ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 1 - 2026

Urban Architectural Landscape Shaping and Sustainable Development Strategies

Zhankuo Yao[1,*]

[1] School of Civil Engineering and Architecture, Linyi University, Linyi, Shandong, 276000, China

SUMMARY: *This study proposes a set of multidimensional perception shaping framework for urban architectural landscape integrating computer vision, spatial analysis and color theory, aiming to provide decision support for landscape protection and sustainable development. Firstly, on the basis of Mask-RCNN, k-means algorithm is used to cluster the anchor frames to improve the accuracy of detection. Then the FCN semantic segmentation branch is improved to improve the accuracy of urban building mask and realize the accurate recognition of different buildings. Then based on spatial analysis techniques such as kernel density and spatial syntax, the spatial characteristics of urban architectural style are revealed. Based on the color partitioning method, the color level of Beijing urban buildings is divided into regions, and the color level data of Beijing urban buildings are mined and analyzed. The improved Mask-RCNN building recognition model achieves an average accuracy mean of 91.36% compared with the current mainstream deep learning network. The spatial features show that the Second Ring and Third Ring are the core areas of urban architectural landscape and the highest road network integration area, respectively, which better constructs the skeleton of landscape perception. The color analysis shows that the dominant colors of roofs and walls are concentrated in the yellow-red system, which provides a sustainable method for the conservation and renewal of urban buildings and historical coordination.*

KEYWORDS: *Mask-RCNN; k-means; kernel density analysis; average nearest neighbor analysis; urban building landscape*

1 Introduction

Economic globalization has extended the flow of energy, commodities, manpower and information across the globe, with international trade and investment constituting the main drivers of globalization. Much of the economic dynamism comes from the development of cities, which has led to an accelerated pace of urban economic development. Globalization has aroused a strong desire among cities to attract investment to enhance city branding, and has caused all parties in cities to respond to this trend [1]. Contemporary China is in the early stage of the “advent of urban society”, and it is estimated that in the next 20 years or so, there will be an influx of about 600 million people into the cities [2]. Such a big change in society will result in a more fierce competition between cities. Under the pressure of such competition, the government must find ways to build the competitiveness of the city and improve the quality of the city [3-5]. Residents' pursuit of the city often only stays in the modern skyscrapers and does not pay attention to the creation of urban characteristics, which leads to the formation of a

*yaozhankuo@126.com
https://doi.org/10.65102/is2026154
Published: 30 April, 2026

thousand cities in the image of the major cities [6]. The characteristic cultural practices of each region are gradually fading, how to give the traditional cultural characteristics of architecture and the spirit of the times at the same time is a difficult problem in the minds of contemporary architects.

Urban architectural style is a visual carrier to show the urban style, which is subject to both natural and social factors [7]. It needs to be perfectly integrated with the status quo of the natural environment, and also needs to assume the role of displaying the humanistic characteristics of the city. The role of architecture is different because of the different spatial levels in the city, such as architecture in the overall macro layout of the city mainly constitutes the role of texture, in the whole urban form plays the role of the wheel corridor, while the overall color of the city also has a decisive impact [8-10]; in the neighborhoods, such as the meso level of the landscape or the role of the landmarks; and at the micro level of the spatial composition, the role of places and the spirit of the show is also not to be ignored. In the micro level, its role on the spatial composition and the spirit of the place can not be ignored [11]. In conclusion, the influence of architecture on urban landscape comes from various aspects and levels.

Scholars define urban landscape as the city's style and appearance, which is a collection of natural factors, social factors, its historical lineage and local customs [12]. The “style” in urban landscape refers to the social and humanistic intention of the city, and the “appearance” refers to the natural conditions, spatial patterns and other physical substances in the city [13]. Literature [14] recognizes urban landscape as a dynamic landscape of evolving social, economic and ecological change and diversity, which operates as an entity of ecological, infrastructural systems and planning needs, and at the same time is the basis for urban development. Literature [15] proposed a method for computational analysis of urban landscape, called "Weighted Visual Exposure Map (WVEM)", which aims to provide support for the review of the landscape in the city. For each point of the city, the value of the landscape will be calculated, and the results can be used as the visibility of the most important selected landmarks in the city image at that position. Literature [16] points out that UNESCO's adoption of Historic Urban Landscapes (HUL) has led to international attention to the importance of urban landscape for urban development, calling for a landscape approach to ensure that cultural heritage policy and management issues are integrated into the broader goals of sustainable urban development. Literature [17] reveals the dichotomy between modernity and tradition in Suzhou's urban landscape from a morphological point of view, which is the basis for rapid urbanization and economic development, pointing out that urban landscape is an intangible value that helps local residents to form a closer sense of local belonging and identity. Literature [18] analyzed the case of urban landscape management in Nursultan, Kazakhstan, and analyzed the condition of green space using GIS technology, pointing out that the sustainable development strategy of urban ecosystems is the main direction of urban landscape.

Architectural style is the core content of urban style, and is the most intuitive part of urban characterization. Architectural style is the sum of physical attributes of buildings (color, height, form, etc.) and their social and humanistic orientation, which is determined by specific historical and cultural background, people's aesthetics, living habits, as well as specific climatic conditions, geographic environment, and other factors, reflecting the cultural connotation and urban characteristics of a city [19, 20]. Literature [21] believes that the urban architectural style is the result of a combination of factors such as nature, humanities and building technology, and analyzes the architectural style of Danzhou City, Hainan Province, which demonstrates the significant characteristics of Danzhou City in terms of urban space and architectural language. Literature [22] believes that the transformation of urban architectural style often accompanies the change of the natural environment, and analyzes in detail a variety of architectural styles such as the national romanticism architecture of northern China, the Moscow model

architecture, and the exotic Qingdao seaside city architecture. Literature [23] analyzed the characteristics of the geographical differentiation of traditional architectural styles by taking the representative characteristic architectural styles in China as a case study, and then explored the influencing mechanism of the geographical differentiation, and found that the natural environmental factors are the cornerstone of the geographical differentiation of traditional architectural styles; the humanistic factors, such as the national culture and the patriarchal ethic, are the internal driving force of the geographical differentiation; and the humanistic factors, such as the population relocation, the war defense, and the trade and economic factors, are the external pushing force of the geographical differentiation. The Literature [24] takes Wuhan as an example for a case study, evaluates the effectiveness of existing architectural features and styles management policies, and develops a comprehensive architectural features and styles management and implementation system, aiming to contribute to the improvement of urban identity management and the protection of architectural landscape diversity, so as to create a more cohesive and culturally resonant urban environment. Literature [25] analyzes the architectural style of Harbin, China, and finds that Harbin, known as the "Oriental Moscow" and "Oriental Little Paris", is characterized by its beautiful buildings in the "Art Nouveau" style, which has the characteristics of subjectivity, popularity, regional differences and regional adaptability. Literature [26] has conducted an in-depth study on the protection and development of urban historical features, emphasizing the importance in urban planning and management, and the study concludes that the protection of urban historical features should be based on a comprehensive understanding of the various forms of internal organization and external relations.

The control of landscape often lags behind the construction practice, making it difficult to realize the dynamic balance between protection and development. In order to cope with the above problems, this paper constructs a set of multi-dimensional and multi-technology fusion research paths, based on the established Beijing streetscape dataset, and carries out example analysis. An improved Mask-RCNN model based on k-means algorithm and semantic segmentation branch optimization is introduced to achieve automated high-precision recognition of urban buildings. Kernel density analysis, average nearest neighbor index and spatial syntax are applied to reveal the spatial distribution pattern of urban architectural landscape and the structural characteristics of street network. At the visual perception level, the color quantization theory is used to analyze the main color composition and distribution characteristics of roofs and walls.

2 Deep learning-based urban building recognition

2.1 Mask-RCNN

Mask-RCNN is an instance segmentation method based on Faster-RCNN by adding semantic segmentation branch FCN.In addition to adding semantic segmentation branch, the method also improves Rol Pooling in Faster-RCNN and proposes Rol Align with better effect.The overall structure of the Mask-RCNN algorithm mainly consists of backbone The overall structure of Mask-RCNN algorithm mainly consists of backbone network, region suggestion network, Rol Align, target detection branch, semantic segmentation branch and so on. The backbone network uses ResNet101 and FPN to extract features, after the backbone network extracts the features, the features are sent to the region suggestion network RPN for processing, the RPN can generate the vertex coordinates of the outer rectangle where the target to be recognized may exist, and then the recommendation region is mapped with the features to get the features of the recommendation region, due to the inconsistency in the size of the recommendation region, the

features are processed to the same size by using Rol Align. Due to the inconsistency in the size of the recommended region, Rol Align is used to process the size of the feature region to the same size, and the features processed by Rol Align are further classified by all-connected-layer processing and edge regression to obtain the external rectangle and categories of the target to be recognized, and finally, FCN is used to further predict the categories of each pixel in the external rectangle of the target to be recognized to generate the target's mask.

(1) Feature extraction network

Feature extraction network, also generally called backbone network, is mainly used for feature information extraction, as the basis of the downstream task, its extraction effect directly determines the performance of the model, in the Mask-RCNN, the use of residual network and FPN combination can be further fusion of features extracted by residual network at different scales.

(2) Regional Proposal Network

RPN slides over the features output from the last convolutional layer of the backbone network through a 3×3 filter, and predicts a number of anchor frames at each position when sliding, considering the diversity of target shapes, RPN presets 9 types of anchor frames with different sizes, and then calculates the coordinate information of the frames and the probability of whether there is a target in the frames through the two fully connected layers, respectively. Finally, the calculated coordinate and category data are sent to RoIAlign for further processing.

(3) RoI Align

RoI pooling is used in Faster-RCNN to transform the feature dimensions in the recommended region into a fixed H×W size feature map. However, RoI Pooling produces bias in region matching due to the presence of two rounding operations of feature dimensions. Due to the loss of feature information caused by the quantization operation, the quantization operation is canceled in RoI Align, and the candidate frame is divided into n regions, and then these n regions are further divided into 4 regions, and the pixel values in the centers of the 4 regions are calculated by the bilinear interpolation method, and then the interpolation result is further processed according to the Maximum Pooling method.

(4) Loss function

Mask-RCNN training consists of two stages, the first stage trains the regional recommendation network RPN, the second stage will be generated by RPN to continue the training of the recommended region through the RoIAlign processing, and further classify the target, border regression and semantic segmentation, the whole training process contains multiple tasks, so the use of multi-task loss, the loss function can be divided into a border regression loss, category loss, semantic segmentation loss three categories, the formula is as follows:

$$L = L_{bbox} + L_{cls} + L_{mask} \tag{1}$$

where the border regression loss L_{bbox} contains the border regression loss in RPN and the border regression loss in the second stage of training, the border regression uses Smooth L1 Loss, the classification loss L_{cls} also contains the border classification loss in RPN and the second stage of the border classification loss, respectively, and the classification loss uses the multiclassification cross entropy loss, and L_{mask} is the semantic segmentation loss, and the loss function uses the binary cross entropy loss.

2.2 Algorithm Optimization

Mask-RCNN as a kind of supervised learning requires a large number of labeled datasets, but

the cost of obtaining and producing labeled datasets is high, and the training data is not enough, the model is easy to be overfitted, in order to solve this problem this paper uses migration learning to train the model, and further trains its own dataset on the basis of the pre-trained model. In addition, this paper optimizes the preset anchor frames and semantic segmentation branch for the region suggestion network RPN in Mask-RCNN, firstly, k-means clustering method is used to obtain anchor frames that are more suitable for this paper's monitoring dataset instead of the original preset anchor frames, and secondly, parallel small convolution kernel convolutional layers and symmetric hopping connection structure are added in the semantic segmentation branch.

2.2.1 Anchor box clustering

The 9 anchor boxes predicted by RPN at the location of the filter are of fixed size. Considering that the fixed 9 sizes of anchor box have a large deviation in overlap against the target edges in different datasets, we choose to use k-means method to cluster the target edges in the dataset to get the appropriate target edges. The clustering steps in this paper are as follows:

a) Select the number of categories for clustering k and initialize k cluster centers (anchor boxes);

b) Calculate the distance of all samples to k anchor boxes, determine which anchor box the sample is closest to by comparing the size of the distance, and categorize the sample in the category with the closest distance, the distance measure in this paper adopts IoU, and the distance calculation formula is as follows:

$$d\left(centroid, sample\right) = 1 - IoU\left(centriod, sample\right) \tag{2}$$

where centroid denotes the clustering center, sample denotes the sample, and IoU denotes the ratio of the intersection and concatenation area of centroid and target. In the formula, the larger the value of IoU, the higher the overlap of the two frames and the smaller the d; conversely, the larger the d and the further the distance. In order to facilitate the calculation of IoU, the coordinates of all sample center points are set to $(0,0)$. The formula for calculating IoU between cluster centers and samples is as follows:

$$IoU\left(centroid, sample\right) = \frac{Area(B)}{Area(A)} \tag{3}$$

c) Update the clustering center, calculate the median of the samples in each category, and use it as the new clustering center.

d) Cycle steps a) and b) to obtain the final clustering results, i.e., k different sizes of target borders.

The validation metrics of the clustering results use Avg_IoU (Average_IoU), which calculates the IoU of individual samples and k clustering centers, then each sample corresponds to k IoU values, and the maximum value is taken from k IoU values, i.e., to find the clustering center that has the highest overlap with the sample, denoted as $IoU_{\max}$, and calculate the The $IoU_{\max}$ corresponding to all samples is averaged to be Avg_IoU, which is calculated as follows.

$$IoU_{\max} = \max_{i \le k} \left\{ IoU\left(centroid, sample\right) \right\} \tag{4}$$

$$Avg\_IoU = \frac{\sum_{j=1}^{n} IoU_{\max}}{n} \tag{5}$$

2.2.2 Semantic Segmentation Branch Optimization

Further feature information is extracted at the semantic segmentation branch using jump joins and parallel convolution. After RoIAlign normalization of the features, a parallel convolution structure is added, which contains four convolution blocks with different number of layers and smaller convolution kernels, which enables multi-scale fusion with a small increase in parameters.

The features are all of the same dimension after RoIAlign processing, denoted as A, and four different convolution blocks are used in the parallel convolution structure to process the features A. After obtaining 4 different scales of features, use Concatenate to splice them to get B, use 1×1 convolution to downsize the feature B (reduce the number of channels) to get the feature C; then use the jump-joining method to further process the feature C, up-sample the feature and splice it with C to get D, and finally, use bilinear interpolation to up-sample the feature C. sampling features D to get the final mask prediction result using bilinear interpolation.

2.3 Identification test and result analysis

2.3.1 Establishment of data sets

With the continuous innovation of hardware technology such as in-vehicle cameras, the use of algorithmic models to perceive and parse street scenes has become a cutting-edge hotspot in the current research field. In the process of in-depth exploration of intelligent recognition of large-scale traditional style buildings, a set of street scene dataset is firstly constructed to support the evaluation of the architectural style recognition model, and the dataset should comprehensively cover the multiple forms and features of the style buildings to ensure that it has sufficient diversity and generalization ability.

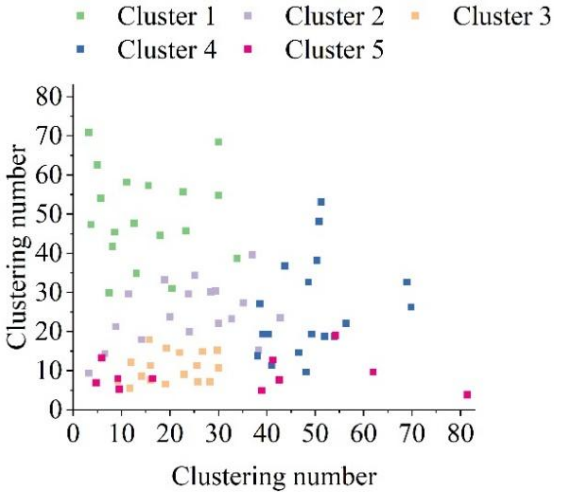
Subsequently, the dataset is meticulously divided into a training set, a validation set and a test set. The role of the training set is to provide sufficient learning samples for the deep learning model, and through a large number of learning iterations, the model can gradually learn the intrinsic laws and feature expressions of traditional style buildings and extract key features from the images. The validation set is mainly used to adjust the hyper-parameters of the model, and by comparing and analyzing the performance of the model in the validation set under different hyper-parameter combinations, the hyper-parameter configurations can be precisely found to optimize the performance of the model. The test set is the key link to evaluate the generalization ability of the model, which is used to comprehensively test the recognition effect of the model in the face of unknown new data, and then evaluate the practicality and reliability of the model. In terms of the data division ratio, the demand of deep learning models for the amount of training data is taken into account, as well as the important role of validation set and test set in model optimization and evaluation. Therefore, this paper adopts the ratio of 80%, 10%, and 10% to divide the training set, validation set, and test set, so that the model can be fully trained and at the same time fully verify its generalization ability.

From the crawled down Beijing Baidu street view images, 660 images are selected as the training set, 83 images are used in the validation set and test set respectively, and the street view dataset is annotated according to the characteristics of the Beijing city building classification organized in this paper, which are labeled as two types of data, namely, traditional buildings

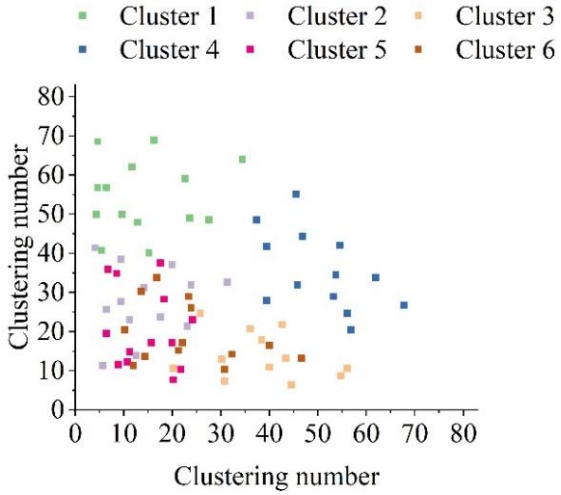
and non-traditional buildings. The annotation tool used in this paper is Labelme, which is a powerful interactive image annotation tool that supports a variety of image annotation tasks, including polygon annotation, rectangular box annotation, point annotation, etc., and is widely used in computer vision and machine learning research. In this paper, the labeled dataset is saved as a file in JSON format, which is convenient for subsequent processing and analysis.

2.3.2 Proportional sizing of a priori frames

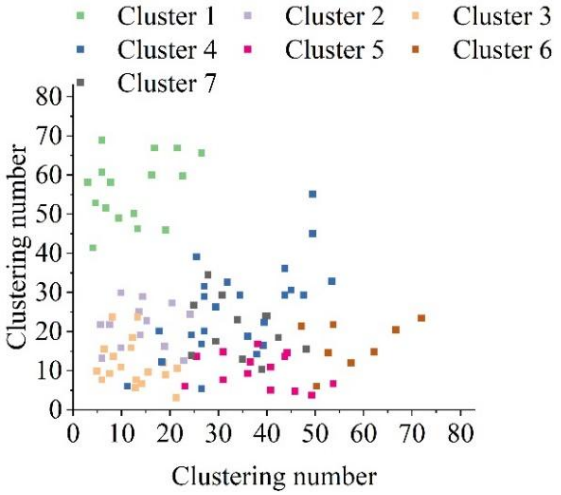
In this paper, the a priori frame proportion size is determined by K-means clustering algorithm,K-means visualization clustering diagram is shown in Fig. 1, Figs. (a)~(d) are the visualization clustering results with the number of clustering clusters of 5, 6, 7 and 8, respectively. As can be seen from Figure 1:When the number of clusters increased to 7, the clustering results no longer have a large change, indicating that the a priori frame proportion is uniformly distributed on the actual data, in line with the size requirements of the actual data, at this time the a priori frame proportional to the size of the (0.06,1,17.55).



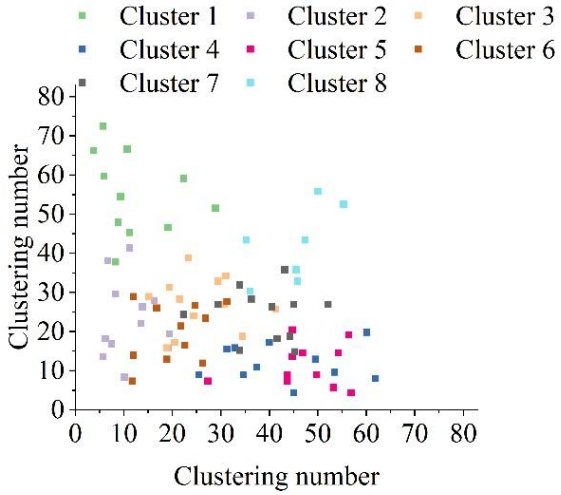
(a)The number of clustering is 5



(b)The number of clustering is 6



(c)The number of clustering is 7



(d)The number of clustering is 8

Figure 1: K-means Visual Clustering Diagram

2.3.3 Model training

The Mask-RCNN in this paper is constructed based on the TensorFlow framework. In the training process, the initial learning rate is 1×10^{-3}, the batch size is 2, the training period is 40, each training cycle is 4000 steps, the momentum factor is 0.8, and the gradient descent is used as the optimization algorithm, with ReLU as the activation function and Adam as the optimizer. With the increasing number of iterations, the learning rate gradually decreases, so that the network continuously converges to the optimal solution. In order to obtain a better training effect and avoid the randomness of the results, the network is trained repeatedly by cycle training, and the optimal weights are selected and applied to the network.

2.3.4 Evaluation indicators

Generally speaking, when the intersection and integration ratio (IOU) threshold is higher, the regression ability of the model is stronger, and the detection results of the model are more consistent with the actual results. When the intersection ratio between the detection frame and the real frame is greater than 0.5, it is considered that the target is predicted successfully. In this paper, the IOU threshold is set to 0.75, and if the IOU is greater than 0.75, the recognition result is correctly detected by the classifier, i.e., True Positive (TP), and vice versa for False Negative (FN); if the classifier detects the error as correct, then it is regarded as False Positive (FP). The main indexes for evaluating the segmentation results of instances are: precision, recall rate, average precision and average precision mean. Among them, precision is the ratio of the number of correctly detected to the number of detected; recall rate is the ratio of the number of correctly detected to the actual number; average precision is the average value of the precision of different individuals in each category; mean average precision is the average of the average precision of the total categories in the validation set, and it is a comprehensive evaluation index of the performance of the model, and the larger the value is, the better the performance of the model is.

2.3.5 Analysis of identification results

In order to verify the effectiveness of the improved Mask-RCNN proposed in this paper, the impact of K-means clustering algorithm and the utilization of lightweight networks MobileNetV2 and PANet network as the backbone network of Mask-RCNN are discussed respectively. Under the same conditions, the recognition results of five networks, Mask-RCNN, Mask-RCNN+K-means, Improved Mask-RCNN, Faster-RCNN, and YOLOv5, are compared for the dataset produced in this paper.

The k-means clustering algorithm is used to adjust the proportion of a priori frames in the original Mask-RCNN, and the networks are trained under the same conditions and compared with the training results of the original Mask-RCNN, and the results are shown in Table 1.

The mean precision, mean recall and mean accuracy of the original Mask-RCNN for the recognition of five types of buildings, namely, royal palace, palace, altar temple, office building and residential building, are 93.82%, 93.40% and 88.06%, respectively, and the mean precision, mean recall and mean accuracy of the Mask-RCNN ten-k-means for the recognition of five types of buildings are were 94.86%, 93.62% and 89.04%. Compared with the original Mask-RCNN, the mean values of accuracy and recall of Mask-RCNN ten k-means are improved by 1.04% and 0.22%, respectively, and the mean value of average accuracy is improved by 0.98%. It indicates that the k-means clustering algorithm can effectively select the proportion of a priori frames that match the actual buildings, reduce the redundancy of rectangular prediction frames, and improve the localization ability of Mask-RCNN for multi-class building recognition.

Table 1: Influence of K-means Clustering Algorithm on Recognition Results

Building type	Model	Evaluation index		
		Precision / %	Recall / %	Average precision / %
Wangfu	Mask-RCNN	91.9	88.4	81.2
	Mask-RCNN+K-means	92.9	90.0	83.5
Palace	Mask-RCNN	96.1	97.1	93.1
	Mask-RCNN+K-means	96.6	97.3	94.0
Temple of altar	Mask-RCNN	94.5	97.7	92.6
	Mask-RCNN+K-means	95.4	97.1	92.8
Office building	Mask-RCNN	94.1	92.7	87.5
	Mask-RCNN+K-means	95.5	93.8	89.9
Residential building	Mask-RCNN	92.5	91.1	85.9
	Mask-RCNN+K-means	93.9	89.9	85.0

Using the same hyperparameters and dataset, we compare the recognition results of the current mainstream deep learning networks Faster-RCNN, YOLOv5, Mask-RCNN, Mask-RCNN+K-means and Mask-RCNN optimized and improved Mask-RCNN based on anchor frame clustering and semantic segmentation branch optimization for the architectural style recognition dataset constructed in this paper, and the results of recognition are shown in Table 2.YOLOv5 takes the least time to reason about 1 image,only 0.275s,but has the lowest average accuracy. In the case of small difference in detection time, Mask-RCNN changes the a priori frame size by K-means clustering algorithm, and its average accuracy is higher than that of Faster-RCNN and the original Mask-RCNN.In contrast, the improved Mask-RCNN has the highest accuracy in detecting the smaller targets, and has shorter inference time, and its average accuracy is higher than the other networks, with the average accuracy being higher than the other networks, with the average accuracy being higher than the other networks. The average accuracy mean value is 91.36%. This shows that the improved Mask-RCNN has higher recognition accuracy and less inference time for architectural style recognition.

Table 2: Comparison of identification results of different networks

Model	Average precision / %					The mean of the average precision / %	Reason time/s
	Wangfu	Palace	Temple of altar	Office building	Residential building		
Faster-RCNN	76.0	93.2	88.0	89.7	84.6	86.30	0.751
YOLOv5	71.6	82.8	85.4	85.4	74.0	79.84	0.275
Mask-RCNN	81.3	93.1	92.8	87.8	86.1	88.22	0.872
Mask-RCNN+K-means	84.0	94.3	92.7	89.4	84.6	89.00	0.870
This model	87.4	96.1	94.5	90.4	88.4	91.36	0.503

3 Physical perception of urban architectural landscape based on spatial analysis technology

3.1 Spatial distribution techniques

3.1.1 Kernel density analysis

Kernel density analysis is a widely used non-parametric estimation method in spatial analysis, used to calculate the density of the elements in its surrounding neighborhood, the method to the location of a particular element point as the center of the distribution of its attributes in the radius of the h circle, in the center of the position of the density of the largest and with the distance and the decay, to the farthest limit of the distance is the value of the weights is zero. The integral of all the densities in the whole region is summed up to the value of the attribute at the center, and the value of the attribute at the independently existing point is 1. The following is the expression of the function side of the kernel density analysis:

$$f_h(x) = \frac{1}{nh}\sum_{i=1}^{n} k\left(\frac{x - x_i}{h}\right) \tag{6}$$

where: $x_1, x_2, \cdots, x_i$ denote independently existing point elements, $k(\cdot)$ is the kernel function, and $x - x_i$ is the distance from the estimated point x to the sample point x_i. In this paper, this method is mainly used to analyze the distribution status and density of urban buildings, and then determine the aggregation of buildings in the city.

3.1.2 Average Nearest Neighbor Analysis

Mean Nearest Neighbor Analysis was originally used as a concept in plant ecology to compare the similarities and differences between observed areas and random distributions, but the method has since been refined to apply to any form of spatial distribution pattern of points, with the following functional expression:

$$Rn = \frac{\sum \frac{d}{n}}{c\left[\sqrt{\frac{\alpha}{n}}\right]} \tag{7}$$

where Rn represents the nearest neighbor ratio, n is the number of points in the study area, d is the distance from each point to the nearest neighbor, α is the area of the study area, and C is the coefficient of fixation. When $Rn \geq 1$ indicates that the spatial points are uniformly distributed without obvious aggregation characteristics; $1 > Rn > 0$ indicates that the points in the space are distributed in an aggregated manner, and the smaller the number, the higher the degree of aggregation; and when $Rn = 0$, the spatial points show absolute aggregation. It is worth noting that the shape of the region has an important effect on the nearest neighbor analysis, the closer to the square the more accurate the analysis results.

In this paper, we use the average nearest neighbor analysis to determine whether the urban buildings in the area are clustered or discrete and the degree of clustering.

3.2 Characteristics of POI spatial distribution

3.2.1 Spatial distribution statistics

Table 3 shows the statistics of the spatial distribution number of POIs of urban architectural styles, and the POIs of scenic spots of official and folk styles are concentrated within the Second Ring Road, and the north side of the Second Ring Road is more densely distributed than the south side. POIs of official style Fengge in the northwest corner of the Fifth Ring Road are densely distributed. The number of POIs for official-style scenic spots is much larger than the number of POIs for folk-style scenic spots within the Second Ring Road, the Second to Third Ring Roads, and the Fourth to Fifth Ring Roads. On the contrary, the number of POIs of scenic spots in folk style is larger than the number of POIs in official style in the fourth ring road. From the spatial distribution density of POIs, it can be seen that the average POI density value is 1.96 POIs/km², i.e., there are about 2 POIs in one square kilometer, and the highest POI density value is 12.77 POIs/km² in the Second Ring Road, which is much higher than the average POI density value of 1.96 POIs/km², and the POI densities in the Second to the Third Ring Roads, the Third to the Fourth Ring Roads, and the Fourth Ring Roads are all lower than the average POI density.

Table 3: Statistics on the distribution of POI space in scenic and scenic buildings

Loop circuit	Official style	Folk style	Population	Area(km²)	Percentage (%)	POI density (Per square kilometer)
Two rings	635	172	807	62.85	9.40	12.77
Two rings to three rings	80	4	84	96.33	14.41	0.85
Three rings to four rings	56	85	141	143.85	21.52	0.96
Four rings to five rings	285	18	303	365.54	54.67	0.81
Total	1056	279	1335	668.57	100	1.96
Percentage (%)	79.10	20.90	100	100		

3.2.2 Spatial aggregation characteristics

In this paper, we use the average nearest neighbor tool in ArcGIS 10.5 software to analyze the overall urban architectural style POI. and the spatial aggregation characteristics of its two subcategories, official style POI and folk style POI. The results contain five values, average closest neighbor distance, expected average closest neighbor distance, closest neighbor index (R), Z score and P value, the smaller the R value the higher the degree of aggregation. Table 4 shows the average closest neighbor distance, expected average closest neighbor distance, closest neighbor index, Z score, P value and spatial distribution type result values for the overall urban architectural style POI and its two subcategories.

The overall spatial aggregation of traditional style scenic spots within the fifth ring road of Beijing is characterized by the fact that the main body is aggregated within the second ring road, and some of them are scattered in the surrounding urban areas. The overall average closest neighbor distance of traditional style scenic POIs is 118.02m, while the expected average closest neighbor distance under the ideal distribution mode is 352.24m, the closest neighbor index R is 0.33, which is less than 1, the Z score is -45.56, and the spatial distribution type is significant aggregation. The overall spatial distribution of scenic POIs in Beijing traditional style is more regional, which is related to the historical development of Beijing city and the type of protection units where the points of interest are located.

The closest neighbor indices R of the scenic attractions POIs in official style and folk style within the fifth ring road of Beijing are 0.36 and 0.39, respectively, both less than 1, and the

absolute values of the Z scores are greater than 2.58, with a p-value of less than 0.01, which passes the 99% significance test, suggesting that both the POIs of the official style and the POIs of the folk style show the spatial distribution characteristics of significant aggregation. Among them, the clustering characteristics of official style scenic POIs and the overall clustering characteristics are the closest, indicating that the official style scenic POIs dominate the overall traditional style scenic POIs, and their spatial share reflects stronger. The closest neighbor index R-value of folk style scenic POIs is much lower than that of traditional style scenic POIs in general and official style POIs, which is related to the fact that Beijing is an ancient historical capital, as well as the historical development process of the city. As the imperial city of the last dynasty of feudal society, Beijing's official style building complexes, which embody the official regime, have a much higher number of POIs representing traditional style scenic spots than folk style POIs due to their high historical and cultural value and the relatively strong protection of them.

Table 4: POI space aggregation analysis

Class	Average closest distance	The average closest distance	R	Z	P	Spatial distribution type
Population	118.02	352.24	0.33	-45.56	0	Significant aggregation
Official stylePOI	136.45	393.44	0.36	-40.36	0	Significant aggregation
Folk stylePOI	265.59	704.35	0.39	-19.78	0	Significant aggregation

3.2.3 Spatial distribution characteristics

Figure 2 shows the results of the kernel density analysis of the traditional style scenic spots POIs in the Fifth Ring Road of Beijing. The traditional style scenic spots POIs in the Fifth Ring Road of Beijing are centered on the Forbidden City Museum, and in the north, along the central axis from Di'anmenwai Street, and in the south, along the axis to Qianmen Street, which forms the spatial agglomeration area for the large-scale aggregation of the first-level traditional style scenic spots POIs in the Second Ring Road. Among them, the belt-shaped gathering area formed with Gongwangfu as the center and Nanluoguxiang as the center, the belt-shaped gathering area formed with Dazhalan commercial area as the center and Hubei Macheng Hall as the center, these two belt-shaped gathering areas are connected with the belt-shaped area formed with the Forbidden City as the center respectively to form the large-scale aggregation of first-class traditional style scenic spots POI in general. This area is the historical and cultural protection zone of the Imperial City of Beijing, which has a deep historical and cultural heritage since ancient times, and the traditional style of the city is remarkable, thus best reflecting the historical style of the city of Beijing.

The city develops radially from the center of the city, and the closer to the center of the city, the earlier the development, and the more ancient buildings there are. Within the second ring road of Beijing, the POIs of traditional style scenic spots are gathered in large scale and mainly concentrated in the northern part of the city, while the southern part of the city is relatively less distributed, and the distribution of the Forbidden City, Jingshan, Shichahai and Nanluoguxiang areas, Dashilan area and Dongsi area are the most concentrated.

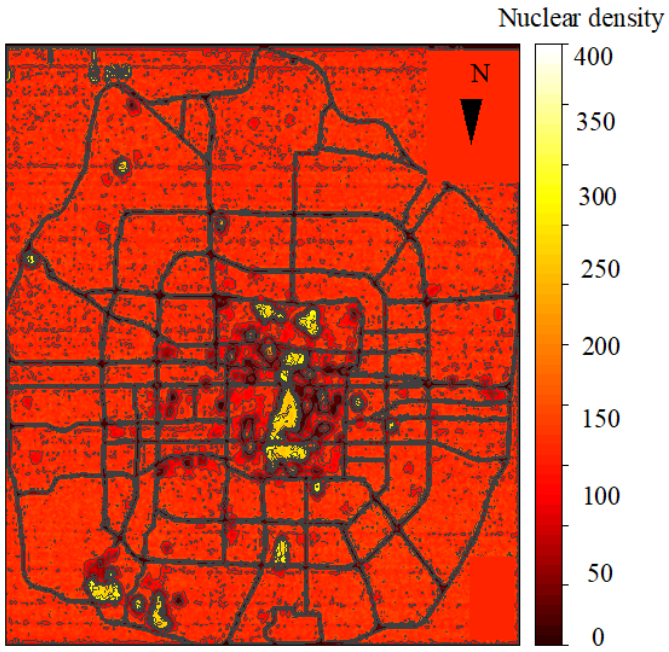


Figure 2: The analysis of the core density of poi in the traditional style of Beijing

Figures 3 and 4 show the results of kernel density analysis of POIs of scenic spots in official style and POIs of scenic spots in folk style in the Fifth Ring Road of Beijing, respectively. Compared with the official style, the number of POIs of scenic spots in folk style is relatively small, and they are mainly distributed in the periphery of the Second Ring Road, which is related to the fact that the buildings in folk style have a low status of power and are protected less than the buildings in official style during the historical changes.

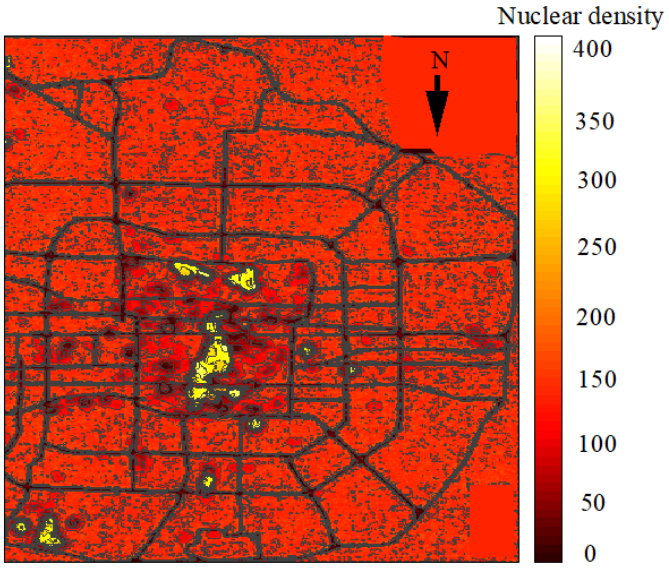


Figure 3: Analysis of the core density of scenic spots in Beijing

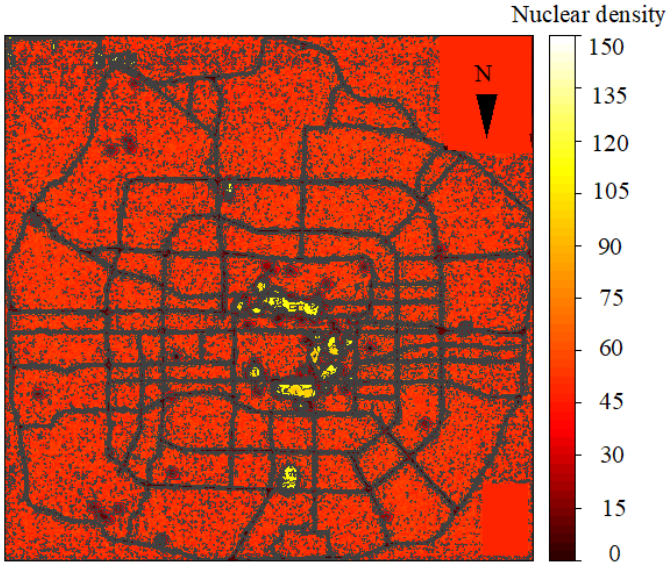


Figure 4: Analysis of the folk style scenic spot poi

3.3 Analysis of urban patterns based on spatial syntax

3.3.1 Space syntax

Based on the basic principle of spatial syntax, this paper compares the urban street patterns of the four main rings, namely the second ring, the third ring, the fourth ring and the fifth ring, horizontally, which can show the spatial characteristics of the four main rings more clearly. Specifically, three measurement variables are used: integration degree, selectivity, and comprehensibility. By analyzing the degree of selectivity and integration of Beijing's loops, the potential for traversing traffic and arriving traffic between Beijing's city streets is explored. Then the degree of comprehensibility of Beijing's urban space is determined by analyzing the degree of global integration and the degree of local integration.

Integration Degree: After determining the radius, the spatial syntactic integration degree is used to predict the arrival traffic potential of the streets.

$$RA_i = \frac{2(MD_i - 1)}{n-2} \tag{8}$$

$$I_i = RRA_i = \frac{n\left[\log_2\left(\frac{(n+2)}{3}\right) - 1\right] + 1}{(n-1)\left|MD_i - 1\right|} \tag{9}$$

$$RRA_i = \frac{RA_i}{D_n} \tag{10}$$

$$I_i = \frac{1}{RRA_i} \tag{11}$$

In the formula where MD refers to the average depth value, n is the number of all the nodes, RA_i is the global integration degree, and I_i represents the local integration degree.

Degree of Selection: The possibility of each road being selected is calculated by the analysis of the software, according to different radii, the possibility of the road being selected by passers-by or vehicles is calculated, which represents the traversing traffic potential of the road.

$$\frac{1}{(N-1)(N-2)} \sum_{j=1,k=1;j\neq k=1}^{N} \frac{n_{ik}(j)}{n_{ik}} \tag{12}$$

Intelligibility: In spatial syntax, intelligibility is defined as the number of linear relationships between the overall integration of axes and the connectivity of axes. It expresses the degree of correlation between the local and the whole of the system, and the spatial intelligibility is judged by the fact that the residents can infer the pattern of the whole space from the local in a certain radius. Comparing the global integration degree and local integration degree, the scatterplot of the comprehensibility analysis can be obtained, when the correlation coefficient R^2 value of the obtained scatterplot $0.3 > R^2 > 0.5$ indicates that the correlation between the two is low, the value of $0.5 > R^2 > 0.8$ indicates that the correlation is moderate, and the value of $R^2 > 0.8$ indicates that the correlation is high.

$$R^2 = \frac{\left[\sum\left(C_i - \overline{C}\right)\left(I_i - \overline{I}\right)\right]^2}{\sum\left(C_i - \overline{C}\right)^2 \sum\left(I_i - \overline{I}\right)^2} \tag{13}$$

The C in the formula denotes the spatial connectivity value and I is the spatial global integration.

3.3.2 Results of the urban pattern analysis

The numerical analysis of spatial syntax is shown in Table 5, and the scatter plots of comprehensibility of different rings are shown in Fig. 5, and Figs. (a) to (d) are the comprehensibility of the second ring, the third ring, the fourth ring, and the fifth ring, respectively. The main characteristics of Beijing's urban pattern are as follows: the urban area of the Third Ring is the largest among the four major rings, and the urban area of the Fifth Ring is the smallest. Spatial syntactic analysis shows that the fifth ring has the highest average value of selectivity, and the smaller area allows the city streets to have strong traversability. The third ring city has the highest average global integration degree and partial integration degree, and the third ring city has a larger area and a rich road network, which makes the third ring capital city streets have the highest potential for reachable traffic and the highest degree of urban recognition.

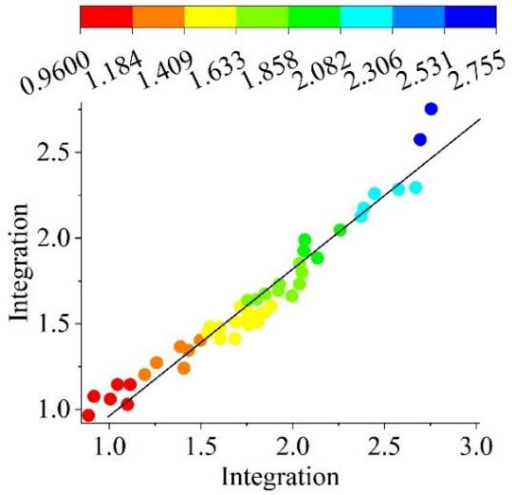
By comparing the global integration degree and local integration degree, the correlation coefficient R2 of the scatterplot is obtained. from the figure, it can be seen that the four cities R2 values are greater than 0.8, which are in the higher level of correlation, indicating that the city's intendability is higher, among which the third-ring city's value is the highest, which is 0.999, indicating that the third-ring's city has the best fit.

The analysis of the spatial syntax shows that the streets with high selectivity and integration of the second ring cities are located in the north of the city, and the north city streets have higher traversability potential and popularity, and are surrounded by buildings that cover the main typical landmarks and religious buildings of the city. The streets with high integration and selection of the third ring are to the southwest, where the main tower, religious and administrative buildings of the third ring are established. After analyzing the spatial system of

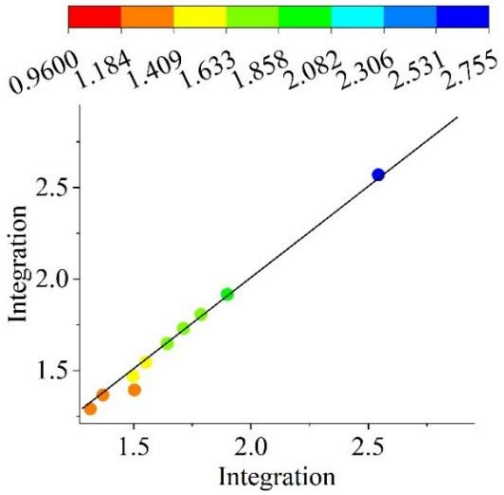
the fourth ring, the street with high integration and selectivity is to the east, and this section of the street has better penetration. After analyzing the spatial system of the fifth ring, the street with high integration and choice is in the northeast, and the street connects the inner and outer streets of the city, which makes it better accessible, and there are also a large number of religious buildings, administrative buildings, and cultural and educational buildings in the surrounding area, and the section of the street has better penetrability. The layout of the four cities shows that military-related buildings, mainly in the left half of the cross intersecting streets, while temples, cultural and religious buildings are interspersed among them.

Table 5: Numerical analysis of spatial syntax

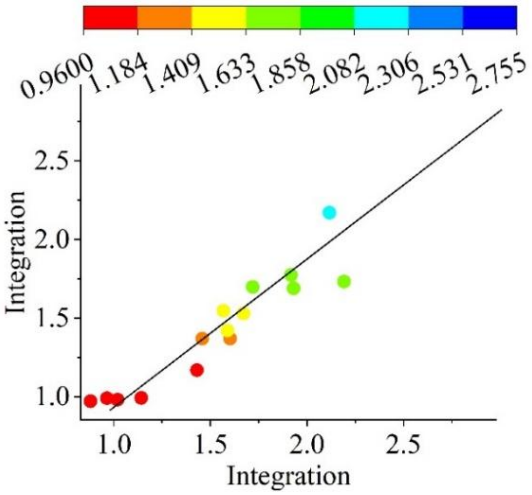
	Global integration	Local conformity	Comprehensibility	Selectivity
Two rings	1.585	1.786	0.932	0.086
Three rings	1.908	1.924	0.999	0.133
Four rings	1.536	1.657	0.938	0.157
Five rings	0.975	1.082	0.915	0.235



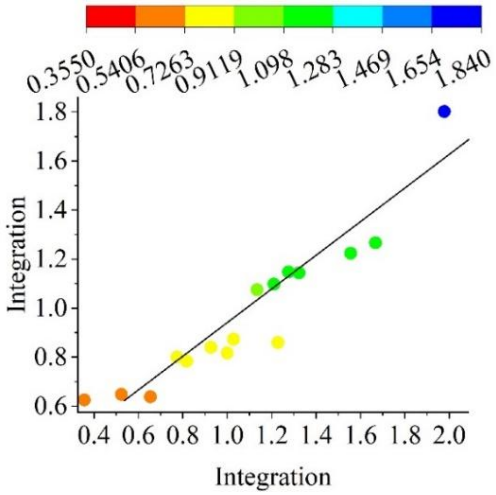
(a)The understanding of the two rings



(b)The understanding of the three rings



(c)The understanding of the four rings



(d)The understanding of the five rings

Figure 5: Comprehensibility

4 Urban landscape perception based on color theory

4.1 Color Related Theories

4.1.1 Color Properties

Color is both an abstract definition and a figurative expressive feature. The first exposure to color was in the process of painting, where words such as shade, light and dark, warm and cool, and ambient colors describe the colors conveyed to us by a still life. Every color is not pure and there are differences between colors. Colorists have defined three elements of color, the attributes of color. They can be summarized as: hue, lightness, and color.

Hue is the primary characteristic of color, that is, all kinds of color phase designation, that is, refers to the name of the color, is to distinguish between a variety of different colors of the most accurate standard, each of the different colors have a different hue, such as red, yellow, blue, green and so on. Hue is determined by the wavelength of light. It is generally based on the pure color on the hue ring.

Brightness refers to the degree of depth of a color, and brightness is a property of every type of color. In painting practice, it is found that the more white pigment is added to a color, the brighter the color becomes, and on the contrary, the higher the black content in a color, the darker the color becomes. The degree of darkness of a color is its brightness.

In painting, it is common to speak of pure or impure colors as the chromatic scale in color science. One of the concepts involved here is that black, white and gray have no chroma. Colors like red, yellow, and blue are the ones that have chroma. The purity status of a color directly affects the sense of excitement that the color gives.

4.1.2 The Menzel color system

Menzel color system is the earliest study of a color expression method, focusing on the study of color classification and calibration, color logic psychology and visual characteristics, etc. Menzel color system using three-dimensional similar to the spherical spatial model of the color of hue, lightness, chromaticity of the three elements of the entire performance, also known as color stereo. Color stereo horizontal coordinates for a variety of colors hue ring, in the horizontal direction, to red (R), yellow (Y), green (G), blue (B), purple (P), red violet (RP) and so on the basis of ten colors, and then each class of hue is divided into 10 equal parts. Vertical axis for the brightness, from low to high for a number of grades, a total of 11 levels, the lowest value of 0, is black, the highest value of 10, is white, the middle of the 1-9 is the so-called gray range. Extending outward from the center of the axis is the purity, the greater the distance from the center of the axis, the higher the corresponding color purity becomes. Conversely, the closer to the center axis, the lower the purity of the color, the closer to gray. The purity on the central axis is 0. In the Munsell color system, due to the specificity of different colors, the highest purity and brightness of different hues have some variability.

The Munsell color system is usually widely used in the process of research and collection of color data, the Munsell color card and the target object color for numerical comparison and matching. Business culture is constantly developing, and the requirements for corporate culture and image have become important indicators. Various types of chain enterprises emerge one after another, in order to maintain a unified business image, such as fast hotels, supermarket chains and other building colors consciously decorated into the same color, color without a unified numerical standard, it is inevitable to produce differentiation of the same color, the Menzel color system can be used in the specification of the business brand color scale value.

4.1.3 Basic principles of color in business

Commercial buildings and other types of architectural color difference lies in its impact on consumers, this impact includes psychological and physiological. Color itself is a medium of transmission, the use of color in the process there are many forms of beauty presentation, the use of these color techniques can be summarized as follows:

(1) Contrast and Harmony

Color contrast is a variety of colors gathered in two-dimensional and three-dimensional environment, compared with each other and produce differences in color relationships. Color contrast through the combination of different colors, strengthen or weaken part of the color, set off the characteristics of other matching colors. Often, in order to attract the visual attention of the target audience, billboards will use colors with a higher intensity of color contrast with the exterior of the commercial building. Harmony refers to a variety of colors through the matching, sequencing, blending and other ways to derive a soft, coordinated and unified color relationship. The method of color harmony is conducive to the diversity of commercial color unity and change.

(2) Symmetrical balance

Symmetry and balance refers to a relatively balanced and stable composition and combination of colors in the visual form. Today's commercial buildings due to the over-pursuit of color splendor, often ignoring the matching with the building's inherent volume, area, form, color, making the symmetry and balance of the building is greatly reduced.

(3) Primary and secondary reality

Primary and secondary reality can be said to be both spatial and distance concepts, as well as the number of contrasts, strong and weak differences in the concept. Macro scale observation of the building as a perspective, will produce the visual first color, that is, the main color, the meso scale for observation, through perception and judgment to get second to the main color of the auxiliary color, micro scale in the building will get the color nodes, that is, the color of the embellishment. The area of color plays an important role in visual communication in the external façade of the building, the commercial building color can be coordinated according to the conditions of the building's volume and spatial environment, the proportion of the color area, so as to control the color of the sense of stability and liveliness.

(4) Rhythm

Rhythm is an auditory description of the regular dynamics of music. The rhythm of color is the expression of visual order, continuity and intermittent movement. Different colors together will produce different rhythmic rhythm. Hue, brightness, purity in accordance with certain basic guidelines, in order to stage changes in the way, through the orderly arrangement of the composition of the color collocation, can reflect the color full of rhythm and rhythm of the sense of flow.

4.2 Analysis of color data at the regional level

The analysis of roof color hue proportion is shown in Fig. 6, according to the roof color mentioning samples of all buildings in Beijing, in which the hue is 62.51% in yellow-red system (YR), 14.34% in red system (R), 6.24% in yellow system (Y), 5.63% in red-purple system (RP), 4.02% in blue-purple system (PB) and so on. It can be seen from the ratio of roof color phases in Beijing that the overall architectural roof color landscape landscape in Beijing has a high degree of unity, and the vast majority of architectural roof colors are in the yellow-red, red, and yellow phases, but there are also a certain number of roofs with colors in the blue-violet phases. From the point of view of the distribution of each color phase in the space, the overall distribution lacks regularity, except for the concentrated distribution in a small area. According to the on-site research, most of the roofs in Beijing are jiageng tiles, wave tiles, slate tiles, flat

roofs, etc. The vast majority of the buildings are consciously participating in the coordination of the architectural style of Beijing, but due to the differences between the old and the new building materials, the influence of the building materials by the natural erosion, and a small number of buildings that do not pay attention to the coordination of the style of the building and the phenomenon of illegal building, etc., the color of the roofs of the buildings presents the color differences in the color hues.

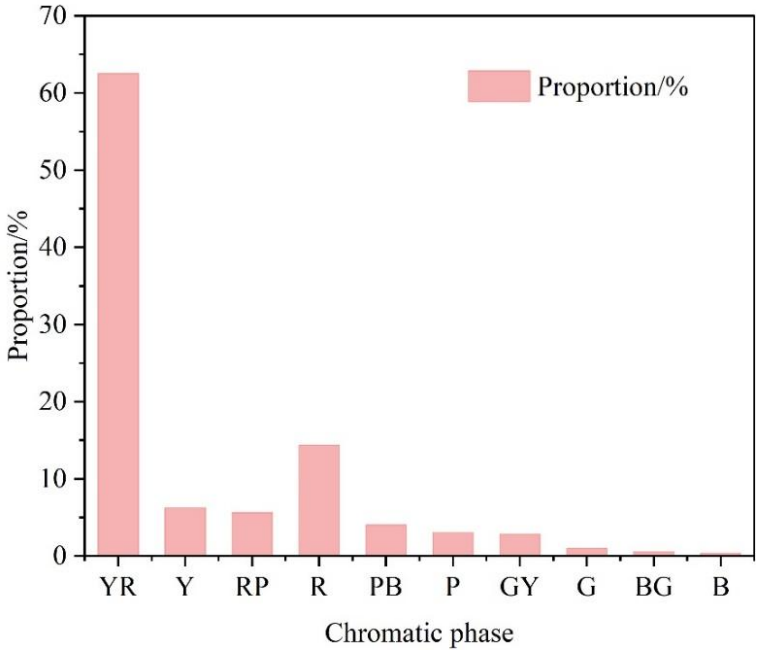


Figure 6: The roof is dirty compared to the analysis

The results of wall color chromatic weighting analysis are shown in Figure 7, based on the color lifting samples of the main color of the wall of all buildings in Beijing, in which the yellow-red system (YR) is 34.46%, the red system (Y) is 15.85%, the red system (R) is 15.74%, the blue-violet system (PB) is 9.94%, the greenish-yellow system (GY) is 8.36%, the reddish-purple system (RP) is 3.78%, the violet system (P) 3.73%, the colorless system (N) 2.35%, and the blue system (B) 2.36%. P) 3.73%, colorless (N) 2.35%, blue (B) 2.36%. From the viewpoint of the ratio of the main color hue of the wall of all buildings in Beijing, the yellow-red system still occupies the largest proportion, mainly influenced by the red brick wall construction materials, but compared with the proportion of the red-yellow system accounted for by the roof color has been reduced a lot, this is mainly related to the architectural form of Beijing, a large number of Western-style buildings are used in the red-yellow system of the roof of the tiles, the wall color is dominated by the stone, plaster and red bricks. From the architectural form caused the difference in the proportion of color.

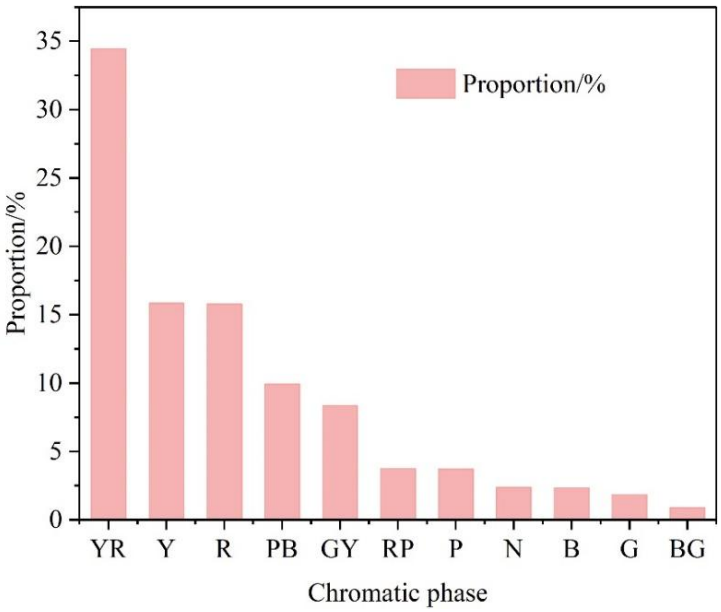


Figure 7: The wall is dirty compared to the wall

Roof color luminance value distribution as shown in Figure 8, from the distribution of roof color luminance value, color luminance value are fluctuating in the range of medium luminance, that is, luminance value of 3-7.5 range of the distribution of the most, the roof building color luminance value is more concentrated in the distribution of the interval of 4.5-5.9. However, the distribution of roof color luminance values above 7.5, i.e., showing high luminance, is not a minority, especially the part above 8.0. To a certain extent, this reflects some of the problems in the color landscape of Beijing, although the distribution of the distribution of high concentration from the perspective of hue, but from the perspective of luminance, the distribution of the concentration of the degree of low relative to the hue, and the distribution of the number of high luminance and low luminance, which has a certain impact on the harmony of the color landscape.

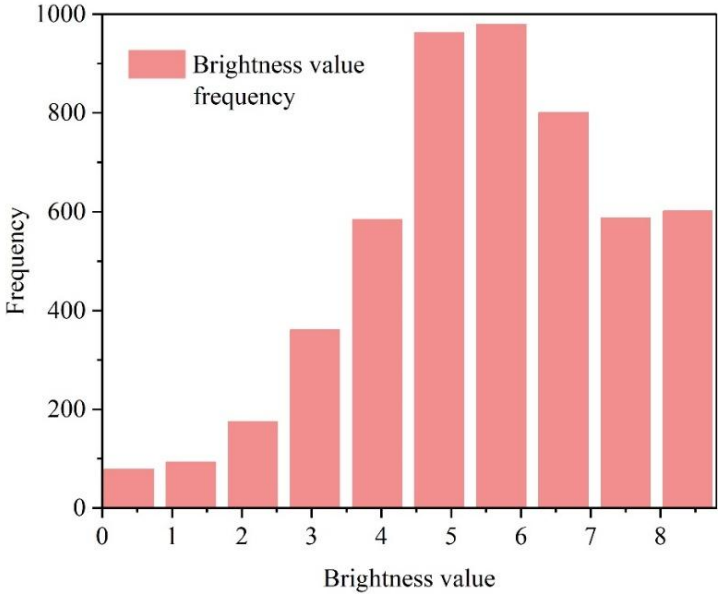


Figure 8: The roof color brightness value distribution

The distribution of wall color luminance values is shown in Figure 9, from the distribution of wall color luminance values, the color luminance values are mainly concentrated in more than 6.0, which belongs to the distribution of high luminance colors, and a large number of color distributions are located in more than 7.5, which is related to the fact that plaster is used in large areas of the walls of many buildings in Beijing. The distribution of luminance values of the main wall color and the roof color show significantly different characteristics. At the same time, the degree of dispersion of the luminance values is large, showing a strong contrast. At the material level, the contrast between the red brick walls and the plastered walls is strong. This is not only a distinctive feature of Beijing, but also an entry point for remediation of the color and landscape problems.

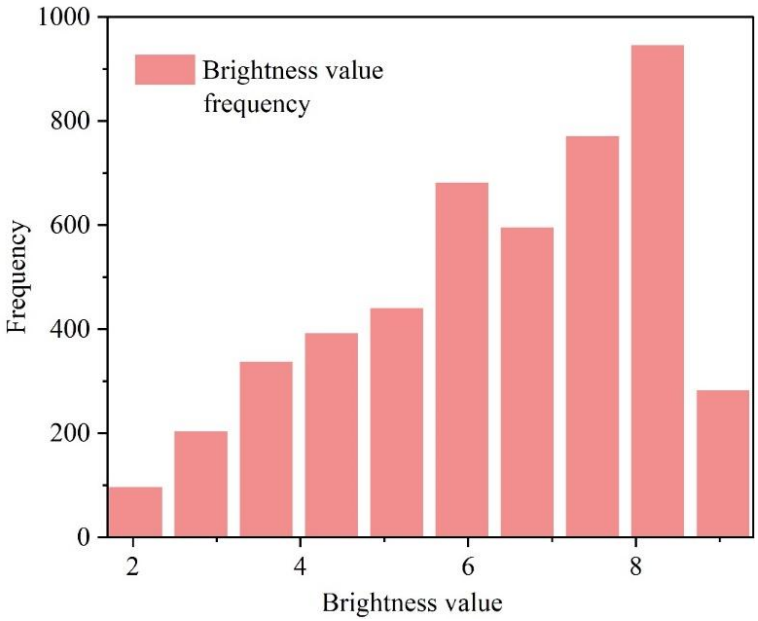


Figure 9: The color brightness of the wall is distributed

The distribution of roof color degree values is shown in Figure 10, from the distribution of roof color degree values, medium and low color degree accounted for the vast majority of the part, that is, the color degree value is lower than 5.0, the number of low color degree of the roof color is higher, that is, the color degree value is lower than 3.0 part of the part. There exists a portion with high colorimetric values. From the numerical distribution characteristics, the overall harmony is high, the color degree value presents if contrast, there is a certain number of roof color color degree value distribution destroys the overall color degree harmony. From the viewpoint of spatial distribution characteristics, the distribution of chromaticity values distinguishes a certain regularity, and it is difficult to further feel the color characteristics and architectural distribution of Beijing from the numerical value part of the color. There is room for optimization and improvement.

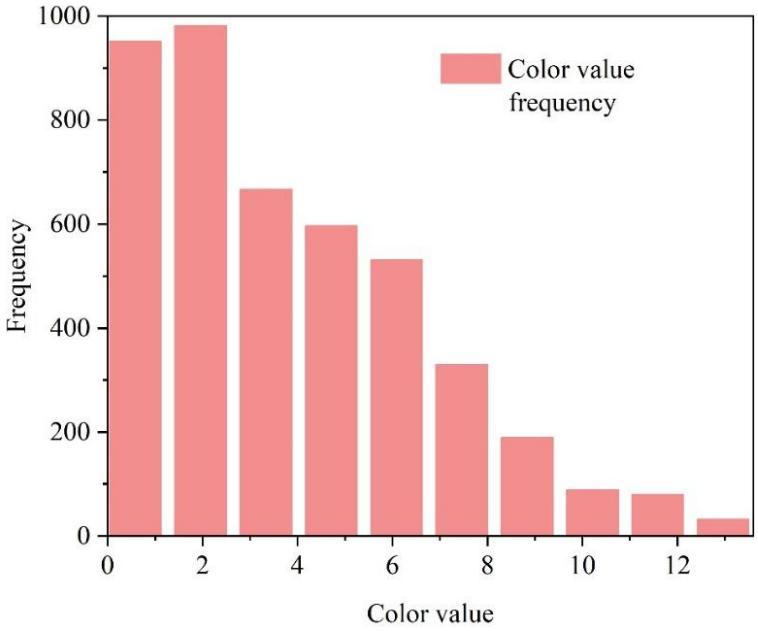


Figure 10: Roof color distribution

Beijing wall color degree value distribution as shown in Figure 11, from the analysis of Beijing wall color degree value, color degree value value distribution in low and medium color degree is dominated, and compared to the distribution of color degree value of the roof color is more inclined to low color degree, that is, less than 4.8 color degree is dominated, and in the low color degree, but also to the color degree value of the value of less than 1.5 is the most values. Only a very small proportion of the color values are above 7.0. This is one of the more distinctive features from roof color. From the spatial point of view of the distribution of chroma values, the distribution lacks regularity and the regional characteristics are not obvious.

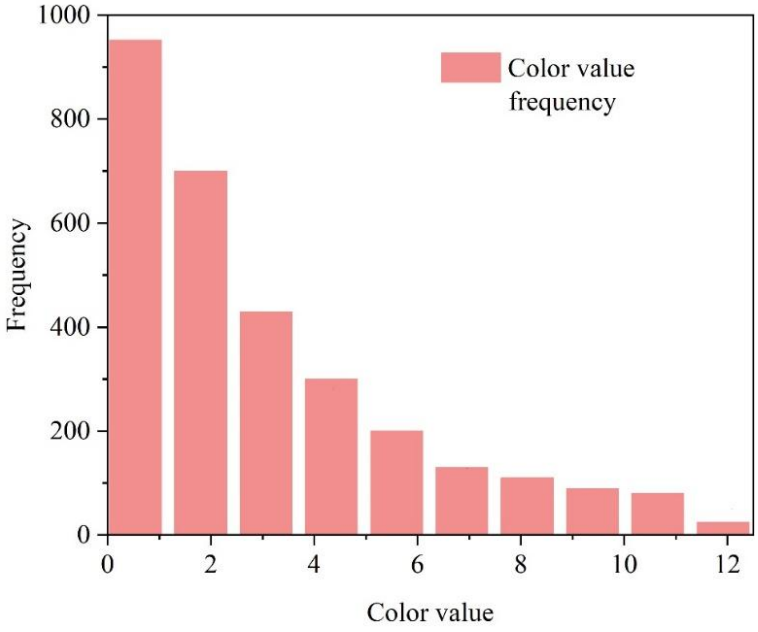


Figure 11: Color value distribution of metope

5 Conclusion

Urban architectural landscape perception is of great significance to the national goal and task of “improving the level of urban design and shaping the city's characteristic landscape”. In this paper, we explore the multi-angle perception of urban architectural landscape based on urban points of interest, urban streetscape images, and architectural color data, relying on deep learning methods and spatial data analysis techniques.

The improved Mask-RCN building recognition model constructed based on the deep learning method performs well on the dataset produced in this paper, achieving an average accuracy mean of 91.36% on the basis of a short inference time, which can be used for the work of intelligently recognizing the building types of large-scale urban building image datasets.

The density of urban building style POIs reaches the highest in the Second Ring Road, with a value of 12.77 POIs/km², reflecting the deep imprint of architectural history in the Second Ring Road. The spatial syntactic analysis further reveals that the Third Ring Road network possesses the highest global integration, suggesting that it has a better transportation potential and is a key structural tool for connecting the historical kernel with the modern urban area and carrying the perception of the landscape. However, the distribution of the landscape shows the phenomenon of dense in the north and sparse in the south, which suggests that spatial equity should be taken into account in shaping the architectural landscape of the city.

The roof and wall colors of Beijing buildings are dominated by the yellow-red color, which accounts for 65.51% and 34.46% respectively. This reflects the collective influence of Beijing's traditional architectural glazed tiles and red bricks, and the sporadic distribution of heterogeneous color phases such as blue-violet system, which reflects the problem of detuning of the landscape caused by the change of building materials with the change of seasons and natural weathering.

Based on the above conclusions, the design of the sustainable development strategy can be completed from the construction of an intelligent multi-data landscape monitoring platform, the implementation of the spatial guidance strategy for the revitalization of the district, and the formulation of the urban color control rules, so as to make the historical heritage of the buildings and the contemporary development of the buildings coexist harmoniously under the sustainable framework.

References

[1] Han, L., Wei, Y., & Xi, G. (2025). Urban vitality in the context of shrinkage: spatiotemporal patterns and influencing factors at urban and regional scales. Humanities and Social Sciences Communications, 12(1), 1-16.

[2] Guan, X., Wei, H., Lu, S., Dai, Q., & Su, H. (2018). Assessment on the urbanization strategy in China: Achievements, challenges and reflections. Habitat International, 71, 97-109.

[3] Chu, Y. W. (2020). China's new urbanization plan: Progress and structural constraints. Cities, 103, 102736.

[4] Komasi, H., Hashemkhani Zolfani, S., Prentkovskis, O., & Skačkauskas, P. (2022). Urban competitiveness: Identification and analysis of sustainable key drivers (a case study in Iran). Sustainability, 14(13), 7844.

[5] Ghahremani, H., Afsari Bajestani, S., McCarthy, L., & Jalalianhosseini, M. (2021). Transformation of urban spaces within cities in the context of globalization and urban competitiveness. Journal of Urban Planning and Development, 147(3), 05021026.

[6] Atkinson, R. (2012). Urban governance and competitiveness: Improving 'urban attractiveness.'. Regieren: Festschrift für Hubert Heinelt, 297-312.

[7] Permyakov, M. B., & Krasnova, T. V. (2019, December). Architectural and design approaches to creation of comfortable urban environment. In IOP Conference Series: Materials Science and Engineering (Vol. 687, No. 5, p. 055062). IOP Publishing.

[8] Ossola, A., Locke, D., Lin, B., & Minor, E. (2019). Greening in style: urban form, architecture and the structure of front and backyard vegetation. Landscape and Urban Planning, 185, 141-157.

[9] Cao, Z., Mustafa, M., & Mohd Isa, M. H. (2024). Traditional Regionalism or Modern Minimalism? Unveiling the Psychological Impact of Architectural Styles in Sustainable Urban Planning. Sustainability, 16(13), 5576.

[10] Abd Elrahman, A. S., & Asaad, M. (2021). Urban design & urban planning: A critical analysis to the theoretical relationship gap. Ain Shams Engineering Journal, 12(1), 1163-1173.

[11] Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., ... & Schiele, B. (2016). The cityscapes dataset for semantic urban scene understanding. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3213-3223).

[12] Cohen, S. (2015). Cityscapes. In The Routledge Reader on the Sociology of Music (pp. 231-244). Routledge.

[13] Keshtkaran, R. (2019). Urban landscape: A review of key concepts and main purposes. International Journal of Development and Sustainability, 8(2), 141-168.

[14] Ananiadou-Tzimopoulou, M., & Bourlidou, A. (2017, October). Urban landscape architecture in the reshaping of the contemporary cityscape. In IOP Conference Series: Materials Science and Engineering (Vol. 245, No. 4, p. 042050). IOP Publishing.

[15] Czyńska, K., & Rubinowicz, P. (2019). Classification of cityscape areas according to landmarks visibility analysis. Environmental Impact Assessment Review, 76, 47-60.

[16] Ginzarly, M., Houbart, C., & Teller, J. (2019). The Historic Urban Landscape approach to urban management: a systematic review. International Journal of Heritage Studies, 25(10), 999-1019.

[17] Jiang, J., Zhou, T., Han, Y., & Ikebe, K. (2022). Urban heritage conservation and modern urban development from the perspective of the historic urban landscape approach: A case study of Suzhou. Land, 11(8), 1251.

[18] Ospangaliyev, A., Utebekova, A., Dosmanbetov, D., Akhmetov, R., & Mazarzhanova, K. (2022). Impact of urban landscaping on improving the sustainable development of the urban environment. The case of Nur-Sultan. Journal of Environmental Management &

Tourism, 13(5), 1459-1466.

[19] Asif, N., Utaberta, N., & Sarram, A. (2019). Architectural styles of Malaysian mosque: Suitability in compact urban settings. In MATEC Web of Conferences (Vol. 266, p. 06001). EDP Sciences.

[20] Tandon, M., & Bano, F. (2023). Comparative critical analysis of modern architectural styles. Indian J. Nat. Sci, 14, 57095-57106.

[21] Zhao, J. (2020, June). Influencing factors of urban architectural style in China: A case study of danzhou city, hainan province. In IOP Conference Series: Earth and Environmental Science (Vol. 510, No. 6, p. 062003). IOP Publishing.

[22] Ivashko, Y., Kuzmenko, T., Li, S., & Chang, P. (2020). The influence of the natural environment on the transformation of architectural style. Landscape architecture and Art, 15(15), 98-105.

[23] WANG, D. G., LYU, Q. Y., WU, Y. F., & FAN, Z. Q. (2019). The characteristic of regional differentiation and impact mechanism of architecture style of traditional residence. Journal of Natural Resources, 34(9), 1864-1885.

[24] Liu, W., Li, D., Pernice, R., Meng, Y., Wang, R., & Lu, C. (2025). Enhancing urban identity through the refined management of architectural styles: insights from Wuhan. Journal of Asian Architecture and Building Engineering, 24(4), 2509-2529.

[25] Ni, X. (2016, December). Study on Architectural Characteristics of New Art Style in Harbin. In 2016 3rd International Conference on Education, Language, Art and Inter-cultural Communication (ICELAIC 2016) (pp. 578-580). Atlantis Press.

[26] Hokhrin, E. V., & Smolkov, S. A. (2020, June). Regeneration of architectural style of the historical environment of the cultural and public downtown. In IOP Conference Series: Materials Science and Engineering (Vol. 880, No. 1, p. 012062). IOP Publishing.