applied sciences

MDPI

Article

City Architectural Color Recognition Based on Deep Learning and Pattern Recognition

Yi Zhuang [1,2] and Chenyi Guo [1,*]

[1] School of Art and Design, Shanghai University of Engineering Science, Shanghai 201620, China
[2] Shanghai International Design Innovation Institute, Tongji University, Shanghai 200092, China; zhuangyi@sues.edu.cn
* Correspondence: guocy516@gmail.com; Tel.: +86-18701729144

Abstract: The collection of information about buildings and their colors is an important aspect of urban planning. The intelligent recognition of buildings using image information plays a significant role in the development of smart cities and urban planning. This thesis proposes a building color-recognition technique based on morphological features utilizing convolutional neural networks and the K-means clustering algorithm of image-recognition technology. The proposed method can identify buildings in images and classify them into two categories, buildings and panoramas, for color extraction and matching. This method involves training convolutional neural networks on deep learning so that the buildings can be differentiated and segmented. Subsequently, the K-means algorithm extracts colors from the segmented building images. The extracted building category, color, and text information were analyzed to obtain a comparison and analysis results of buildings and panoramas. The results demonstrated that the system is capable of accurately segmenting buildings, as well as extracting colors from both buildings and panoramas separately. It can also contribute to the extraction and presentation of color schemes in smart city planning and provide valuable insights for the future development of urban colors.

Keywords: deep learning; pattern recognition; convolutional neural network; building recognition; building color

check for updates

Citation: Zhuang, Y.; Guo, C. City Architectural Color Recognition Based on Deep Learning and Pattern Recognition. *Appl. Sci.* **2023**, *13*, 11575. https://doi.org/10.3390/app132011575

Academic Editors: José António Correia and Jose Machado

Received: 15 August 2023
Revised: 1 October 2023
Accepted: 19 October 2023
Published: 23 October 2023

1. Introduction

China's White Paper on Smart Buildings, published in July 2021, emphasized the need to promote digital innovation and building scenarios for smart cities and strengthen digital applications. In the development and construction of cities, the forms of the development of buildings, including their color, shape, and arrangement, significantly shape the urban landscape and play a crucial role in facilitating its growth. Within the scope of design psychology, color holds particular importance as it influences the perception and image of a building, ultimately reflecting the character and identity of a city. Therefore, the use of image-recognition technology in the detection and analysis of the colors of urban buildings is an efficient way to understand a city's impression. Providing valuable data and theoretical support for future urban development and planning endeavors, it has become a crucial step in developing smart cities.

Existing research in this field can be categorized into three main areas. The first area focuses on building recognition within the context of smart city planning. Zou Zhichong et al. [1] discovered that identifying model buildings based on their plan outline features partially reflects the functional characteristics of said buildings. In their study, they identified features of the building's texture and subsequently classified the building. Additionally, Yan Longxu et al. [2] researched the use of satellite images to extract building contours and classify urban areas into urban and rural construction zones, studying the urban morphology within these areas. However, the proposed region extraction algorithm in their study was found to be complex and have a high demand for accurate building region identification. In the area extraction algorithm, the authors proposed a relatively cumbersome approach with high accuracy requirements, requiring extensive training to ensure the accuracy of the building area extraction. The second area of research concentrates on algorithms for building recognition. Zhang Hua et al. [3] employed MAEU-CNN technology to improve building recognition. However, they encountered the challenges of under-classification and misclassification during the training process, causing the authors to explore the incorporation of building color and shape as additional elements for classification. In a separate study, Xia et al. [4] proposed a CNN framework that successfully retrieved and differentiated building edges, overcoming limitations in image recognition clarity. Similarly, Taoufiq et al. [5] classified buildings based on two benchmarks: Functional purpose and architectural style. They established a predetermined hierarchy within the overall model and differentiated buildings based on color identification. The final area of research focuses specifically on the identification of building colors. Zijian Shi et al. [6] collected data on the colors of residential areas in the functional core of Beijing using GIS imagery, aiming to predict the overall color composition of the area based on the colors of the collected buildings. However, their study lacked a detailed analysis of the specific colors and details of the buildings within the area. Another study conducted by Jin et al. [7] highlighted the potential for the positioning of original buildings based on their exterior color. Based on convolutional neural networks, David Knox and his team identified regional buildings, pinpointing the positional differences between new and older structures within images [8].

Considering the social background and existing research, the extraction of architectural colors holds significant importance in developing smart cities. These extracted colors have diverse applications, including architectural recognition, urban form evaluation, and urban planning. However, previous research has primarily focused on utilizing color elements to assist in building edge detection or reverse positioning of buildings based on key colors. There is a notable gap in methodologies specifically concentrating on extracting building colors. Therefore, this thesis used urban landscape images to extract their architectural elements and colors, which were then applied to coordinate urban architecture with landscape colors. Deep learning and convolutional neural network algorithms were employed to extract and isolate buildings from scene images. By applying distinct segmentation techniques, the colors of both the buildings and the panoramic images were extracted separately, resulting in a more precise identification of building colors and architectural landscapes within the scenes. The proposed building color recognition method presented in this thesis supports the foundational data of urban planning and is a complementary tool to the theory of smart city planning.

2. Background and Related Work

Smart city planning is driven by the concept of "smart," using cutting-edge information technology to further the digitalization of cities and ensure the sustainable growth of various industries, as well as the ecological environment, while creating a city where citizens can engage in fulfilling activities [9]. To promote integration and creation in the planning and construction of cities, there is a need to improve cities' visual aesthetics, while color plays a pivotal role in capturing the essence of their form. This thesis proposes an image-recognition technique that combines deep learning, pattern recognition, convolutional neural networks, and K-means clustering algorithms to effectively analyze, calculate, segment, and extract colors from architectural images.

Deep learning is mainly active in feature extraction, and more complex feature expressions can be gradually synthesized with deep learning. Pattern recognition is primarily the process of describing, classifying, judging, and recognizing various things in the real world using a computer imitation of the human brain. Guo Chunhua et al. used Newton's deep learning optimization algorithm to enhance the accurate collection of building information and utilized a collaborative filter algorithm to provide users with maintenance methods [10]. Huang Yong et al. utilized deep learning to assess and rate the damage state of buildings, providing valuable assistance for post-disaster reconstruction [11]. Hyeongmo Gu and Seungyeon Choo used façade datasets and deep learning to automatically mark façade data and effectively generate large-scale data sets to understand the characteristics of the street [12]. Young-ha Shin et al. used PointNet++ to extract and segment buildings and found that the two echoes of the laser pulse could be fully identified and removed in said buildings [13]. The systems proposed in the thesis of Shi-Jinn Horng and Pin-siang Huang can identify the different states of a product with effective identification of the products [14]. Cai Wei and his team enhanced the precision of building recognition by extracting building contours from high-resolution satellite images [15]. Using deep learning within Faster RCNN, Zheng Lijuan and her team achieved superior detection outcomes in drone remote sensing imagery [16]. The majority of building recognition methods rely on the extraction and identification of building contours from satellite or aerial imagery. Gao Zhiheng and colleagues approached building recognition using Internet maps, translating the recognized images to text via the CRNN algorithm [17]. Depending on the application of architecture or plane, deep learning can also be applied to various fields. Ahmed Jawad A. Aibdairi et al. classified people into race and ethnic groups in their study. The recognition system, based on FPGA (DE1-SOC) methods, could accelerate recognition and reduce power [18]. Domenico Buongiorno et al. used non-destructive detection methods to detect and classify welding defects. Combining different classifiers and three different feature groups was able to achieve good results in predicting weld resistance defects [19]. Wei Zhuang improved accuracy by labeling the critical points of human bones in the case of human dynamics [20].

The role of the convolutional neural network primarily focuses on the input and output layers of data. By importing and reading captured images, deep learning algorithms analyze numerous building images from various angles to extract their specific features. Based on these features, the buildings within the images are then framed, and the building area is thus identified in the output layer and distinguished from the background area. Wu Yanqing et al. designed a picture retrieval framework with a fusion of attention mechanisms for images of different patterns in overseas Chinese architecture, improving accuracy [21]. Santo et al. classified traffic signs on streets and highways to detect the differences between traffic signs and studied some factors affecting detection [22]. Wu Bei and Xiao Li implemented a method of extraction of crop neural networks using the color characteristics of the color histogram feature extraction algorithm [23].

Deep learning, as evidenced by the references, primarily focuses on its versatile applications across different domains and optimizing diverse datasets [24,25]. In the realm of architecture, the predominant emphasis lies in extracting building outlines from satellite imagery to enable easy demarcation of urban regions. However, there has been comparably less emphasis on leveraging deep learning techniques to explore and enhance urban architectural aesthetics. Within the realm of smart city planning, based on deep learning, convolutional neural networks enable the evaluation of the appearance of building texture, style, and color through image or video data, facilitating real-time monitoring and intelligent planning of smart cities. Additionally, in the age of iterations, it ensures the functional and visual effectiveness of urban architecture, enhancing the city's integration and promoting its distinctive characteristics while upgrading and updating the related intelligent computing models for city perception. Moreover, monitoring building patterns provides valuable guidance for subsequent planning and design improvements in smart city development, as well as predictive insights for personalized city development.

Applying the K-means clustering algorithm in smart city planning enables the extraction of a specific number of colors from urban buildings, allowing for the analysis and calculation of their color ratios. In smart city planning, the development of digitalization is essential, and the intelligent visual representation of the city also plays a vital role. Color serves as the language of a city, reflecting its cultural traditions and spirit. Well-chosen urban colors contribute to the city's development and communicate its character. Collecting urban architectural colors facilitates the creation of an accurate portrayal of the city in the metaverse, forming a comprehensive color scheme based on specific color data. Lang Xianmei et al. classified and analyzed traffic driver's behavior using the K-means algorithm [26]. Tu et al. enhanced and upgraded the grayscale model of each clustering class based on the K-means algorithm [27]. Guo Chaofan et al. segmented a corn blade image based on the K-means algorithm to reduce the impact and improve accuracy using the HSV color space's H component [28]. In the realm of remote sensing imagery, Ibrahim El rube utilized K-means and PHQ to balance image quality with expedited execution time [29].

By employing image-recognition techniques based on convolutional neural networks and K-means clustering algorithms, the proposed approach leverages the deep learning of architectural features to automatically identify urban buildings and extract their colors and background colors. This establishes a comprehensive system for color extraction in urban buildings. From the perspective of related research, existing deep learning-related technologies are suitable for various types of recognition, such as architecture and facial recognition, and have achieved good recognition results. This paper selected buildings through deep learning and pattern recognition to extract their colors after selecting them. In the future progression of smart cities, the extracted data results from this system can provide valuable insights for city or regional planning, further enhancing the development of smart cities in terms of visual aesthetics and urban management.

3. Research Methodology

Shanghai, a major city in China, is undergoing meticulous planning and regulation of its architectural colors as an essential landscape component, pioneering the integration of smart urban planning systems. Presently, the utilization of these smart systems is centered around defining urban zones, artificial facilities, and traffic detection. However, the recognition of architectural colors is still limited, often requiring color collection through photographs or manual measurements. Even though differentiated planning has been applied per area, there remains a gap in showcasing diverse architectural colors or overall color cohesion in urban planning guidelines. This research primarily investigated Shanghai's buildings, capturing their images and segmenting the structures and landscapes within intricate photos created by blending architectural and natural scenery. Through advanced computer image processing methods, colors were extracted. This approach offers a visual means to assess the harmony between architecture and landscape colors across various settings, laying a practical foundation for color planning. Concurrently, this system enhances the color identification aspect of the smart city planning system. The associated image processing technologies encompass deep learning, convolutional neural networks, and K-means clustering algorithms.

3.1. Deep Learning

The urban building recognition system begins by collecting relevant information about buildings and applying deep learning techniques to the collected images. Extensive testing has shown that utilizing approximately 20–30 images enables the system to accurately identify characteristic buildings. In this system, 30 images are selected for deep learning. These images are then inputted into the system, and the building features are extracted and analyzed using an algorithm. Based on these features, the system assigns a building type to each image. Figure 1 illustrates this process. The system focused on three types of buildings in this case: Residential buildings, special library buildings, and office buildings, as depicted in Figure 2. This thesis adopted a single-layer neural network as its chosen implementation technique, enhancing the deep learning capacity for handling complex functions by incorporating non-linear functions into the activation function. Subsequent chapters provide detailed explanations of each step involved in the process.

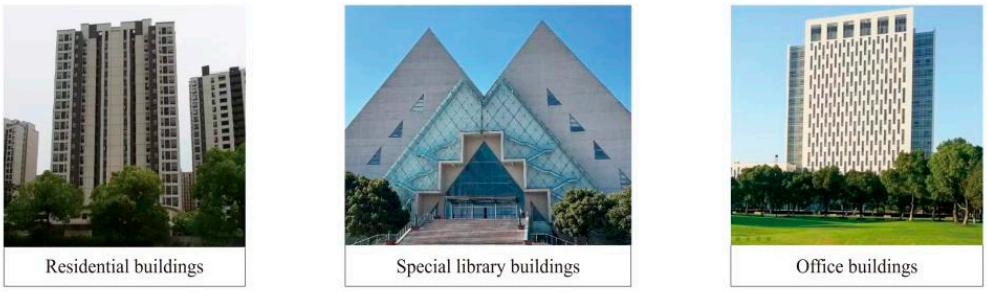

Figure 1. Residential buildings, special library buildings, and office buildings.

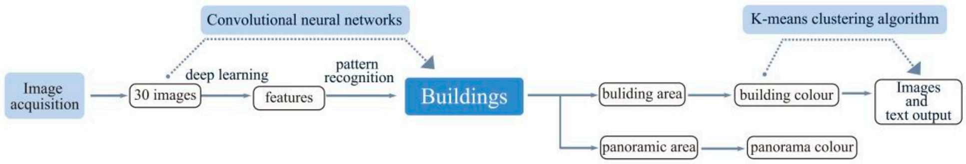

Figure 2. Project flow chart.

Based on the building recognition framing, the buildings are categorized into two areas: The building area and the panoramic area. The original image is then segmented and stored based on the building recognition framing; the image is therefore stored in two parts: The building image and the panoramic image. Finally, using the K-value setting, the system extracts colors from both images and assigns corresponding color labels to each image after determining the building color and the panorama color.

YOLO's primary goal is feature extraction. YOLO-V4 introduced a markedly different feature extraction method than its predecessors (from YOLO-V1 through to YOLO-V3). In contrast, YOLO-V5 incrementally refined performance, building on YOLO-V4′s foundations [30]. Given the requirements of this thesis regarding building recognition, YOLO-V5 was chosen as the platform to deploy deep learning and discern architectural images.

3.2. Explanation of the Convolutional Neural Network Algorithm

3.2.1. Single-Layer Neural Network

In this deep learning algorithm, feature extraction is accomplished using a single-layer network structure. During training, the algorithm captures many feature values, stated as x_1, x_2, to x_n, and applies the function a = g(z) to evaluate the features of various objects. This enables the detection of buildings and facilitates the distinction between different building categories. Equation (1) illustrates this process, where $z = WT + b$ represents the linear regression prediction, which is a prerequisite needed to introduce the Sigmoid function.

$$z = W^{\Lambda}Tx + b \; (x = x_1 x_2 \ldots\ldots x_n) \tag{1}$$

3.2.2. Activation Functions

Using an activation function, which introduces a non-linear function into the algorithm, enables to learn complex functions in depth. In this thesis, both the Sigmoid (2) and tanh (3) activation functions were utilized, as they exhibit S-shaped saturation characteristics. The addition of the ReLU (4) slope function further enhances the effectiveness of neuronal screening. The Sigmoid function, represented as $f(x) = 1/(1 + e^{-x})$, produces a constant positive output value. Consequently, weight updates are restricted to the same direction, affecting the convergence speed. On contrary, the tanh function, $(e^x - e^{-x})/(e^x + e^{-x})$, centered around zero, facilitates rapid convergence without significantly influencing the loss value. Additionally, the ReLU function, $f(x) = \max(0, x)$, allows a subset of neurons to produce zero output, leading to a sparse network. This mechanism helps address overfitting issues and reduces parameter interdependence.

$$f(x) = \frac{1}{1+e^{-x}} \quad (2)$$

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (3)$$

$$f(x) = \begin{cases} \max(0,x), x \geq 0 \\ 0, x < 0 \end{cases} \quad (4)$$

3.2.3. Logistic Regression-Based Gradient Descent

The non-linear function requires gradient descent using logistic regression. In the logistic regression approach, a suitable prediction function is chosen, such as the Sigmoid, tanh, or ReLU function, as described above. The three functions divide the distribution values into two distinct regions, as depicted in Figure 3. A loss function is then constructed to quantify the discrepancy between the prediction function and training data labels. This loss reflects the difference between the predicted and actual selections in the training set. By considering the "loss" of all training data, we can evaluate the deviation of the training data from the actual categories. Through an optimization process, the best parameter value is determined. Smaller parameter values correspond to higher accuracy of the prediction function and yield improved training results [24].

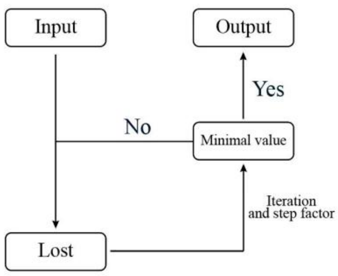

Figure 3. Logistic regression gradient training graph.

3.3. Analysis of the K-Means Clustering Algorithm

At the core of the K-means clustering algorithm lies the process of partitioning data into K-groups [31]. Initially, K-objects are randomly selected as cluster centers, and the algorithm proceeds to calculate the distances between the data points and K. The greater the similarity between the center of the cluster (K) and the surrounding data, the greater the distance, and vice versa.

To compute the distance, this paper employed the Euclidean distance (5) measure, which quantifies the distance between a cluster center and the distributed data. The formula for determining the distance of a data object from the cluster center within the space is illustrated in Equation (5):

$$(x,y) = \sqrt{\left((x_1-y_1)^2 + (x_2-y_2)^2 + \ldots + (x_n-y_n)^2\right)} = \sqrt{\sum_{I=1}^{n}(x_i-y_i)^2} \quad (5)$$

where x is regarded as the data object, y is regarded as the cluster center, and n is regarded as the dimension of the data object. The clusters generated by this classification calculate the distance to the other uniformly worthy minimum points within the cluster, treating the center of mass as the reference point until it no longer changes.

4. Framework

4.1. Image Data Processing

The initial step in building recognition involves gathering image data of building objects. This thesis explored two approaches to test the effectiveness of architectural feature recognition. The first approach involves using approximately 500 images of buildings and landscapes, without any specific categorization, in an attempt to identify general buildings. All 500 images used in this test experiment were captured by the author. In the second method, the same type of building is photographed from all angles, and around 30–50 images of the building are identified. In the first method, some identification of architectural features can be achieved. However, this approach yields relatively low accuracy in correctly identifying buildings in panoramic images and fails to classify different types of buildings effectively. Therefore, the second approach was adopted in this thesis to photograph individual buildings and classify them, which involves capturing images of each type of building, such as residential buildings, special library buildings, and office buildings, from various angles. By focusing on specific building types and conducting initial identification of building areas, the system was trained to study the building objects. First, for each type of building, 30–50 images were selected to train the deep learning model on building feature extraction. The image processing procedure primarily involved training the deep learning model on buildings and extracting features specific to each building type. Figure 4 illustrates the training process for different types of buildings, where rectangular bounding boxes were utilized to locate buildings in the images and collect their respective features for subsequent classification. Figure 4 provides a glimpse of some images used for training the feature recognition of administrative buildings.

Figure 4. Image of the training phase of office building feature recognition.

4.1.1. Building Recognition

Once the model was established, three types of buildings were randomly photographed from various angles and under different weather conditions to evaluate the accuracy of building recognition after training. The system's interface is designed. By utilizing the recognition function of the system, the buildings were imported for detection, resulting in successful identification of the buildings in the images and clear categorization of the objects. Additionally, in the presentation of the resulting experimental images, the accuracy of the building recognition was calculated based on the deep learning results and displayed in the rectangular bounding boxes. The accuracy value ranged from 0 to 1, with a higher value indicating a higher accuracy in building recognition.

The experimental results show that the recognition accuracy varied when dealing with building images containing occluded objects or captured from different angles. Notably, images with larger occlusion areas tended to exhibit lower recognition accuracy. Figure 5 displays this scenario, where the accuracy of building recognition was only 0.32 due to the large occlusion caused by trees. Conversely, as shown in Figure 6, the accuracy increased as the occlusion area of the building decreased, reaching 0.71 for images with minimal occlusion. Other shooting factors and environmental variables, such as distance, proximity, and exposure intensity, did not significantly affect the accuracy of the building recognition, apart from the occlusion area.

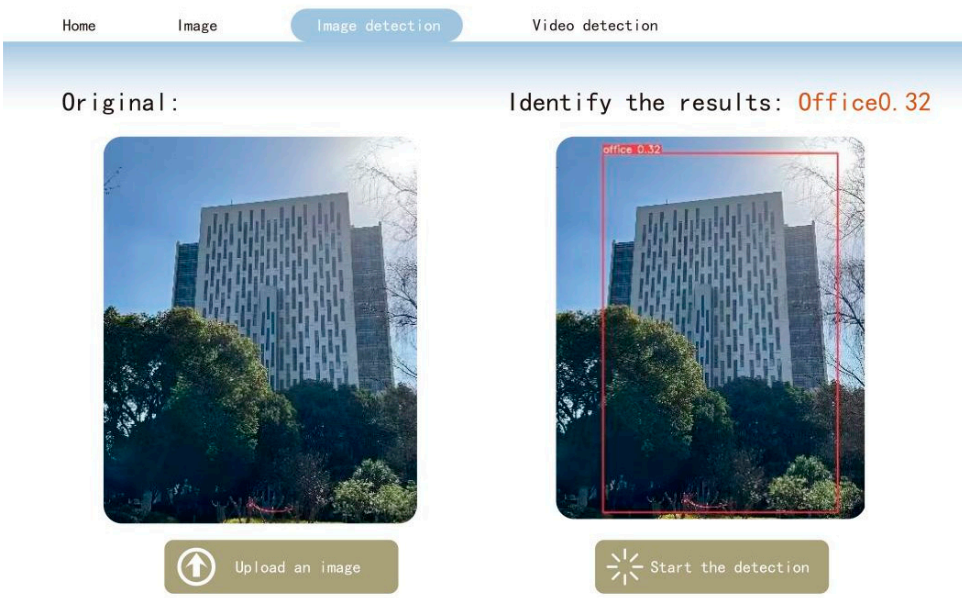

Figure 5. Inspection drawing of a covered office building.

Figure 6. Inspection of an unobstructed office building.

4.1.2. Data Analysis

The system primarily relies on YOLO-V5 as its underlying framework, and the binary cross-entropy loss function is utilized during the computation process. Training the model produced promising results, as evidenced by the small difference between the predicted and true values. The training outcomes, as shown in Figure 7, indicate a notable reduction in localization loss, classification loss, and confidence loss, all falling below 0.02. This signifies diminishing errors and accurate target detection. Throughout the training process, the accuracy and recall rates trended upward, with a gradual reduction in the disparity between the actual and predicted selections. When the Intersection over Union (IoU) threshold was set to 0.5, the average accuracy (mAP) was able to reach an impressive 95% for the given category.

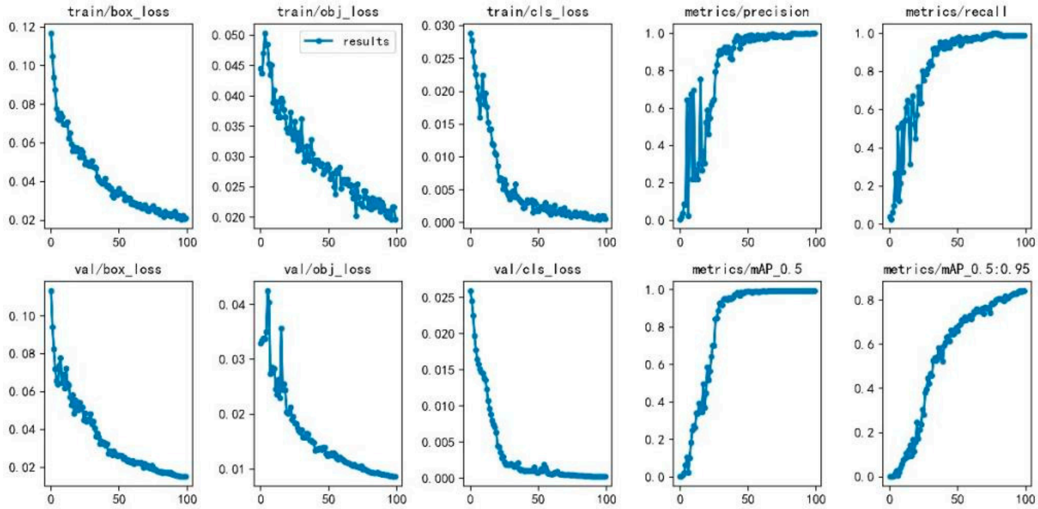

Figure 7. Building detection recognition training results.

4.2. Urban Architecture Color Extraction Methods and Processes

4.2.1. Image Splitting

Following the building recognition process and preceding the building's color extraction, the system incorporates an image-segmenting operation. This involves segmenting the image based on the building's edge frame. The portion of the image contained within the building recognition frame is saved as the building image, while the remaining part is stored separately as the panorama image. Subsequently, in the color extraction process, the colors are extracted independently from the two split images.

4.2.2. Building Color Extraction Based on the K-Values

Considering the specific characteristics of architectural colors, the system sets the default value for the K-means clustering algorithm at 5 in this color extraction. However, in this system, K-values are open values. This provides flexibility to the user by allowing them to manually adjust the K-value during the image recognition process. The K-means clustering algorithm is employed to classify the colors in architectural and panoramic images, which extracts the mean values of the K-class colors for each image. The system interface is depicted in Figure 8.

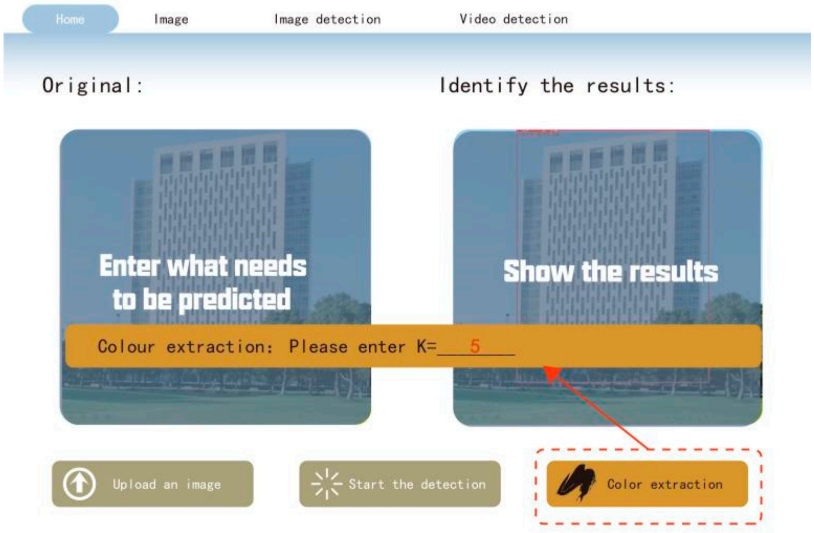

Figure 8. Building color extraction diagram.

4.2.3. Color Identification and Representation

In this test, the colors that occupy the largest proportion of the building's color in the K-category or close to it are extracted. During the color extraction process, the system identifies and displays the building types within the designated display. It then proceeds to identify the colors from the two types of images based on the building and panoramic images segmented using the system. After previewing the image, the system displays the colors present in the image area and calculates the proportion of each color. Figure 9 illustrates the results of the color extraction for an office building with a specified K-value of 5. The preview section clearly distinguishes between the building color and the panorama color. The extracted color bars visually represent the color scheme, with the size of each bar reflecting the proportion of the color. In the text section, the building type, color code, and specific area percentage are provided. For instance, in the case of office buildings, the color area usage of #DDEDF2 amounts to 38.846%.

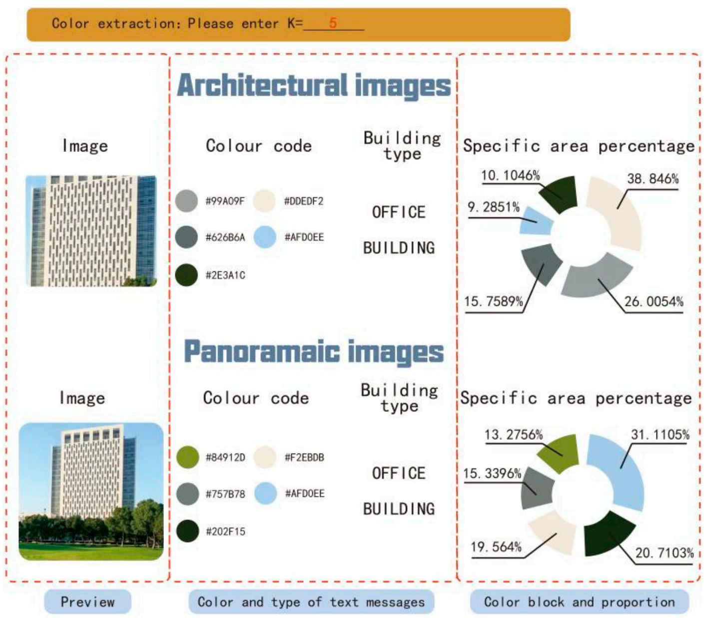

Figure 9. Graph of the training results for building color extraction.

5. Conclusions and Future Work

5.1. Discussion and Conclusions

In recent years, urban buildings have been continuously expanded and constructed, and the identification of various design elements of buildings has also been essential, such as color. The identification of the color of a building requires multiple factors, such as its location and usage area. Combining design factors and objective color values can improve the effectiveness of the recognition results, thus enabling more accurate planning and prediction of urban building colors based on the results.

In earlier research, Gao Zhiheng and his team [9] innovated by harnessing Internet technology, while Young-Ha Shin and colleagues [16] utilized cutting-edge methods like PointNet++ for building extraction in expansive settings. Most research primarily zeroes in on capturing building edge contours, especially for demarcating areas within administrative boundaries. This thesis, however, sought to fine-tune the interplay between buildings and their surroundings by employing environmental imagery to localize structures, subsequently recognizing and analyzing their colors and delivering pertinent results. Building upon this concept, the system presented in this thesis extracts building colors using image-recognition technology. Image-recognition technology is used in the system to differentiate the buildings in the panorama and to capture the colors of the buildings and the panorama image, respectively. Using deep learning on a set of 30–50 specified building images, the system can capture features, demonstrating high accuracy in identifying and classifying buildings. Additionally, it provides an intuitive color-matching effect, identifying the color number and area of use by identifying the building and panorama colors based on the K-values. The combination of convolutional neural networks and K-means clustering algorithms was proven to be feasible and efficient for building recognition and color extraction. The results can also contribute to the extraction and presentation of color schemes in smart city planning and provide valuable insights for the future development of urban colors.

5.2. Limitations and Future Work

There are still areas for improvement in building recognition accuracy algorithms. Specifically, challenges arise when dealing with shadows, as segmenting shaded areas may lead to inaccuracies in color extraction. Moreover, the recognition of special buildings presents difficulties. In this case, while residential buildings can be identified relatively well within the same building type, special library buildings like libraries require separate collection and learning of their unique images and features. Lastly, the building recognition technique in this thesis is tailored for single buildings and has not been trialed on clustered structures. Future algorithm enhancements will focus on addressing these limitations.

Author Contributions: The authors confirm their contribution to the paper as follows: Study conception and design: Y.Z.; data collection: Y.Z.; analysis and interpretation of results: Y.Z. and C.G.; draft manuscript preparation: C.G. All authors have read and agreed to the published version of the manuscript.

Funding: This thesis was supported by the Shanghai Planning Project of Philosophy and Social Science (study on color planning system of new City area in Shanghai, 2021EWY004).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Zou, Z.; Zhang, Q.; Zhai, F. An Urban Residential Buildings Recognition Model based on Building Footprint Geometry Feature Identification. *Archit. Cult.* **2021**, *7*, 225–228.
2. Yan, L.; Tu, H.; Wang, D.; Zhang, S.; Liu, L.; Zhang, Y.; Zhang, Y.; Wang, Y. Exploring City Examination Using Deep Learning-based Building Detection. *Shanghai Urban Plan. Rev.* **2022**, *162*, 39–46.
3. Zhang, H.; Zheng, X.; Zheng, N.; Shi, W. Building Extraction from High Spatial Resolution Imagery based on MAEU-CNN. *J. Geo-Inf. Sci.* **2022**, *24*, 1189–1203.
4. Xia, L.; Zhang, J.; Zhang, X.; Yang, H.; Xu, M. Precise Extraction of Buildings from High-Resolution Remote-Sensing Images Based on Semantic Edges and Segmentation. *Remote Sens.* **2021**, *16*, 3083. [CrossRef]
5. Salma, T.; Balázs, N.; Csaba, B. HierarchyNet: Hierarchical CNN-Based Urban Building Classification. *Remote Sens.* **2020**, *12*, 3794.
6. Shi, Z. Research on the colour of modern residential buildings in the functional core area of the capital based on GIS. *North China Univ. Technol.* **2022**, 121. [CrossRef]
7. Jin, R.; Long, Y. *Architectural Skin Colour Recognition Algorithm Based on Moiré Principle Analysis*; Atlantis Press: Amsterdam, The Netherlands, 2015; pp. 1248–1252.
8. Knox, D.; Xue, B.; Zhang, M. Convolutional Neural Networks with Mask Shift Loss for Improving Building Outlines Detection. *Procedia Comput. Sci.* **2023**, *222*, 636–645. [CrossRef]
9. Cai, F. The theory and practice of smart city planning in the era of big data. *Cheng Shi Jian She Li Lun Yan Jiu* **2023**, *431*, 150–152.
10. Guo, C.; Zang, Y.; Gao, Y. Building Protection Data Release Planning Based on Multifeature Deep Learning. *Sci. Program.* **2022**, *2022*, 2765486. [CrossRef]
11. Huang, Y.; Yu, J.; Lin, X.; Zhong, J.; Li, H. Research on intelligent evaluation of building damage state based on deep learning. *J. Nat. Disasters* **2023**, *32*, 148–158.
12. Gu, H.; Choo, S. Method for Constructing a Façade Dataset through Deep Learning-Based Automatic Image Labeling. *Appl. Sci.* **2022**, *12*, 7570. [CrossRef]
13. Shin, Y.-H.; Son, K.-W.; Lee, D.-C. Semantic Segmentation and Building Extraction from Airborne LiDAR Data with Multiple Return Using PointNet++. *Appl. Sci.* **2022**, *12*, 1975. [CrossRef]
14. Horng, S.-J.; Huang, P.-S. Building Unmanned Store Identification Systems Using YOLOv4 and Siamese Network. *Appl. Sci.* **2022**, *12*, 3826. [CrossRef]
15. Cai, W.; Wen, X.; Tu, Q.; Guo, X. Research on image processing of intelligent building environment based on pattern recognition technology. *Vis. Commun. Image Represent.* **2019**, *61*, 141–148. [CrossRef]
16. Zheng, L.; Ai, P.; Wu, Y. Building Recognition of UAV Remote Sensing Images by Deep Learning. In Proceedings of the IGARSS 2020–2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020; pp. 1185–1188.
17. Gao, Z.; Yuan, S.; Gong, Y. A Study of Internet Map Building Recognition Based on AI Technology. In Proceedings of the 2020 IEEE 6th International Conference on Computer and Communications (ICCC), Chengdu, China, 11–14 December 2020; pp. 2323–2328.
18. AlBdairi, A.J.A.; Xiao, Z.; Alkhayyat, A.; Humaidi, A.J.; Fadhel, M.A.; Taher, B.H.; Alzubaidi, L.; Santamaría, J.; Al-Shamma, O. Face Recognition Based on Deep Learning and FPGA for Ethnicity Identification. *Appl. Sci.* **2022**, *12*, 2605. [CrossRef]
19. Buongiorno, D.; Prunella, M.; Grossi, S.; Hussain, S.M.; Rennola, A.; Longo, N.; Di Stefano, G.; Bevilacqua, V.; Brunetti, A. Inline Defective Laser Weld Identification by Processing Thermal Image Sequences with Machine and Deep Learning Techniques. *Appl. Sci.* **2022**, *12*, 6455. [CrossRef]
20. Wei, Z. *Research on Detection of Violent Abnormal Behavior Based on Deep Learning*; Harbin University of Science and Technology: Harbin, China, 2022; p. 000258.
21. Wu, Y.; Gao, L.; Zeng, Z. Research on Architecture Image Retrieval Based on Convolutional Neural Network and the Attention Mechanism. *J. Wuyi Univ. (Nat. Sci. Ed.)* **2023**, *37*, 57–65.
22. Santos, D.; Silva, F.; Pereira, D.; Almeida, L.; Artero, A.; Piteri, M.; de Albuquerque, V. Real-Time Traffic Sign Detection and Recognition using CNN. *IEEE Lat. Am. Trans.* **2020**, *18*, 522–529. [CrossRef]
23. Wu, B.; Xiao, L. Research on Crop Intelligent Image Recognition and Classification Based on Convolution Neural Network. *Agric. Mech. Res.* **2023**, *45*, 20–23+29.
24. Teoh, T.T.; Rong, Z. Convolutional Neural Networks. In *Artificial Intelligence with Python. Machine Learning: Foundations, Methodologies and Applications*; Springer: Singapore, 2022; pp. 261–275.
25. Gao, J. *Deep Learning: Convolutional Neural Network Technology and Practice*; China Machine Press: Beijing, China, 2020.
26. Lang, X.; Zhao, Z.; Xiong, G. The Analysis of Traffic Drivers' Behavior based on Kmeans. In Proceedings of the 2017 International Conference on Computer Science and Artificial Intelligence (CSAI '17). Association for Computing Machinery, New York, NY, USA, 5–7 December 2017; pp. 232–236.
27. Tu, T.N.; Van, D.D.; Hoang, H.N.; Van, T.V. A Method to Enhance the Remote Sensing Images Based on the Local Approach Using KMeans Algorithm. In *Advances in Information and Communication Technology*; Springer: Cham, Switzerland, 2017; Volume 538, pp. 41–52.

28. Guo, C.; Wang, X.; Shi, C.; Jin, H. Corn Leaf Image Segmentation Based on lmproved Kmeans Algorithm. *J. North Univ. China (Nat. Sci. Ed.)* **2021**, *42*, 524–529.
29. El rube', I. Image Color Reduction Using Progressive Histogram Quantization and Kmeans Clustering. In Proceedings of the 2019 International Conference on Mechatronics, Remote Sensing, Information Systems and Industrial Information Technologies (ICMRSISIIT), Ghana, 20–22 December 2019; pp. 1–5.
30. Hussain, M. YOLO-v1 to YOLO-v8, the Rise of YOLO and Its Complementary Nature toward Digital Manufacturing and Industrial Defect Detection. *Machines* **2023**, *11*, 677. [CrossRef]
31. Yang, J.; Zhao, C. Survey on K-Means Clustering Algorithm. *Comput. Eng. Appl.* **2019**, *23*, 7–14.