If an observation is consistent or similar to neighboring regions, it indicates the existence of a positive spatial correlation. Conversely, if there are obvious differences or dissimilarities, they are regarded as negative spatial correlations. The global spatial autocorrelation can reflect the overall spatial characteristics of a region, and its calculation formula is as follows:

$$\text{Global Moran's I} = \frac{n\sum_{i=1}^{n}\sum_{j=1}^{m}\mathcal{W}_{ij}\left(\mathcal{X}_i - \overline{\mathcal{X}}\right)\left(\mathcal{X}_j - \overline{\mathcal{X}}\right)}{\sum_{i=1}^{n}\sum_{j=1}^{m}\mathcal{W}_{ij}\sum_{i=1}^{n}\left(\mathcal{X}_i - \overline{\mathcal{X}}\right)^2} \tag{9}$$

In the provided equation, n represents the total number of regional units, $\mathcal{W}_{ij}$ is the weight matrix, and $\left(\mathcal{X}_i - \overline{\mathcal{X}}\right)$ and $\left(\mathcal{X}_j - \overline{\mathcal{X}}\right)$ represent the discrepancy between the observed values of individual spatial units within a given regional unit and the overall mean, respectively. To ascertain the significance of the global autocorrelation index, the standardized Z-value is typically employed as a measure, which is calculated in the following manner:

$$\mathrm{Z_{score}} = \frac{\mathrm{I} - \mathrm{E(I)}}{\sqrt{\mathrm{VAR(I)}}} \tag{10}$$

In this context, E(I) represents the expected value of autocorrelation, VAR(I) is the variance, and the Z-score is employed as a threshold for the standardized statistic, which is utilized to quantify the extent of discrepancy between the actual and expected values.

The global Moran's I index assumes a value range of [−1,1], defined as Moran's I > 0 ($p < 0.05$), which indicates the presence of a positive spatial correlation of residents' emotions. The magnitude of this correlation is inversely proportional to the value, with a larger value indicating a more pronounced spatial correlation. Conversely, Moran's I < 0 ($p < 0.05$) signifies the existence of a negative spatial correlation of residents' emotions, whereby a smaller value corresponds to a greater spatial variability. When Moran's I is equal to or nearly equal to zero, it indicates the absence of spatial autocorrelation in residents' emotions in neighboring areas, which are, instead, distributed randomly.

The global spatial autocorrelation analysis was employed to assess the degree of the autocorrelation of densities across the survey area. However, due to the differences in the degrees of autocorrelation between individual spatial units in the survey area and their

neighboring areas, the global analysis was unable to elucidate these subtle local variations in sufficient detail. Consequently, in order to more accurately comprehend the local characteristics of spatial autocorrelation, it was necessary to conduct a further local spatial autocorrelation analysis. The objective of this local spatial autocorrelation analysis was to investigate the correlations between each spatial location and its neighboring locations within the survey area with respect to a specific attribute. This analysis method allows for the accurate identification of spatial differences due to spatial correlation, as well as the delineation of spatial hotspots and high-incidence areas for a given attribute. This enables a more comprehensive understanding of the characteristics and patterns of spatial distribution. In this study, the Getis–Ord Gi* method (cold spot and hot spot analysis method) will be employed for a detailed analysis. Its calculation formula is as follows:

$$\text{Localized Moran's I} = \frac{(\mathcal{X}_i - \overline{\mathcal{X}})\sum_{j=1}^{m} \mathcal{W}_{ij}(\mathcal{X}_j - \overline{\mathcal{X}})}{\frac{1}{n}\sum_{i=1}^{n}(\mathcal{X}_i - \overline{\mathcal{X}})^2} \tag{11}$$

2.9.2. Least Squares and Geographically Weighted Regression Models

In the field of regression analysis, traditional ordinary linear regression models tend to presuppose global consistency among variables. However, this approach ignores the local heterogeneity or region-specific characteristics that variables may exhibit. While Ordinary Least Squares (OLS) can provide reasonable and effective estimates for quantifying the effects of color complexity and coordination on affective scores in many statistical contexts, with corresponding regression coefficients and significance levels, the inherent heterogeneity and spatial autocorrelation of spatial data limit the applicability of OLS. In contrast, geographically weighted regression (GWR) is capable of capturing the heterogeneity of spatial data, thereby being a powerful tool for studying spatial nonstationarity. The GWR model reveals the local spatial relationships at a specific scale and the driving factors. This model incorporates the geographic location of the data as a component of the regression parameters and considers the impacts of spatial weights in a comprehensive manner. In this study, we employed both the OLS and GWR models to calculate the color complexity and coordination indexes and the six categories of emotional perception values. These calculations were based on the six categories of emotional perception values derived from related studies. We then conducted a regression analysis on the explanatory variables and the explained variables. The aim of this analysis was to explore the relationship between the color complexity and coordination of streetscapes in the Lixia District of Jinan City and the emotional perceptions of its residents. The calculation formula is as follows:

$$Y_i = \beta_0(\mathcal{U}_i + \mathcal{V}_i) + \sum_k \beta_k(\mathcal{U}_i + \mathcal{V}_i)\mathcal{X}_{ik} + \varepsilon_i \tag{12}$$

In this model, Y_i denotes the affective perception score of cell i; $(\mathcal{U}_i, \mathcal{V}_i)$ represents the horizontal and vertical coordinates of the spatial location of cell i; $\beta_0(\mathcal{U}_i + \mathcal{V}_i)$ signifies the model regression intercept of change; $\beta_k(\mathcal{U}_i + \mathcal{V}_i)$ is the regression coefficient of the explanatory variables; $\mathcal{X}_{ik}$ is the kth explanatory variable on cell i; and ε_i is the random error term of the model.

2.10. Model Validation and Evaluation

In order to guarantee the precision of the model prediction outcomes, this study employed a deep learning analysis of urban streetscape image data to select 30 representative sampling points for field validation. These sampling points were selected to encompass a diverse range of streetscape environments and color characteristics. The selected sampling points exhibited a high building density, frequent street activities, and complex color distributions (see Figure 9. for examples of the sampling points). The sampling points were chosen to ensure diversity and differences in the visual characteristics of the streets, thereby guaranteeing the inclusion of various urban streetscapes.

At each sampling point, field research and interviews with local residents were conducted in order to capture the color characteristics of the streetscape, as well as residents' emotional responses to the streetscape (Figure 10). The process was largely consistent with the emotional scores predicted by the color model, and was mainly used to compare the differences between the model's predictions and residents' actual emotional responses.

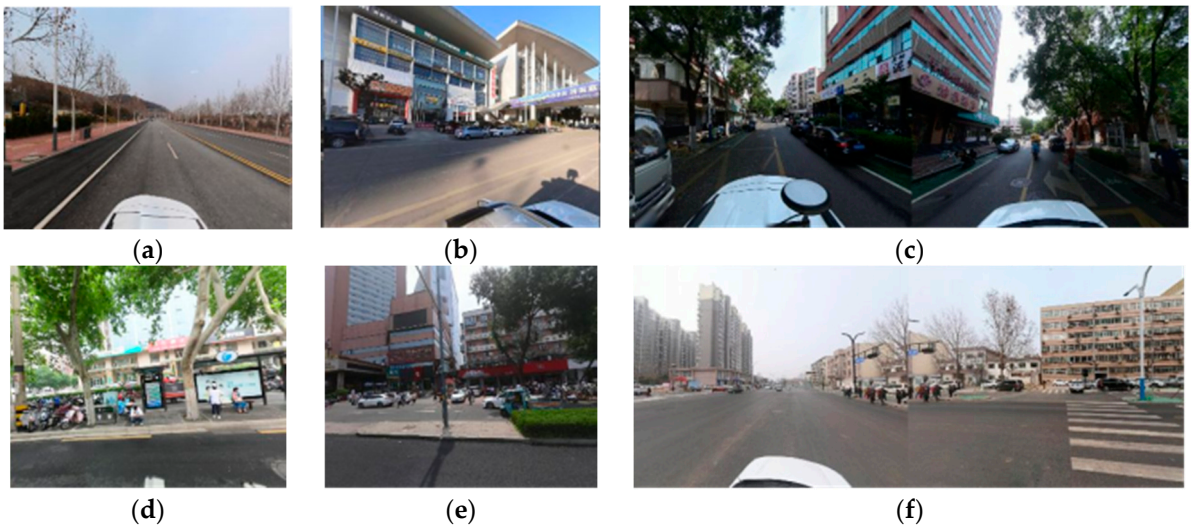


Figure 9. Examples of sampling points. (**a**) View of the city's main street, showing the wide road and tree-lined surroundings. (**b**) A view of the traffic outside a shopping centre, showing the bustling activity of the city. (**c**) A view of a shopping street showing a variety of shops and pedestrians on the street. (**d**) Street view near a public transport station showing the convenience of city life. (**e**) Streetscape of a commercial area showing a busy scene with a mix of pedestrian and vehicular traffic. (**f**) Panoramic view of an urban intersection showing traffic lights and road layout.



Figure 10. Examples of data collection and forecasting.

The results of the validation demonstrate that commercial areas characterized by a high degree of color complexity, despite their strong visual appeal, were reported by some residents to evoke a sense of visual fatigue. This finding aligns with the conclusions of the study. In contrast, residents in areas of a high color coordination in residential zones typically reported a high sense of security and comfort, thereby further substantiating the positive effect of color coordination on emotional perception. This approach allows us to more accurately validate the model, particularly in the context of complex streetscapes and the emotional perceptions of residents.

3. Experiments and Results

3.1. City Color Analysis

In this study, color information was extracted based on the K-means clustering algorithm, and the color type with the highest percentage in each fishing net was calculated. This resulted in the creation of a color spatial distribution map of Lixia District (Figure 11). The main color types identified were gray and blue-green. In particular, gray was the

predominant color in the central region of Lixia District, exhibiting a notable concentration trend. Conversely, bright blue-green was present but more sporadic in distribution, failing to form an evident concentration area. A further analysis of the data reveals that the color gray was concentrated in the center and west of Lixia District. This is due to the presence of roads, businesses, and administrative offices in these areas, such as the Quancheng Road business district, the Daming Lake neighborhood, and governmental agencies. Additionally, buildings in these regions are predominantly constructed with gray cement, glass, and steel structures. Similarly, roads in these areas are primarily made of gray material. Collectively, these factors contribute to the formation of the gray tone observed in the central city. The blue-green color was predominantly observed in parks, squares, and residential areas, particularly in the green-rich northeastern and southern regions of the city and in some commercial areas. This reflects the positive outcomes of urban greening and ecological environment construction.

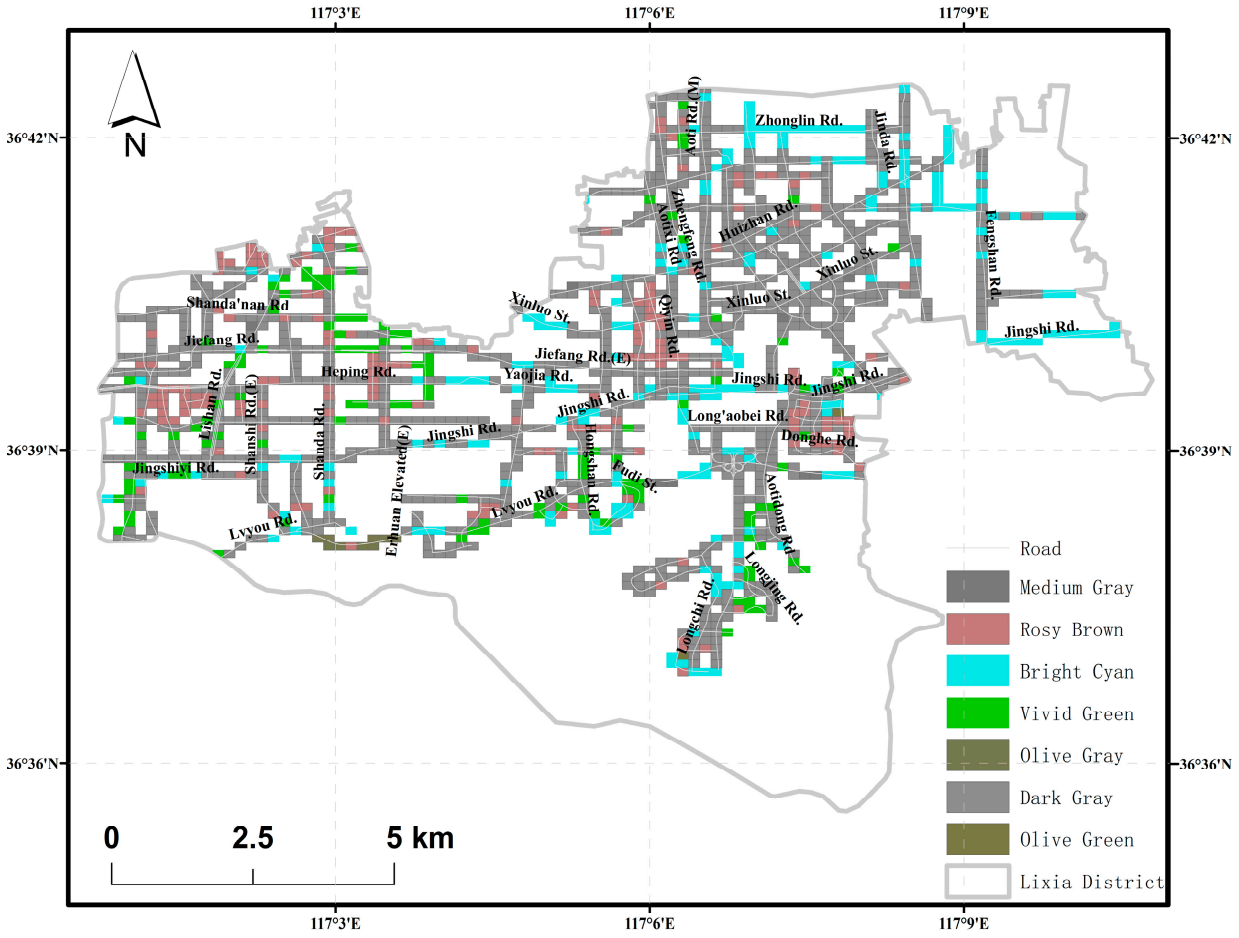


Figure 11. Color distribution map of Rixia District.

The complexity and coordination formulas were employed to conduct a quantitative assessment of the color complexity and coordination of Lixia District, with the resulting values being presented in the form of a spatial visualization (Figure 12). The findings indicate that the majority of areas within Lixia District exhibit elevated levels of color complexity and an overall low color coordination. In particular, a high level of color complexity is indicative of a visually rich environment characterized by a diverse and contrasting palette of colors. The areas exhibiting the highest degree of color complexity were concentrated in the city center and in the vicinity of major traffic arteries. This is due to these areas' high density of buildings, the prevalence of commercial activities, and the abundance of advertisements, signboards, and other visual elements, which collectively contribute to

a diverse and dynamic visual environment. The color coordination was generally low, indicating a lack of visual cohesion in many areas. Conflicting color combinations were prevalent, particularly in the city center and commercial zones, where the prevalence of commercial advertisements and diverse architectural styles impairs color coordination. This phenomenon is further compounded by the use of different colors for buildings along major traffic routes.

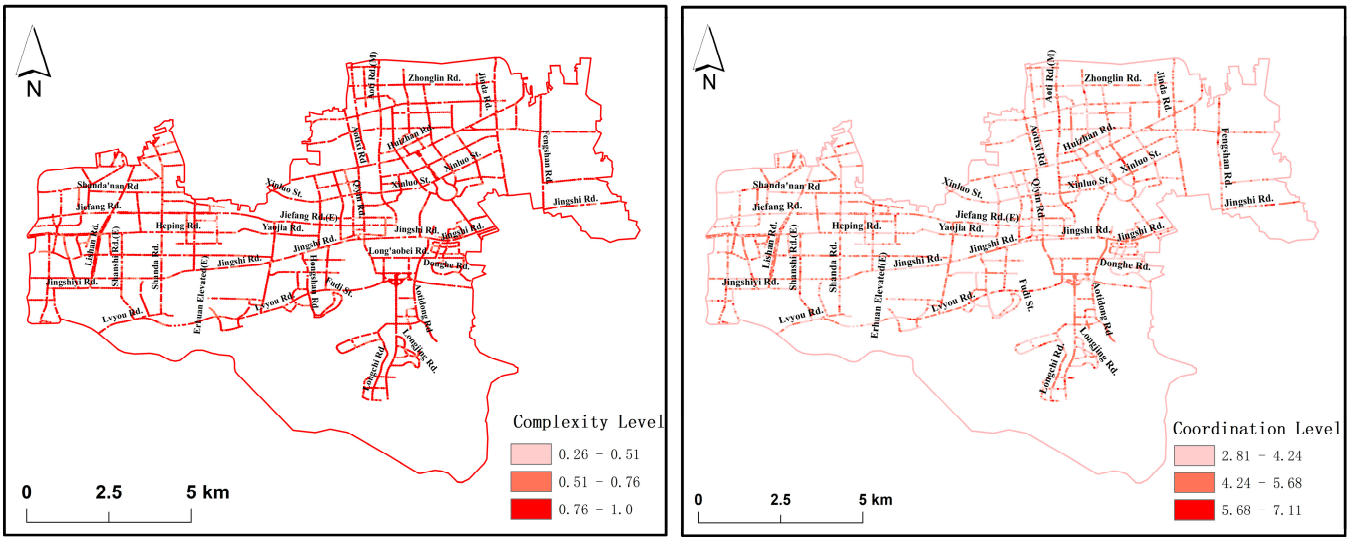


Figure 12. Visualization of color complexity and coordination.

3.2. Emotional Perception Analysis

3.2.1. Emotional Relevance Analysis

The Pearson correlation analysis of the six categories of emotional indicators in this study demonstrated a statistically significant correlation between two of the emotional perceptions (Figure 13). The data revealed a highly positive correlation between beauty and richness and a highly negative correlation between beauty and depression. Boredom demonstrated a moderately positive correlation with depression, a moderately negative correlation with safety and wealth, and a strong negative correlation with liveliness. Depression exhibited a strong negative correlation with safety and wealth, and a moderately negative correlation with liveliness. Liveliness displayed a moderately positive correlation with safety and wealth, while depression showed a moderately negative correlation with these two variables. Safety and wealth demonstrated a strong positive correlation. Classifying beauty, liveliness, safety, and richness as positive affective perceptions and boredom and depression as negative affective perceptions, it can be observed that residents perceived positive and negative perceptions in a similar manner [26].

To more intuitively demonstrate the intensity of residents' perceptions of the six distinct emotions, the intensity of each emotion perception was divided into five levels according to the principle of equal intervals, and six different colors were utilized to differentiate the six types of emotional perceptions (Figure 14). The residents' emotional perceptions exhibited a notable spatial distribution pattern. The spatial distribution of the perceptions of beauty and liveliness was analogous, originating from the center and exhibiting a radial diffusion trend from weak to strong. The perceptions of safety and wealth formed a "low–high–low" circular distribution pattern around the center. The perceptions of boredom were distributed gradually from west to east. The perception of boredom gradually increased from west to east, forming a gradient of "low–high". In contrast, the perception of depression was relatively scattered and, overall, similar to the perception of boredom, showing a gradient of "low–high–low" from west to east.

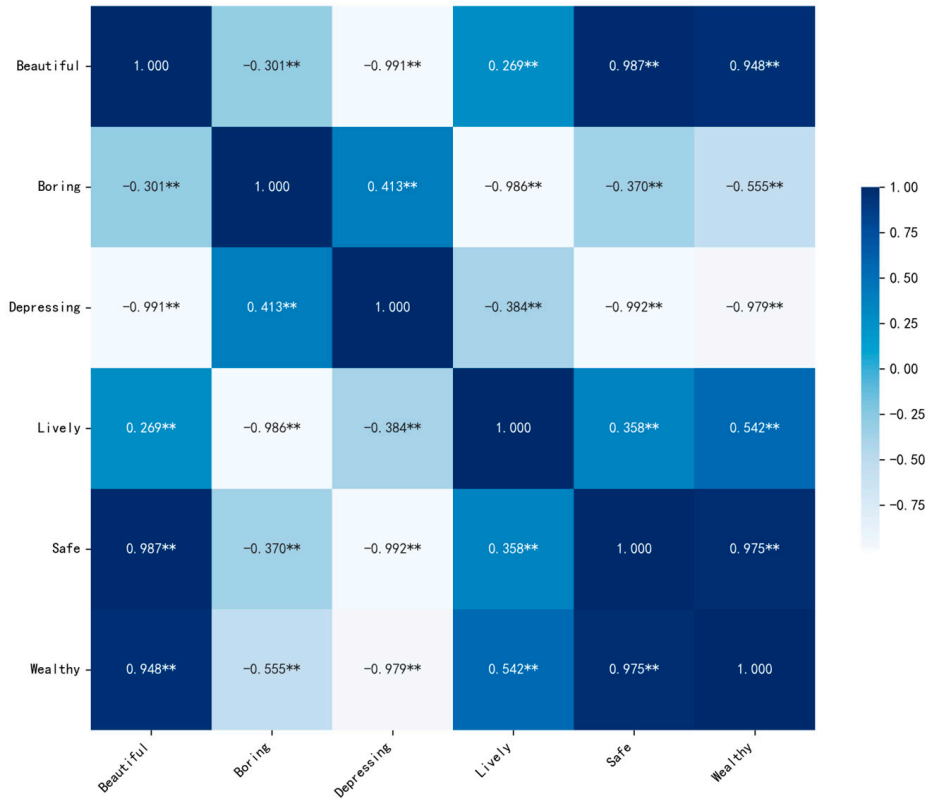


Figure 13. Pearson correlation coefficients among the 6 perceptual indicators. **. Correlation is significant at the 0.01 level (two-tailed).

A detailed examination of the actual situation reveals that Lixia District, as the central area of Jinan, is characterized by a high population density, robust commercial activities, and a considerable volume of traffic, which collectively contribute to a heightened prevalence of depression among residents in the central area. However, as the distance from the central area increases, the living environment of residents becomes relatively relaxed. This is accompanied by increases in green space, natural landscape, and public facilities, which contribute to a gradual decrease in the perception of depression and an increase in the perceptions of liveliness and beauty. Coupled with the high pressure of security management in the central area of Lixia District, these factors contribute to a more positive overall perception of the area. In Lixia District, the sense of security is low, and this extends outward to the residential area and high-grade community, where the security situation is better and residents' sense of security and affluence are significantly increased. This extends further outward to urban–rural areas or slower-developing areas where law and order management and infrastructure may be insufficient and the sense of security and affluence decrease once more. In the western area, due to its proximity to the old city and major commercial areas, which offer a plethora of living and recreational facilities, residents exhibited a low level of boredom. However, as the area expands to the east and gradually enters newly developed zones with a limited number of commercial establishments, residents displayed a notable increase in boredom.

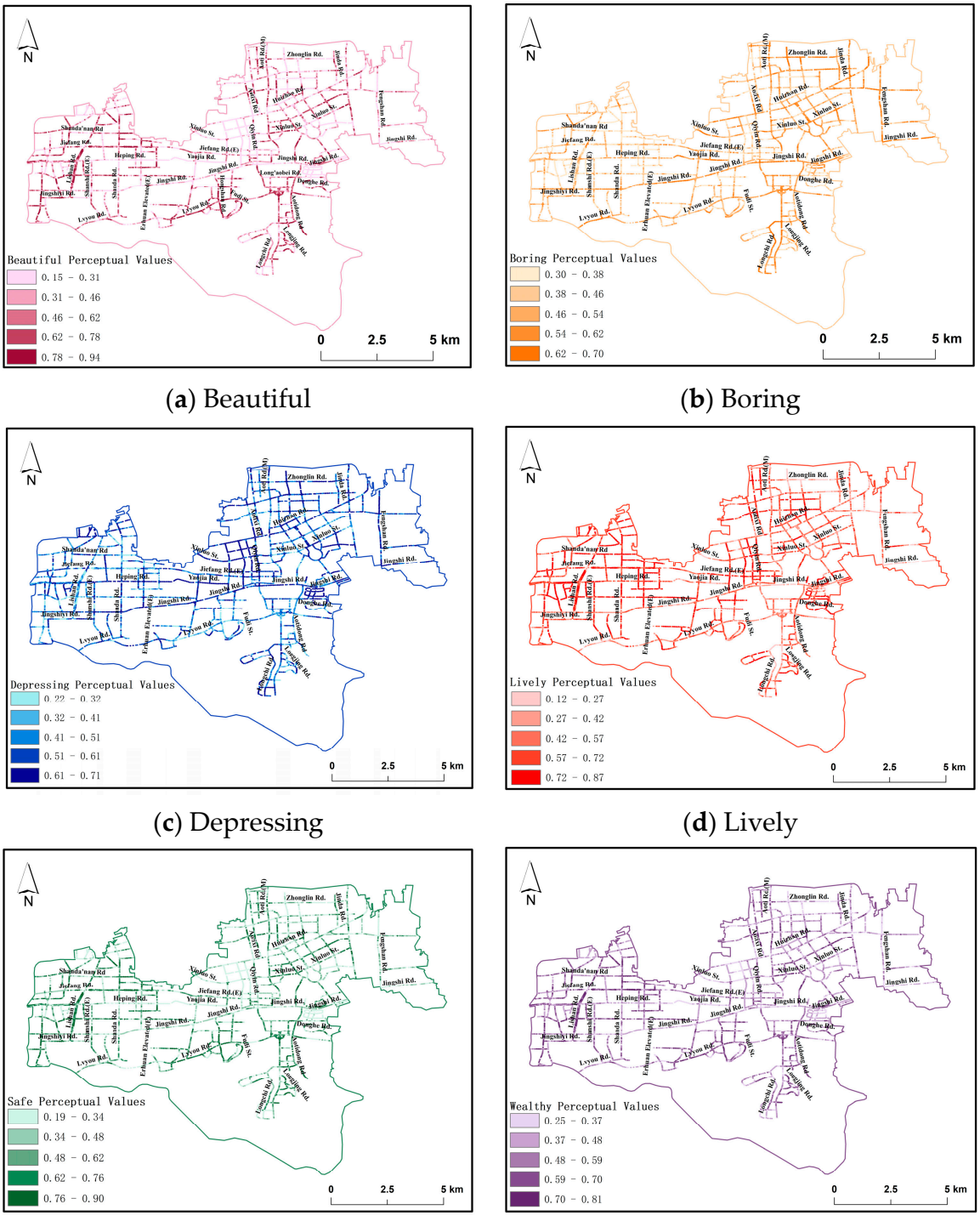


Figure 14. Spatial distribution of emotional perception.

3.2.2. Spatial Autocorrelation Analysis of Emotions

As any elements in space are connected to each other, the correlation is stronger for elements that are in closer proximity and weaker for elements that are farther away from one another [39]. This concept was applied to the mapping of various types of emotional perceptions to streets, where the spatial agglomeration of these perceptions and their spatial correlations were analyzed using the global Moran's index method (Table 1). The spatial distribution of these emotional variables and their significance can be evaluated using

the Moran's I index, Z-value, and *p*-value, as indicated in the table. The *p*-value for all emotional variables is 0.000, indicating that the spatial autocorrelation is significant at the 0.01 level of significance. In sum, the distribution of these affective variables in geospatial space is not random, but rather exhibits a regular aggregation. This clustering pattern is unlikely to occur randomly, with a probability of less than 1%. This provides strong empirical support for further affective spatial research.

Table 1. Spatial autocorrelation and high/low clustering analysis.

Emotional Perception	Spatial Autocorrelation Statemen			High/Low Clustering Analysis Statement		
	Moran's I	Z	P	General G	Z	P
Beautiful	0.549	49.162	0.000	0.000022	19.866	0.000
Boring	0.667	59.697	0.000	0.000019	8.569	0.000
Depressing	0.556	49.845	0.000	0.000018	0.650	0.516
Lively	0.662	59.327	0.000	0.000019	9.550	0.000
Safe	0.545	48.816	0.000	0.000020	16.990	0.000
Wealthy	0.552	49.507	0.000	0.000019	12.804	0.000

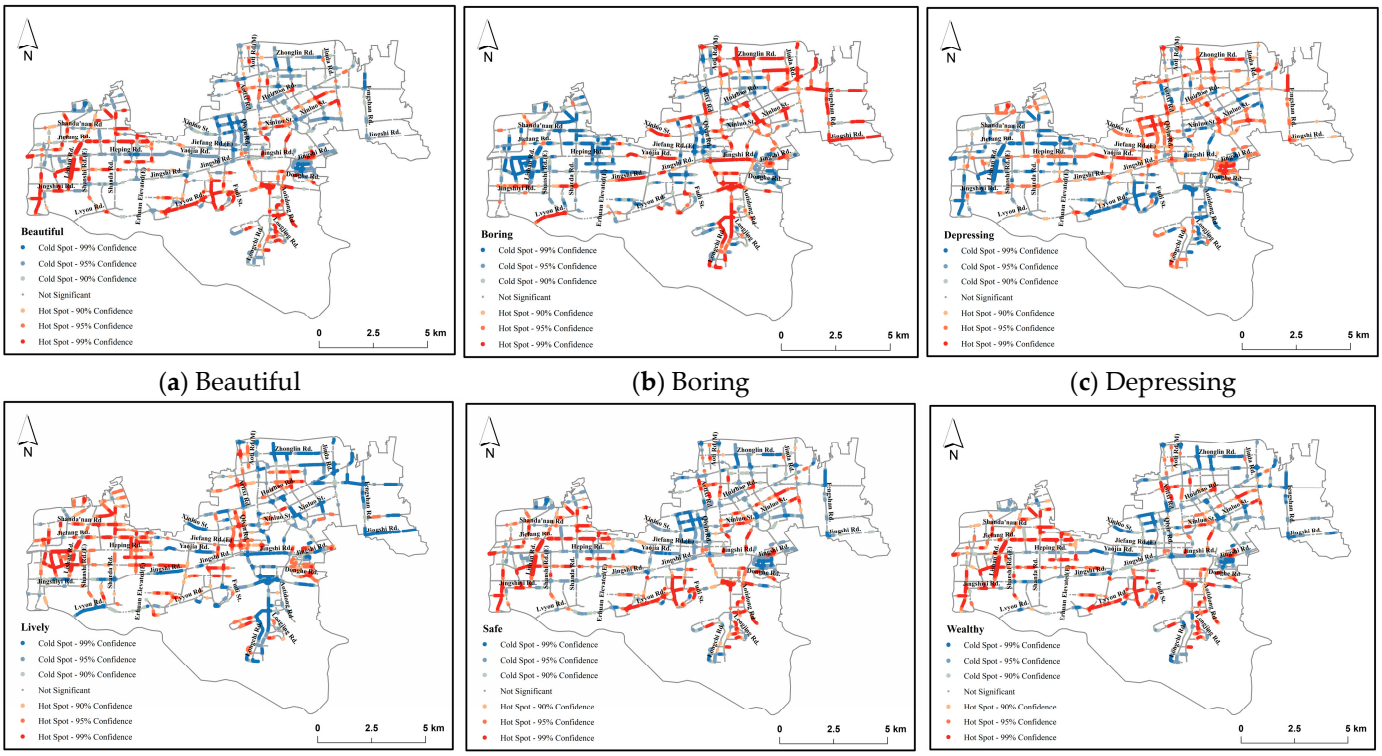
3.2.3. High/Low Cluster Analysis

The Moran's I index is unable to ascertain whether the spatial data are clustered with high or low values. In 1992, Getis and Ord proposed General G coefficients, which are similar to the global Moran's I. These coefficients test the statistical significance of spatial autocorrelation using Z-scores. However, there is a distinction between the two approaches, as follows: a positive Z-score implies the existence of high/high clustering, while a negative Z-score indicates low/low clustering [40].

The analysis of the spatial distribution of emotional perceptions using high/low clustering analysis (Table 1) demonstrated that distinct emotional indicators exhibited disparate clustering characteristics in the spatial distribution. Perceptions of beauty, liveliness, safety, boredom, and affluence demonstrated elevated clustering values. In contrast, no significant clustering of depressing perceptions was observed within Lixia District, suggesting that these perceptions were more evenly distributed across the areas within the district. The impacts of environmental factors (e.g., traffic congestion, building density, etc.) on the perception of repression varied across different areas, exhibiting a more dispersed spatial pattern than the concentrated distribution observed for other affective perceptions.

3.2.4. Hot Spot Analysis

The application of a hotspot analysis allows for further delineation of the specific locations where the cold and hot spots of various types of emotional perceptions occurred spatially (Figure 15). The distribution of cold and hotspot aggregations of positive and negative emotions exhibited a similar pattern, with the hotspot areas of positive emotional perceptions being concentrated in the area of Shunqiao Road, the area of Jiefang Road, the area of Daming Lake Road, and the area of Tourism Road. The positive affective perceptions of residents and tourists are enhanced as a result of the high-quality urban planning, sufficient greenery, perfect public facilities, and natural and cultural landscapes. In contrast, the negative affective perception hotspots were primarily situated in areas with high traffic and industrial activities in Lixia District, including the central portion of Jingshi Road, Xinluo Street, Jiefang East Road, and the northeastern section of South Industrial Road. Notwithstanding the robust economic activities of these areas, their high levels of commercial, transportation, and industrial operations, the pervasive noise and air pollution, the suboptimal environmental quality, and the inadequate greening and landscaping design lead residents and tourists to exhibit diminished positive perceptions and heightened negative emotions.



(**a**) Beautiful (**b**) Boring (**c**) Depressing

(**d**) Lively (**e**) Safe (**f**) Wealthy

Figure 15. Distribution of perceived hot and cold spots.

3.3. Analysis of Color Indicators and Emotional Perception

3.3.1. Effect of Color Complexity–Coordination on Emotion Perception

The correlation analysis of color complexity and coordination with various types of emotional perceptions (Table 2) reveals the following: color complexity is positively correlated with boredom and depression, negatively correlated with liveliness and richness, and exhibits no significant correlation with safety and beauty, and color coordination has a significant positive correlation with beauty, safety, and richness, a negative correlation with depression, and exhibits no significant correlation with boredom and liveliness. The emotional experiences of individuals are, to some extent, influenced by the characteristics of colors. An increase in color complexity has been observed to correlate with increases in feelings of boredom and depression, while a high color coordination has been linked to enhanced perceptions of beauty, safety, and richness.

Table 2. Color–emotion perception correlation analysis.

	Beautiful	Boring	Depressing	Lively	Safe	Wealthy
Color Complexity	−0.016	0.231 **	0.041 **	−0.186 **	−0.006	−0.029 *
Color Coordination	0.188 **	0.003	−0.178 **	−0.017	0.180 **	0.156 **

**. Correlations are significant at the 0.01 level (two-tailed). *. At the 0.05 level (two-tailed), the correlation is significant.

A Random Forest Regression (RFR) analysis was conducted to investigate the potential impacts of streetscape color complexity and coordination on residents' emotional perceptions. The assessment of the relative importance of the features of the random forest indicated that the total relative contribution of the two to emotional perceptions was normalized to 100%. The results revealed slight differences in the sensitivity of each emotion to color complexity and coordination (Figure 16). Specifically, color coordination was more effective than color complexity in shaping residents' perceptions of beauty. In contrast, color complexity played a more significant role in stimulating the emotion of

boredom. Additionally, color coordination had a stronger influence on the formation of repressed emotions, while color complexity was more critical in stimulating lively emotions. Furthermore, color coordination similarly surpassed the influence of complexity on both safe and rich perceptions.

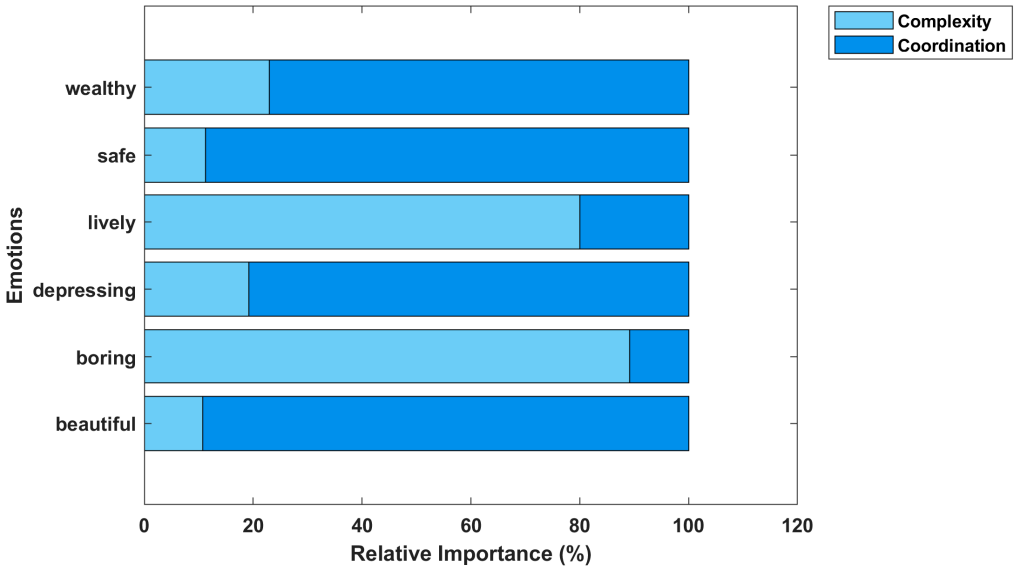


Figure 16. Importance scores of color metrics on emotional impacts.

The assessment of feature importance determines the significance of color complexity and coordination in emotional perception. The results of deep learning are made more intuitively interpretable by further plotting partial dependency graphs to demonstrate their marginal effects [26]. As illustrated in Figure 17, the effects of both on the six categories of emotions exhibit complex nonlinear patterns.

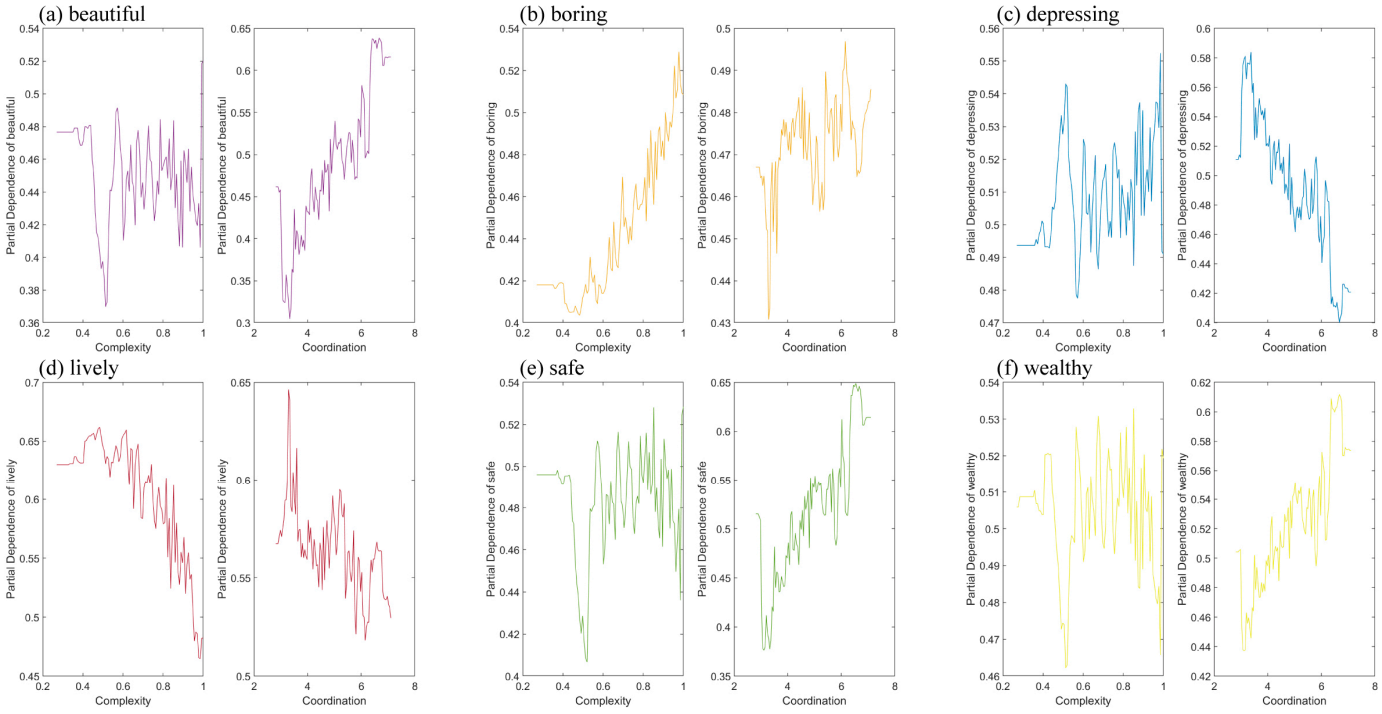


Figure 17. Bias dependence of color complexity and coordination on perception.

(1) The perception of beauty increases and then decreases with color complexity, reaching a peak at 0.2 and then decreasing to the lowest at 0.52, with the remaining values exhibiting significant fluctuations. The color coordination degree is in a "V" shape, with a peak near 7.5 and a slight decrease.

(2) Boredom is observed to be relatively stable and low at lower levels of color complexity. However, it is noted to fluctuate with an increase in complexity after 0.4. Additionally, color coordination is observed to fluctuate within the range of 4–8, yet it is noted to be stable overall.

(3) The sense of depression remains stable at lower complexity levels, but exhibits significant fluctuations following an increase in complexity. The degree of coordination displays an "inverted V" pattern, with low values exhibiting a high volatility and high values demonstrating stability.

(4) The lively value demonstrates an "inverted U" type of decrease in relation to color complexity, with a peak value of approximately 0.45. The coordination degree is observed to be below three when there is a notable increase in growth, with fluctuations above three and currently at the trough.

(5) The formation of safe and emotionally rich knowledge is influenced by color complexity and coordination, exhibiting trends that are analogous to those observed in beauty. No significant differences are evident.

The effects of color complexity and color coordination on the perception of emotional states were analyzed (Figure 18). Darker areas represent lower emotional values, while lighter areas indicate higher emotional values [26]. The results demonstrated a significant correlation and synergy between the two variables in influencing emotional perception. The perception of boredom was found to be strongest when both color complexity and color coordination were simultaneously high. Conversely, there was a significant negative interaction of perceptions of beauty, safety, and richness with lower perceptions of these emotions when both were simultaneously high. Additionally, a significant positive interaction was observed for perceptions of depression, with the strongest perceptions of this emotion occurring when both color complexity and coordination were high. The effect on perceptions of liveliness was more complex, with perceptions of liveliness being strongest when color complexity was moderate and coordination was moderate, and weakest when complexity was high and coordination was low.

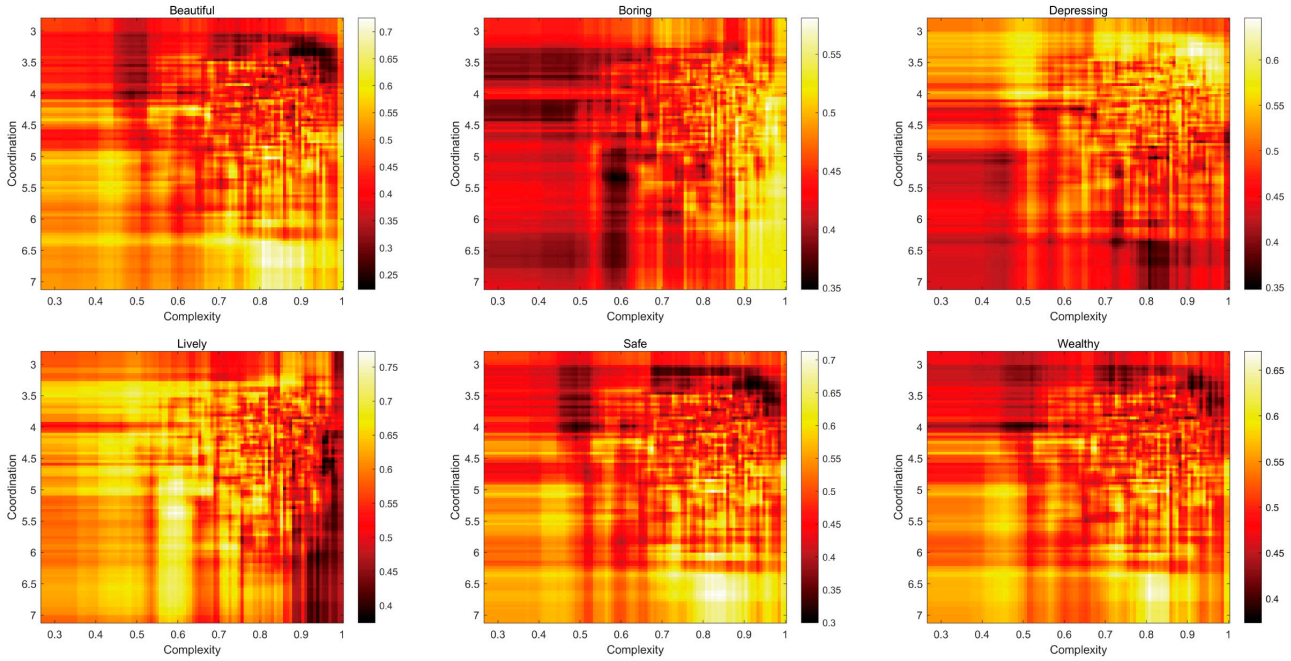


Figure 18. Color complexity–coherence interaction on perception.

Analysis of the bias dependence and interaction dependence graphs reveals that color complexity and color coordination exert complex non-linear effects on emotional perception, with a significant interaction between the two variables. The combination of different color features has the capacity to significantly modulate or enhance the perception of emotions, having positive and negative effects. A higher color complexity and coordination were associated with lower ratings of positive perceptions, including beautiful, safe, and rich. Conversely, a lower complexity and coordination were linked to higher ratings of positive emotions. A moderate complexity and coordination were linked to enhanced ratings of positive perceptions, including lively and beautiful. Furthermore, the effect of color features on perceptions of safety and richness followed a similar trend to that of perceptions of beauty, confirming the significant role of color in emotional expression.

3.3.2. OLS Regression Analysis Versus Geographically Weighted Regression Analysis

The emotional experiences of residents and tourists are significantly influenced by color complexity and coordination. However, OLS is unable to fully account for regional differences in response to color features. GWR analysis addresses this limitation and elucidates the spatial heterogeneity of color features on emotional scores. To gain a deeper understanding of the specific effects of color complexity and coordination on emotional perceptions and their spatial differences, this study employed both OLS and GWR models to conduct regression analyses with color complexity and coordination as independent variables, respectively (Table 3).

Table 3. OLS and GRW regression analysis.

		OLS							GRW	
		Ratio	**P**	**Robust_pr**	**Koenker (BP)**	**Adj R2**	**AICc**	**Jarque–Bera**	**Adj R2**	**AICc**
Beautiful	Complexity	−0.000481	0.988	0.988	0.024 *	0.03	−920.36	0.000 *	0.15	−1518.78
	Coordination	0.066143	0.000 *	0.000 *						
Boring	Complexity	0.230475	0.000 *	0.000 *	0.582	0.05	−9025.48	0.000 *	0.24	−10,058.43
	Coordination	0.00341	0.113	0.117						
Depressing	Complexity	0.034976	0.063	0.063	0.017*	0.03	−6152.9	0.000 *	0.15	−6769.57
	Coordination	−0.035787	0.000 *	0.000 *						
Lively	Complexity	−0.362119	0.000 *	0.000 *	0.316	0.04	−2569.97	0.000 *	0.23	−3600.65
	Coordination	−0.009735	0.021 *	0.023 *						
Safe	Complexity	0.017164	0.534	0.533	0.078	0.03	−2492.61	0.000 *	0.15	−3071.09
	Coordination	0.053784	0.000 *	0.000 *						
Wealthy	Complexity	−0.021679	0.259	0.257	0.028*	0.02	−5961.1	0.000 *	0.15	−6578.42
	Coordination	0.031954	0.000 *	0.000 *						

AIC is the Akaike Information Criterion, which measures the goodness of fit of a statistical model; R2 is the goodness of fit. Koenker (BP) is used to determine whether the explanatory variables of the model have a consistent relationship with the dependent variable in both geographic and data space; Jarque–Bera is a test of the skewness and kurtosis of the data, which evaluates whether the assumption that the given data follow a normal distribution with unknown mean and variance is held, with * indicating that the coefficients next to each other are statistically significant [41].

After accounting for spatial heterogeneity, the R^2 values of the GWR model were significantly higher than those of the OLS model. Additionally, the AICc values were all smaller than those of the OLS model. The GWR model demonstrated a superior performance to the OLS model in terms of the interrelationship between color and emotion, as well as in terms of the fitting effect. While the Koenker (BP) statistic was not significant for some emotional scores, this does not preclude the existence of spatial heterogeneity. The GWR model, which allows regression coefficients to vary spatially, was better able to capture this spatial heterogeneity and more effectively explain the role and spatial

variability of the influence between color complexity and coordination and emotional perception [41].

The results of the GWR model's coefficient statistics (Table 4) indicate that color complexity and coordination exert a notable influence on emotional perceptions, with considerable variability being observed. The complexity of color has a significant positive effect on perceptions of beauty, boredom, depression, security, and richness and a significant negative effect on perceptions of liveliness. Similarly, the coordination of color has a significant positive effect on perceptions of beauty, boredom, security and richness and a significant negative effect on perceptions of depression and liveliness. In particular, a high level of color complexity is likely to result in visual fatigue and confusion, as well as increased feelings of boredom and depression. Conversely, a moderate complexity can enhance environmental recognition and diversity, as well as a sense of security and wealth. Complex color schemes are often associated with luxury and affluence. However, excessive complexity has the opposite effect, diminishing the sense of liveliness. Color harmony positively contributes to perceptions of beauty and safety, enhances overall visual harmony and comfort, and reduces feelings of depression. For example, coordinated colors are often used in healthcare settings to alleviate patients' emotions.

Table 4. Coefficient statistics of emotion–color GWR analysis.

Emotional Perception	Color Complexity		Color Coordination	
	Average Value	Standard Deviation	Average Value	Standard Deviation
Beautiful	0.024	0.192	0.058	0.025
Boring	0.155	0.098	0.004	0.015
Depressing	0.008	0.109	−0.031	0.014
Lively	−0.215	0.190	−0.009	0.028
Safe	0.049	0.154	0.047	0.022
Wealthy	0.022	0.108	0.028	0.013

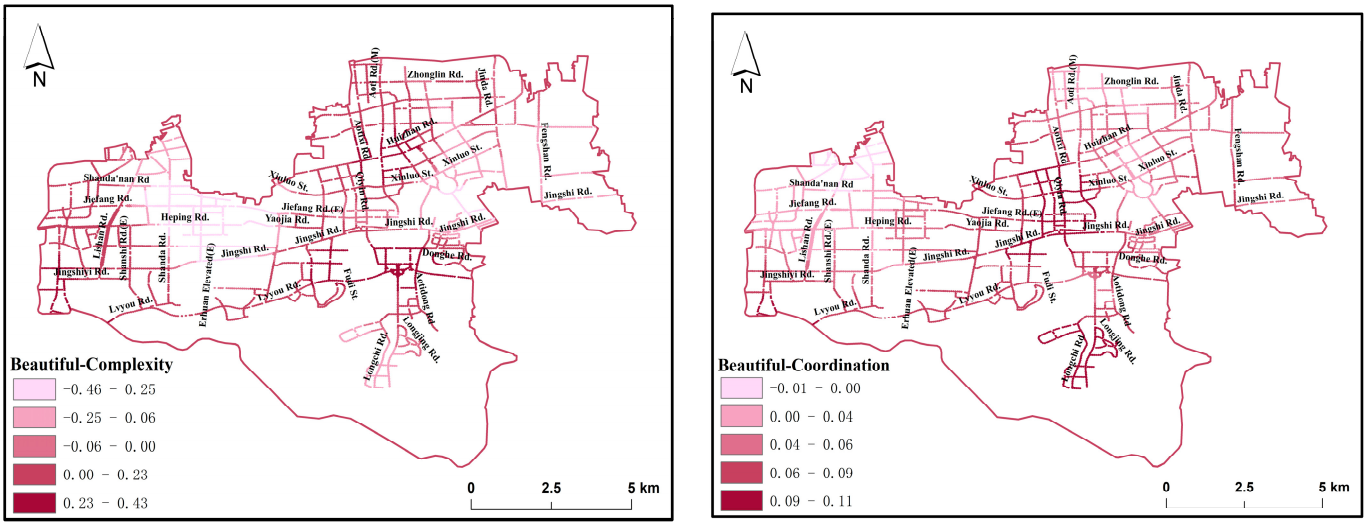
(a) The positive and negative signs of the GWR coefficients reflect the potential enhancing or inhibiting effects of color on emotion perception and (b) the standard deviation of the GWR measures the extent to which the different variables affect the differences in emotion perception, with a larger standard deviation indicating a greater difference in the degree of emotion perception.

To demonstrate the spatial variability of the regression coefficients in a more intuitive manner, the regression coefficients for color complexity and coordination with respect to each emotion score were visualized (Figures 19–24). The results indicated that color complexity and coordination exerted varying degrees of influence on different emotional perceptions in different types of city streets.

In the western area of Lixia District, which is dominated by Quancheng Road, Lok Yuen Street, West Wenhua Road, West Qianfushan Road, and Shunqiao Road, as well as in of Lixia District's central area, color complexity has a positive effect on the perception of beauty. In the area dominated by Heping Road, Jingshi Road, Century Avenue, and Longteng Road, the color complexity has a significant negative effect on the perception of beauty, particularly in the area dominated by Jiefang Road and Heping Road. Conversely, in the area south of Jingshi Road, the color coherence has a positive effect on the majority of the area, especially in the area south of Jingshi Road, where this positive effect is the most significant. In the area south of Jingshi Road, the presence of color coordination has a favorable impact on the majority of the area, particularly in the area south of Jingshi Road.

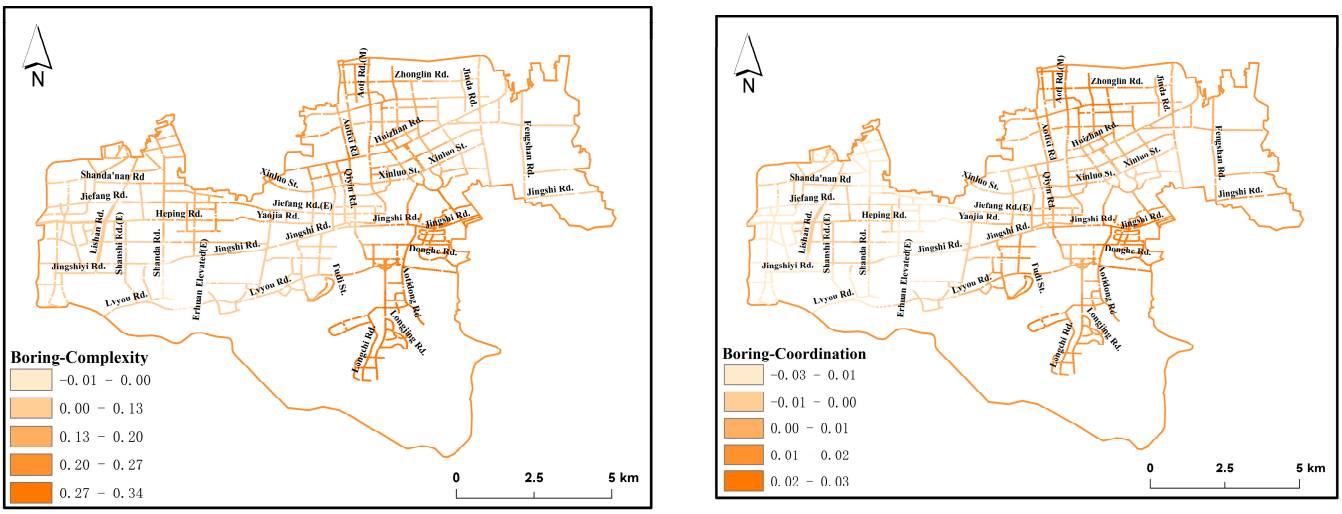
In the majority of locations within Lixia District, an increase in color complexity is associated with a reduction in boredom perception. This is particularly evident on Longding Avenue, where the lack of visual stimulation and monotonous environment contribute to this effect. In the area of Yanzi West Road and Kaiyuan Tunnel, the influence of color complexity is either weaker or even negative; the influence of color coordination on boredom perception demonstrates some spatial heterogeneity in Lixia District. In the area dominated by East Heping Road and East Jiefang Road, color coordination has a negative

effect on boredom perception. However, in the southeastern area, the positive effect of color coordination on boredom perception is most significant, because the color scheme in these areas is more harmonious and the environment as a whole gives people a pleasant visual experience, thus reducing boredom.



(**a**) Beautiful–Complexity (**b**) Beautiful–Coordination

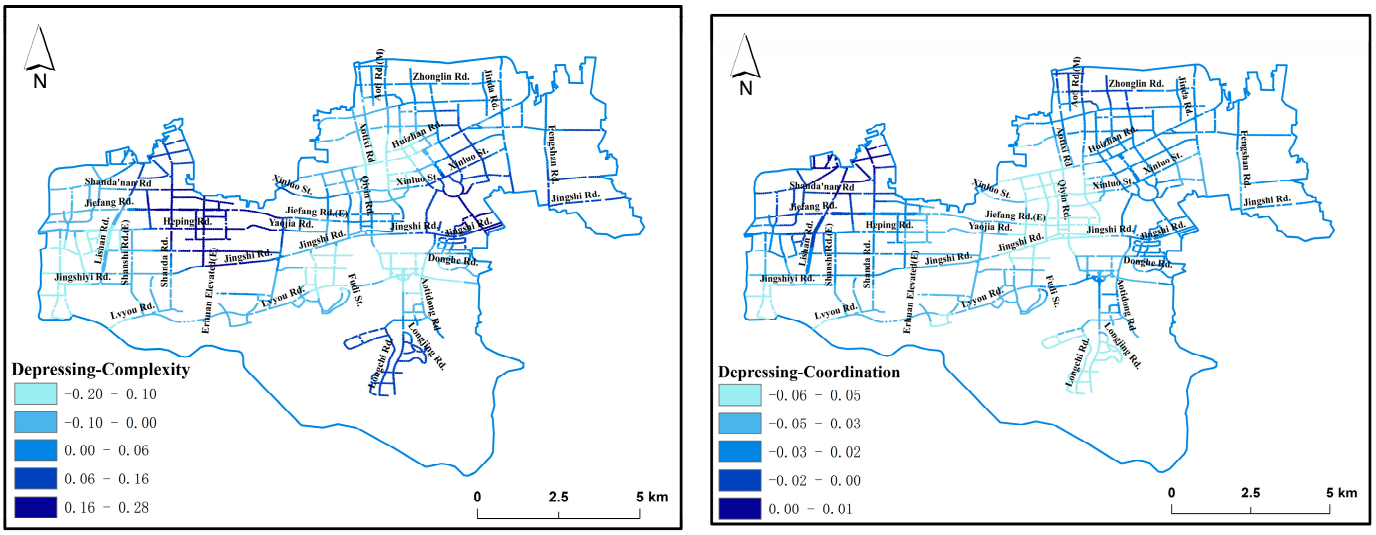
Figure 19. Visualization of GRW analysis of beautiful perception–colors.



(**a**) Boring–Complexity (**b**) Boring–Coordination

Figure 20. Visualization of GRW analysis of boring perception-colors.

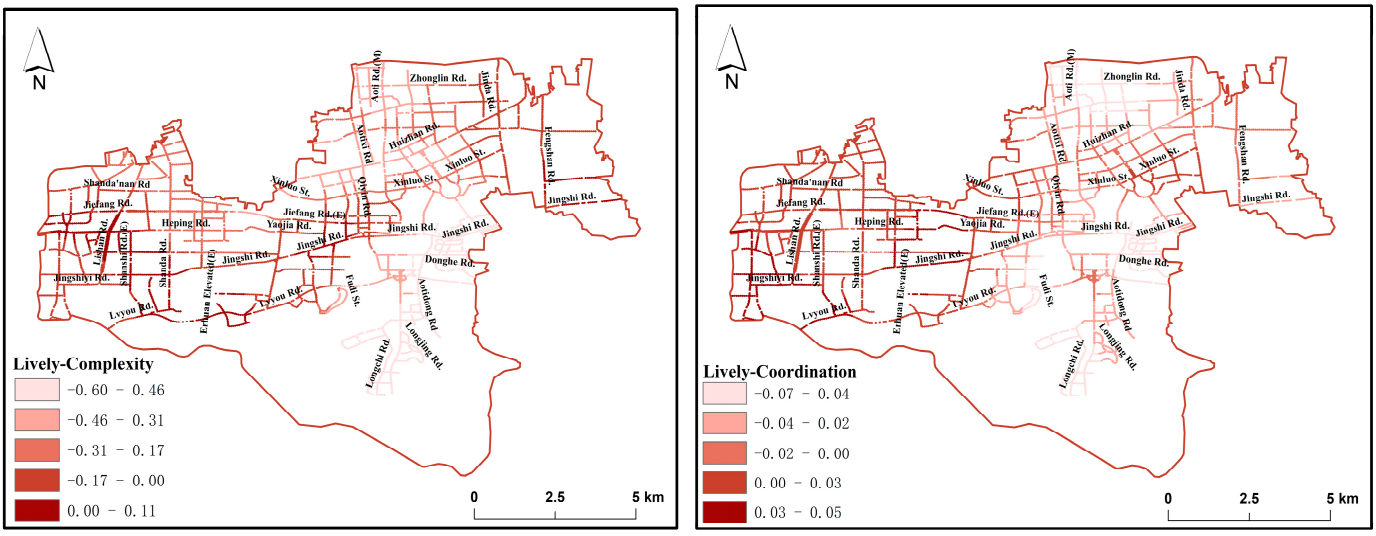
In the majority of locations within Lixia District, an increase in color complexity is observed to positively influence the perception of repressed emotions, particularly in the area surrounding Heping Road. Conversely, in locations such as Shunqiao Road, Tourism Road, and Aoti Middle Road, the impact of color complexity is either diminished or even inverse, suggesting that the color complexity in these areas exerts a less pronounced influence on the perception of repressed emotions, and may even attenuate the sensation of suppression. In the majority of areas within Lixia District, color coordination has a detrimental effect on the perception of repressed emotions, particularly in the central region. In contrast, the effect of color coordination is found to be less pronounced in the northwestern region, which is dominated by Daminghu Road.



(**a**) Depressing–Complexity (**b**) Depressing–Coordination

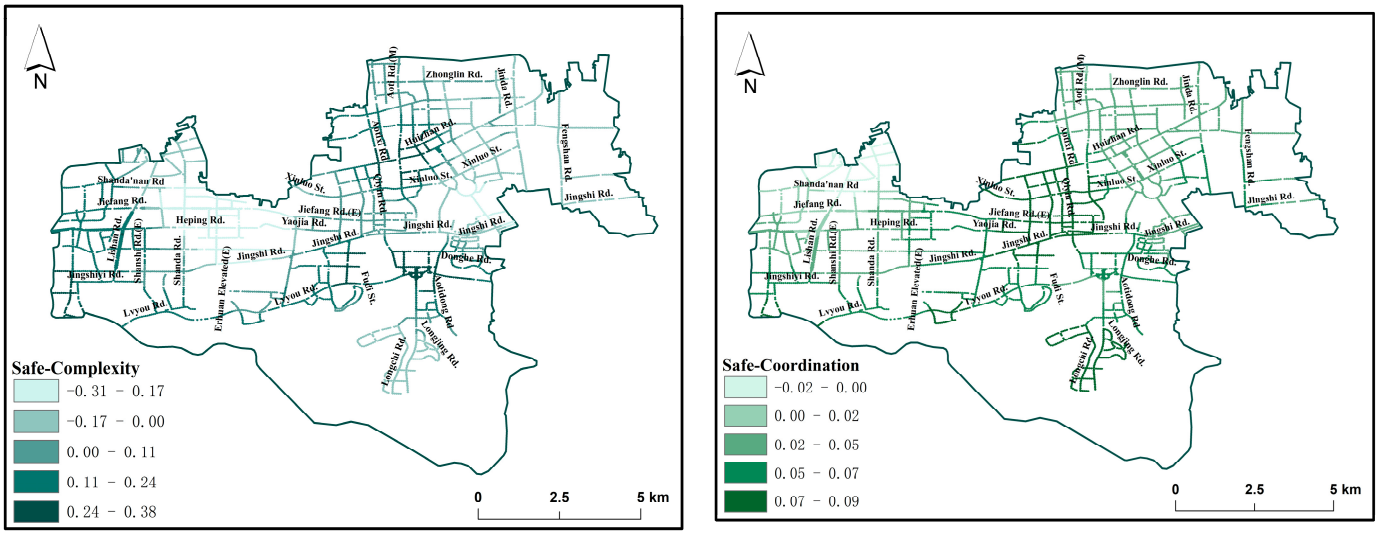
Figure 21. Visualization of GRW analysis of depressing perception–colors.

In the southwestern region, centered on Longding Avenue, and the central region near Aoti Road, an increase in color complexity is found to have a significant negative effect on feelings of liveliness. However, in the southwestern region, especially Lok Yuen Street, an increase in color complexity is found to have a significant positive effect on feelings of liveliness. Conversely, beyond Pulp Spring Road, there are notable regional variations in the impact of color coordination on feelings of liveliness. In the western region, particularly in the vicinity of Shunqiao Road, color coordination exerts a pronounced positive influence on feelings of liveliness. In contrast, in the eastern region, especially in the northern area and in the vicinity of Longteng Road, color coordination evinces a pronounced negative effect on feelings of liveliness.



(**a**) Lively–Complexity (**b**) Lively–Coordination

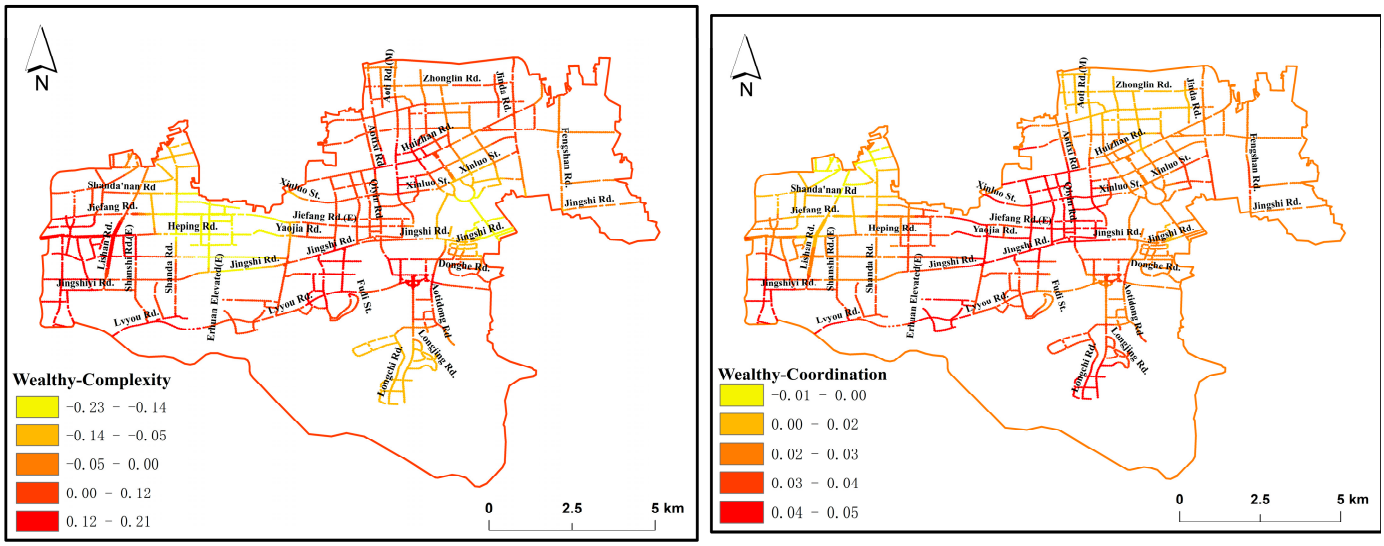
Figure 22. Visualization of GRW analysis of lively perception–colors.



(**a**) Safe–Complexity (**b**) Safe–Coordination

Figure 23. Visualization of GRW analysis of safe perception–colors.

In the western area, which is dominated by Lok Yuen Street, and the area dominated by Long Ao Road, the presence of color complexity has a positive effect on the perception of safety emotions. In the area dominated by Heping Road and the northeastern area dominated by South Industrial Road and Century Avenue, the presence of color complexity is found to have a negative effect on the perception of safety emotions. This negative effect is particularly pronounced in the case of Heping Road. From the perspective of the influence of color coordination on the perception of safety emotions, the vast majority of areas in Lixia District demonstrate a positive influence of color coordination on the perception of safety emotions, particularly in the central area dominated by Jiefang East Road and the southern area dominated by Longchi Road. However, in the northwestern area dominated by East Minghu Road and Linongzhuang Road, color coordination has a negative effect on the perception of safety emotions.



(**a**) Wealthy–Complexity (**b**) Wealthy–Coordination

Figure 24. Visualization of GRW analysis of wealthy perception–colors.

In the area west of Shanda Road and the central area dominated by Ou Ti Zhong Road and Jiefang Dong Road, the complexity of color has a significant positive effect on

perceptions of emotionality, particularly evident on Quan Cheng Road, Lok Yuan Street, and Sun Gao Road. However, in the area dominated by Heping Road and the southern area dominated by Longchi Road, color complexity has a negative effect on the perception of emotionality. With regard to the effect of color harmony, the majority of areas in Lixia District demonstrate a positive effect, particularly the central area dominated by Xinluo Street and West Olympic Sports Road. Conversely, a negative effect is predominantly concentrated around East Minghu Road.

A comparison of the results of the OLS regression and GWR analyses reveals that the former provides an overall average effect of color complexity and coordination on affective perceptions, but fails to reflect regional differences. In contrast, the latter reveals the significant spatial heterogeneity of color features' impacts on affective perceptions.

4. Discussion and Conclusions

The alleviation of negative emotions and the enhancement of positive emotions among residents have long been identified as key strategies for improving the quality of life in urban areas. The ongoing advancement of big data technology has facilitated analyses of the multifaceted factors influencing residents' emotional experiences, taking into account the distribution patterns of diverse environmental elements at the spatial scale. In the context of rapid urbanization, we focus on Lixia District of Jinan City as a case study. This study employs the comprehensive utilization of urban streetscape big data and deep learning methodologies. It focuses on the emotional perceptions of street residents as the core research object, utilizing urban streetscape big data to ascertain the location perceptions of expansive urban areas. Furthermore, it adopts urban color characteristics as the research perspective, elucidating the influence mechanisms of the complexity and coordination of colors in the streetscape on the emotional perceptions of residents. The principal findings of this study are as follows:

(1) The predominant color types in Lixia District of Jinan City are gray and blue-green. Gray is distributed in a more concentrated manner, while blue-green is distributed in a more scattered manner. Additionally, the majority of areas in Lixia District exhibit a high color complexity and low color coordination.

(2) There is a high degree of correlation between emotional perceptions, with a high positive correlation between beauty, security, and richness and a moderate positive correlation between boredom and depression. The prevalence of depression and liveliness is higher in the central area, while the perception of safety and richness is lower in this region. As distance from the center increases, residents' depression levels decline, liveliness levels rise, and emotional distributions exhibit a regular aggregation phenomenon. Positive emotions are consistent in areas of a high spatial value, while negative emotions display the opposite spatial distribution characteristics.

(3) The complexity and coordination of colors are pivotal elements influencing the emotional perceptions of residents. A greater degree of color complexity tends to provide a more diverse visual stimulus, which, in turn, attracts the attention of residents, particularly in commercial areas and entertainment venues. However, an excess of complexity may result in visual fatigue, which may, in turn, elicit feelings of depression and unease. In contrast, a moderate color complexity has been found to stimulate vitality and pleasure, thereby inducing feelings of relaxation and ease. Meanwhile, color coordination has been demonstrated to have a more pronounced positive effect on residents' emotional experience. Highly coordinated color combinations have been shown to contribute to a more harmonious and aesthetically pleasing streetscape, and may enhance residents' sense of security and happiness. In contrast, poorly coordinated colors have been found to increase visual confusion and lead to emotional discomfort and tension.

This study makes a significant contribution to both the theoretical and practical understanding of residents' emotional perceptions. Theoretically, the study validates the argument that "environmental color harmony can enhance the visitor experience", which

is in line with existing research results on residents' emotions. Furthermore, this study enriches the theoretical framework of the impact of color features (combinations) on cities. In terms of methodology, the objective of this study was to propose a quantitative methodology that enables the automated and efficient measurement of human perceptions of large-scale urban environments. This approach involves the combination of street-level imagery, representing specific locations, and deep learning methods to facilitate the comprehension of images at a high level of abstraction. By undertaking a multi-level analysis of color distribution, color complexity, color coordination, and residents' emotional perceptions in Lixia District, Jinan City, it was possible to derive an overall impact of urban color on emotional perception. Furthermore, this study furnishes urban leaders in Lixia District, Jinan City, with residents' visual preferences for street colors, thus providing a crucial reference point for the optimization of streetscape design.

However, there are still some shortcomings in this study, as follows: (1) Only a part of Lixia District in Jinan City was selected as the study area, so the study could be further expanded to the main urban area of Jinan City to conduct a more comprehensive analysis of emotions and colors in the future. (2) Color extraction only used simple color clustering, which was able to identify the main color components in an image, but was unable to distinguish between different semantic objects in an image, which may have led to low accuracy. It is suggested that future research should combine more advanced semantic segmentation models to further improve segmentation accuracy. (3) In this study, only urban spatial and color features were considered, and the effects of variables such as air pollution, weather, noise, and other related variables were not considered, which could be further improved in the future.

In sum, this study offers significant insights into the impact of urban color on residents' emotions and provides practical recommendations for urban planning and design. In the future, urban planners may appropriately adjust the complexity and coordination of street colors based on the functions and characteristics of different areas, thereby enhancing the visual experience of cities and the emotional well-being of their residents. For instance, in commercial districts, color complexity should be moderately increased to stimulate vitality, whereas in residential districts, emphasis should be placed on the beauty of color harmony to create a more peaceful and comfortable living environment. This differentiated strategy of color management will help to enhance the aesthetics and livability of cities and promote sustainable urban development. Further research is needed to explore the influences of the interactions between color and other environmental factors on emotional perceptions, providing a scientific basis for refined urban design.

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

Funding: This research was funded by the National Natural Science Foundation of China, grant number 41801308.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The street view images utilized in this study were sourced from the Baidu Map Street View service through API programing in the public domain: https://api.map.baidu.com/lbsapi/viewstatic.htm.

Acknowledgments: The authors thank the anonymous reviewers for their constructive comments.

Conflicts of Interest: Author Weikang Cui was employed by the company Inspur Smart City Technology Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Chen, C.; Li, H.; Luo, W.; Xie, J.; Yao, J.; Wu, L.; Xia, Y. Predicting the effect of street environment on residents' mood states in large urban areas using machine learning and street view images. *Sci. Total Environ.* **2022**, *816*, 151605. [CrossRef] [PubMed]
2. Beenackers, M.A.; Kruize, H.; Barsties, L.; Acda, A.; Bakker, I.; Droomers, M.; Kamphuis, C.B.M.; Koomen, E.; Nijkamp, J.E.; Vaandrager, L.; et al. Urban densification in the Netherlands and its impact on mental health: An expert-based causal loop diagram. *Health Place* **2024**, *87*, 103218. [CrossRef] [PubMed]
3. 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]
4. 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]
5. Meichen, D. Quantitative contrast of urban agglomeration colors based on image clustering algorithm:Case study of the Xia-Zhang-Quan metropolitan area. *Front. Archit. Res.* **2021**, *010*, 692–700.
6. Rui, Q.; Cheng, H. Quantifying the spatial quality of urban streets with open street view images: A case study of the main urban area of Fuzhou. *Ecol. Indic.* **2023**, *156*, 111204. [CrossRef]
7. Rang, W.; Liang, H.; Wang, Y.; Zhou, X.; Cheng, D. A unified hybrid memory system for scalable deep learning and big data applications. *J. Parallel Distrib. Comput.* **2024**, *186*, 104820. [CrossRef]
8. Fu, L.; Li, J.; Chen, Y. An innovative decision making method for air quality monitoring based on big data-assisted artificial intelligence technique. *J. Innov. Knowl.* **2023**, *8*, 100294. [CrossRef]
9. Lindal, P.J.; Hartig, T. Architectural variation, building height, and the restorative quality of urban residential streetscapes. *J. Environ. Psychol.* **2013**, *33*, 26–36. [CrossRef]
10. Subiza-Pérez, M.; Vozmediano, L.; Juan, C.S. Green and blue settings as providers of mental health ecosystem services: Comparing urban beaches and parks and building a predictive model of psychological restoration—ScienceDirect. *Landsc. Urban Plan.* **2020**, *204*, 103926. [CrossRef]
11. Nisaito, Y. *Deep Learning for Beginners*; People's Posts and Telecommunications Publishing House: Beijing, China, 2018.
12. Shao, Y.; Ye, D.; Ye, Y. A Study on Large-Scale Measurement of Street Interface Penetration Rate Based on Street View Data and Deep Learning—A Case Study of Shanghai. *Urban Plan. Int.* **2023**, *38*, 39–47.
13. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *42*, 386–397. [CrossRef] [PubMed]
14. Zhao, H.; Shi, J.; Qi, X.; Wang, X.; Jia, J. Pyramid Scene Parsing Network. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 6230–6239. [CrossRef]
15. 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]
16. Verma, D.; Jana, A.; Ramamritham, K. Predicting human perception of the urban environment in a spatiotemporal urban setting using locally acquired street view images and audio clips. *Build. Environ.* **2020**, *186*, 107340. [CrossRef]
17. Fan, Z.; Bolei, Z.; Liu, L.; Yu, L.; Fung, H.H.; Hui, L.; Carlo, R. Measuring human perceptions of a large-scale urban region using machine learning. *Landsc. Urban Plan.* **2018**, *180*, 148–160.
18. Yang, R.; Deng, X.; Shi, H.; Wang, Z.; He, H.; Xu, J.; Xiao, Y. A novel approach for assessing color harmony of historical buildings via street view image. *Front. Archit. Res.* **2024**, *13*, 764–775. [CrossRef]
19. Chen, H. Deep Learning-Based Extraction of Quantity and Spatial Relationships of Streetscape Features. Master's Thesis, Hebei University of Engineering, Handan, China, 2022.
20. Gong, F.Y.; Zeng, Z.C.; Zhang, F.; Li, X.; Ng, E.; Norford, L.K. Mapping sky, tree, and building view factors of street canyons in a high-density urban environment. *Build. Environ.* **2018**, *134*, 155–167. [CrossRef]
21. Ning, Z.; Yun, C.; Yuebin, W.; Bo, H.; Takis, M.P. Land-Use Mapping for High-Spatial Resolution Remote Sensing Image Via Deep Learning: A Review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 372–391.
22. Liang, Q. Research on Semantic Segmentation Algorithm for Urban Street View Images Based on Convolutional Neural Networks. Master's Thesis, Changchun University of Technology, Changchun, China, 2024.
23. Zhao, Y. Research on the Semantic Segmentation Method of Urban Street Scene Based on DeepLabv3+. Master's Thesis, Nanjing University of Posts and Telecommunications, Nanjing, China, 2023.
24. Yang, H. Research on automatic classification method of building style based on real image. Master's Thesis, Beijing University of Civil Engineering and Architecture, Beijing, China, 2024.
25. Gu, H.; Jiang, H. Quantitative Control Method of Urban Colour Primary Tone with Digital Technology Support--Taking Ankang Urban Colour Planning and Design as an Example. *Planners* **2013**, *29*, 42–46.
26. Ziyin, Q.I.; Junyi, L.I.; Zhe, H.E.; Xiping, Y. The Influence of Urban Streetscape Color on Tourists' Emotional Perception Based on Streetscape Images. *J. Geo-Inf. Sci.* **2024**, *26*, 514–529.
27. Liu, C.; Li, J.; Li, X. An example of the research design of the image region of interest extraction method based on HSV colour space. *Softw. Eng.* **2024**, *27*, 1–5.
28. Liu, J.; Li, D. Adaptive K_means image segmentation method based on Lab colour space. *Mach. Des. Manuf. Eng.* **2018**, *47*, 23–27.
29. Jinan Municipal People's Government. Available online: http://www.jinan.gov.cn/col/col129/index.html (accessed on 18 April 2024).

30. Jinan Municipal Bureau of Statistics. Available online: http://jntj.jinan.gov.cn/art/2024/3/28/art_18254_4750937.html (accessed on 28 March 2024).
31. Ye, Y.; Zhang, L.; Yan, W.; Zeng, W. A Framework for Measuring the Quality of Street Greening from a Humanistic Perspective—A Large-Scale Analysis Based on Baidu Street View Data and Machine Learning. *Landsc. Archit.* **2018**, *25*, 24–29.
32. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
33. Dubey, A.; Naik, N.; Parikh, D.; Raskar, R.; Hidalgo, C.A. Deep Learning the City: Quantifying Urban Perception at a Global Scale. In *Computer Vision—ECCV 2016*; Springer: Cham, Switzerland, 2016; pp. 196–212.
34. 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] [PubMed]
35. Garcia-Garcia, A.; Orts-Escolano, S.; Oprea, S.; Villena-Martinez, V.; Garcia-Rodriguez, J. A Review on Deep Learning Techniques Applied to Semantic Segmentation. *arXiv* **2017**. [CrossRef]
36. Wang, X.; Guo, Y.; Wang, S.; Cheng, G.; Wang, X.; He, L. Rapid detection of incomplete coal and gangue based on improved PSPNet. *Measurement* **2022**, *201*, 111646. [CrossRef]
37. Liu, H.; Li, Y.; Li, Y.; Liu, X. Research on building change detection method based on improved PSPNet model for high resolution remote sensing images. *Surv. World* **2024**, *2*, 1–4.
38. Pan, Y.; Qiu, L.; Wang, Z.; Zhu, J.; Cheng, M. Unravelling the association between polycentric urban development and landscape sustainability in urbanizing island cities. *Ecol. Indic.* **2022**, *143*, 109348. [CrossRef]
39. Zhao, Z.; Luo, H. GIS-based Spatial Evaluation Research on Cultural Facilities--Taking Zhifu District of Yantai City as an Example. *Urban. Archit.* **2023**, *20*, 79–83.
40. Li, M.; Zhong, F.; Peng, Z. Research on Rural Residential Base Remediation Strategies in Karst Areas Based on Cluster Analysis--Taking Xifeng County of Guiyang City, Guizhou Province as an Example. In *Spatial Governance for High-Quality Development—Proceedings of the Annual Conference on Urban Planning in China 2021*; China Architecture Industry Press: Beijing, China, 2021; pp. 779–791.
41. Wang, Z.; Wang, X.; Xie, X.; Xiao, M.; Wu, Y.; Li, X. Influencing factors of green space landscape pattern evolution and its spatial differences in Fujian Province based on the GWR model. *J. Northwest For. Univ.* **2022**, *005*, 242–250.

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.