sustainability

Article

Application of Intelligent Technology in Facade Style Recognition of Harbin Modern Architecture

Linlin Shan [1,*] and Long Zhang [2,*]

[1] College of Fine Arts and Design, Tianjin Normal University, Tianjin 300387, China
[2] College of Computer and Information Engineering, Tianjin Normal University, Tianjin 300387, China
* Correspondence: shanlinlin@tjnu.edu.cn (L.S.); zhanglong@tjnu.edu.cn (L.Z.)

Abstract: The judgment of facade styles is an important part of the sustainable utilization and restorative process of historical architectures. Contemporary Harbin needs the help of modern architectural facade forms in the planning of the famous historic city, especially with the facade renovation of old architectures with non-cultural heritage. This paper discusses the possibility of applying advanced image recognition algorithms to the classification of the modern Harbin architectural facade styles and argues that the keys to the classification and positioning of the styles are the forms, the details, and the decorative patterns of the architectural facades, together with the deformation and the quantitative variation factors of the facade decoration symbols. Based on the conventional classification method, the facade styles of Harbin modern architecture were divided into 12 categories after data analysis. To better capture the overall structure information and the style features of the local components in the architectural images, the group convolution and the dilated convolution were added into the ResNet model, and then, the improved channel attention mechanism was introduced to construct a novel CA-MSResNet model. The CA-MSResNet model could more accurately identify the morphological elements and the style categories of the architectures, and the average accuracy reached 87.5%. These techniques, with their promising results, are expected to be further applied in the future research on the sustainable utilization and renovation of Harbin modern architecture.

Keywords: facade style; intelligent recognition technology; ResNet; Harbin modern architecture

check for updates

Citation: Shan, L.; Zhang, L. Application of Intelligent Technology in Facade Style Recognition of Harbin Modern Architecture. *Sustainability* **2022**, *14*, 7073. https://doi.org/10.3390/su14127073

Academic Editors: Aliakbar Kamari, Carl Peter Leslie Schultz, Steffen Petersen and Marc A. Rosen

Received: 24 April 2022
Accepted: 3 June 2022
Published: 9 June 2022

1. Introduction

Harbin modern architecture basically refers to the buildings built in Harbin, China, during the historical period of modern social development (1840–1949) [1]. Modern Harbin is in the political situation of forced opening, and the architecture in this period has the typical characteristics of foreign cultural input. It is a special process in which the "integration and transformation" and "inclusion and innovation" of multi-source cultures are intertwined. From the perspective of the evolution of architectural style, "the integration of Chinese and foreign styles" and "the modern transformation of traditional architecture" are the essence of the evolution of the modern architectural style in Harbin. The modern architecture of Harbin is in the transition period of intersection between Chinese and Western and the replacement of old and new architectures; that is, the cultural collision of interwoven Chinese and Western architectures has also experienced the historical connection of modern architecture. The spatiotemporal relationship associated with it is intricate, and most of the modern architecture is still retained today; it has become an important composition of urban architecture today and an important embodiment of the planning and utilization of famous historical cities [2]. In recent years, the protection of the historical architectures in Harbin has been increasing year by year, and more attention has been paid to the sustainable use of the historical architectures and their application value in the restoration and renovation of old architectural facades. In 2020, the National Development and Reform Commission of China and the Ministry of Housing and Urban-Rural Development jointly issued the Notice on Further Strengthening the Management of Urban and Architectural Styles, emphasizing the strengthening of urban and architectural style management and the continuation of the urban context and reflecting the urban spirit. We shall clearly respect historical and cultural relics rather than the dismantlement of historical architectures and traditional dwellings [3]. Harbin modern architecture will play an increasing role in the planning of this famous city. The relevant research on how to make sustainable use of the modern architectural heritage and the restoration and renovation of the old buildings in Harbin has also increased rapidly.

Architectural style is a symbol of the architectural features and the cultural connotations. The morphological characteristics of architectural style are mainly reflected in the facade appearance, the space combination, the color application, the material selection, and the indoor environment of the architecture [4]. Architectural styles vary according to the constraints of politics, society, economy, building materials, and the construction technology of the times, as well as the influence of architectural design ideas, views, and artistic literacy [5]. The modern architecture of Harbin is mainly influenced by exotic factors, which form the opportunity and the impetus for urban construction and architectural modernization and cover almost all the Western architectural styles [1,6]. So, their architectural styles are complex and diverse and mixed with Chinese and Western features. With regard to the sustainable use of the contemporary Harbin urban design and the renovation and restoration of the old architectural facades, it is a complex work to quickly and effectively classify them. It has a great relationship with human subjective understanding and knowledge structure, and it is difficult for common users and the public to distinguish them quickly and accurately. Therefore, we often see that the architectural style characteristics do not match the design content in some projects of modern architectural utilization and the renovation of the old architectural facades, resulting in obstacles that obstruct the communication of professionals and in public misunderstandings. Furthermore, in recent years, Harbin has preliminarily completed the construction of urban expansion, and the restoration and renovation of old architectures has become an important part of the future urban construction. The establishment of style in the restoration and renovation design of the old architectures is crucial. If a variety of style design attempts are made on old architecture, they are used to identify the design effect and the degree of integration with the surrounding environment. When they are selected by policy makers and users, they will achieve better results. At the same time, they will greatly increase the cost of manual design. The rapid preliminary style design for a type of architecture, using computer intelligent recognition technology, can effectively solve this problem; it is one of the key technologies for realizing the automatic classification of architectural facade styles. Therefore, it is necessary to use computer deep learning or other machine learning algorithms to quickly and accurately determine the style of the architectural facade. Considering this, we first carried out image acquisition, data analysis, and the processing of the facade images of the modern architecture in Harbin, followed by the analyzing of the forms, the details, and the decorative patterns of the architectural facades, as well as the deformation and quantitative variation factors of the facade decoration symbols. We further completed the classification and confirmation of the architectural facade styles of Harbin modern architecture; on this basis, the computer intelligent recognition technology was applied in order to realize the automatic identification of the architectural facade styles.

The main contributions of this paper are as follows: (1) the classification criteria of Harbin modern architecture were determined based on the conventional classification method in modern architectural history and data analysis; (2) a platform for the picture library of modern Harbin architectural facade styles was established for the effective management, data analysis, and sustainable utilization of the renovation and restoration; (3) to better capture the overall structure information and the style features of local components in the architectural images, a novel CA-MSResNet model was proposed based on the classical

ResNet model. The CA-MSResNet model could more accurately identify the morphological elements and style categories of the architectural facades and reached a promising level of accuracy.

2. Literature Review

The classification and determination standard of the modern Chinese architectural style has always been the key direction of Chinese architectural history research. Through the search of Chinese literature, we found a total of 183 relevant Chinese articles in CNKI [7], and the research on the determining criteria of architectural style has shown a significant upward trend in the past decade. We used the keywords "Chinese architectural style" to search in WoS (Web of Science) and found a total of 63 relevant English articles. Through reading and analysis, it was found that the relevant literature mainly focused on the aspects of architectural history and heritage protection, architectural form characteristics, and style classification, as well as the classification of decorative patterns. The specific relevant research on architectural style classification was mainly carried out using the following two aspects.

2.1. Historical Dimension Classification Method of Architectural Fields

Xu S.B. believed that the imitation of the Western style by Chinese modern architecture almost covers all the Western architectural trends of thought at that time, of which the architectural styles of Qingdao, Dalian, and Harbin are relatively singular and those of Shanghai, Tianjin, and Hankou appear to be of a mixed situation. Since the late 1920s, with the development and spread of modernist architecture in Europe and the United States, there has also been a trend of transition from new architecture in China to modernist architecture [6]. Feng L. believed that the architectural style of the modern national form is another type of modern architectural style in China; such architectures are generally divided into three types, which are antique style, mixed style, and "modern" style [8]. Lai D.L. studied and analyzed many important cases of modern Chinese architectural style from the perspective of artistic expression and cultural connotation by using historical methods and classified the architectural styles according to the development process of modern Chinese architectural aesthetics [9]. Zhang F.H. believed that modern Chinese architectural style can be divided into four categories: inheritance type, impact type, early onset type, and delayed type [10]. Zheng S.L. collated the styles of modern architecture in Shanghai into four categories: the Western new classical style, the combination of modern high-rise and traditional Chinese style, the modern architectural expressionist style, and the decorative art style [11].

Hou Y.B. believed that the main styles of modern architecture in Harbin include the Russian style, the new art movement, eclecticism, and the Japanese modern style [1]. Liu S.F. introduced various styles of modern architecture in Harbin from the perspective of urban development [12]. Chang H.S. introduced in detail the construction process and the forms of modern architecture in Harbin under the influence of the Russian style [13]. The forms, styles, and construction histories of modern architecture in Harbin are introduced in the book *Harbin Historical Architecture*, compiled by the Harbin Urban and Rural Planning Bureau [14]. The Russian architectural theorist Krakin described the forms of Russian churches and other religious buildings in Harbin in his works [15]. Nishizawa, a Japanese architectural history theorist, divided the modern Chinese architecture in Harbin into two categories: one is "Chinese style" architecture, and the other is "Chinese Baroque" architecture [1].

Through the above literature analysis, it can be seen that the focus of professional scholars in the historical dimension on the classification of modern Chinese architectural styles is on whether the facade forms of architecture are satisfied with a certain style, as well as on the further determination of decorative symbol styles; this is followed by the space, the materials, and the colors of the architecture. This shows that the facades and the detailed decoration of the architecture play a key role in the judgment of the modern Harbin architectural style.

2.2. Rational Technical Dimension of Architectural Style Classification

In recent years, the development of computer deep learning technology has opened up a new era for image processing, image recognition, and classification. Its effectiveness and rapidity are difficult to achieve by traditional methods. Although deep learning technology is rarely applied in architectural design, it has great potential in architectural design. The image classification and the recognition of architectures are increasingly dependent on new technologies. The overall solution can be tested and simulated through powerful computer technology, and this is a working method that can be called "new empiricism". It provides a significant collection of the constraints, innovations, and ideologies of architecture and provides unlimited possibilities for the architects' creations at the level of memory, characteristics, and experience [16,17]. Beyond the idea of architectures as instances, the possibility that architectures can be categorized based on the architectural style is explored in [18], and characteristic features with semantic utility for the different architectural styles from the architectural dataset are mined. These features are of various scales and provide an insight into what makes a particular architectural style category distinct. Zhao P. used a feature-extraction module to achieve architectural style classification. Firstly, Deformable Part-based Models (DPM) are used to remove the elements that are not related to classification and to capture the representative elements of the architectures, and then, these elements are sent to the feature-extraction module. The performance of several classifiers is tested, and the best SVM classifier is selected to output the final accuracy [19]. The architecture classification and recognition based on deep learning technology is an effective research method. This method relies on intelligent photography equipment, mobile phones, mobile robots, and other terminals and provides accurate and reliable recognition results for architecture classification based on the collection of high-quality architectural image data [20]. Llamas evaluated the use of deep learning techniques, especially the convolutional neural networks (CNN), for analyzing cultural heritage images, and it is believed that the application of these technologies can make a significant contribution to the digital literature of cultural heritage [21]. Then, they used CNN to classify some elements of interest in the architectural images with architectural heritage value. They also created a new dataset, which was opened to the public, and achieved promising results in accuracy [22]. In the field of urban transformation, Wei, W.X. believes that artificial intelligence not only enriches the content of architectural design but also makes the learning process of architects on historical architectural styles simpler and more efficient [23]. Xia B. proposed a method for the classification and prediction of residential architectural styles. In this study, some site economic factors, such as construction time, location, building height, plot ratio, greening ratio, and so on, are selected as independent variables, and the neural network model is trained to predict morphological elements and style categories with an average accuracy 77.2%. The research shows that the curved volume, the richness of the tones, and the shape of the roof are the most important morphological variables for distinguishing the style categories, and the building height is the most important economic factor for style positioning [4]. Zhang R. studied the pattern recognition of a certain decorative detail of the building and established a dataset of architectural decoration details. On this basis, an effective visual recognition method based on CSV-Net was proposed to achieve a preferable representation ability for distinguishing different architectural detail patterns in practical scenes [24]. Xu H. used deep learning technology to identify and match the styles of street view building areas and established a mapping relationship between the building area images and the building outlines in order to construct the generation method of a large-scale urban architectural style map in detail [25]. Yi Y.K. adopted a CNN model for classifying house styles in the United States. Although the final prediction accuracy is not high due to the lack of image datasets, the trained model can make reasonable predictions in limited test sets. The results show that the correct definition of style is of great significance in improving the accuracy of recognition [26].

Through the above research, we find that it is possible to realize the rapid and efficient recognition of architectural style using the deep learning technology. Additionally, the accuracy of architectural style classification is highly guaranteed by the premise of a large number of accurate image data.

3. Methodology

Unlike the previous manual classification methods, this study introduced the neural network model to construct an intelligent recognition system of modern Harbin architectural facade styles. First of all, high-quality architectural images were collected through the internet and manual shooting. After the preliminary completion of the images, they were manually sorted and analyzed to establish a platform for the picture library of the modern Harbin architectural facade styles. Then, the selected images were preprocessed, and the facade elements identified for architectural style were symbolized. After that, advanced deep learning algorithms were adopted to carry out the architectural facade style recognition, and the accuracy of the recognition was improved through the training of the image dataset of architectural facades, the architectural details, and the CAD drawings of the architectural decorative patterns. Through the following research methods, we efficiently realized the classification of the modern Harbin architectural facade styles.

3.1. Classification Method of Modern Harbin Architectural Facade Styles

Through the analysis of the relevant literature and field data, we concluded that the modern architectural style of Harbin could be divided into two parts: Western-style architecture and traditional revival architecture. The dominant styles in the Western-style architecture are the Russian style, the new art movement, the eclectic, and the Japanese style, followed by classical, Renaissance, Gothic, Baroque, Byzantine, and Decorative art. With regard to the traditional revival architecture, we focused on Chinese retro doctrines and Chinese eclecticism. The Chinese retro doctrines are represented by the imitation Chinese traditional architecture, such as the Harbin Confucian temples, the Jile Temple, etc. Such architecture strives to maintain the volume and overall contour of Chinese classical architecture and to maintain the tripartite form of steps, roofs, and overall architecture without exceeding the basic form of the classical architecture. Others try to attach large roofs to foreign forms of construction, in the so-called Chinese style [27]. The eclecticism of Harbin modern architecture is mainly based on folk architecture, such as that of dwellings, shops, and stores, and most of them are built by Chinese craftsmen with traditional Chinese architectural processes because they do not understand the form and technology of Western architecture. However, to pursue new styles, Western architectural elements and decorative patterns will be added to the architectural facades to form the so-called Chinese–Western integration architecture, such as the Harbin Daowai modern architectural style street [12].

According to our existing datasets and the characteristics of each style, 12 Harbin modern architectural styles were carefully selected. Focusing on these 12 styles, there is also a major reason that these style features are sufficient to distinguish between the different architectures, but at the same time, they also have similarities. This provides a good example for the prediction models to more accurately identify various styles. As mentioned earlier, in modern Harbin, foreign style architecture generally had an eclectic style, but they also had some collective components themselves and obviously had a certain style as the main tone, such as Gothic, classical, Renaissance, Baroque, etc. The similarities they show are difficult for computer identification, but some differences will be found in each of the architecture details and decorative patterns. For example, there are similarities in the roofs, eaves, doors, and windows between classical architecture, eclectic architecture, and Renaissance architecture, but there are more obvious differences in the balconies and decorative patterns; there are also similarities in the decorative patterns, eaves, doors, and windows between imitation Chinese traditional architecture and Chinese–

Western integration architecture, but there are obvious differences in the cornices and dormers (Table 1). We tested our model to see if it could identify these differences, and it is important to enhance the model to accurately identify these similarities and differences.

Table 1. Analysis of Features of Modern Harbin Architectural Styles.

Style	Roof\Dome	Eave	Window	Porch	Balcony	Column	Cornice	Decorative Pattern	Parapet Wall	Dormer Window
Classical	√		√	√		√				√
Eclectic	√	√	√	√		√		√	√	√
New art movement	√	√	√	√		√				√
Decorative art			√							
Gothic	√		√	√						√
Renaissance	√	√	√	√	√	√		√	√	√
Byzantine	√	√	√	√						
Baroque	√	√	√	√	√	√		√	√	√
Russian Style	√			√						√
Japanese Style										
Imitation Chinese traditional	√		√				√			
Chinese–Western integration		√	√	√	√	√		√	√	

3.2. Manual Processing Method of Image Data

3.2.1. Data Collection and Image Preprocessing

Firstly, the early image classification and identification was carried out on the basis of the extraction of the architectural facade style feature images, which were mainly completed by manual shooting and internet collection. After the completion of the image collection, according to the architectural facade style and the architectural use attributes, a platform for the picture library of the modern Harbin architectural facade styles was constructed to carry out the digital classification of the data collection for the development of the research, in order to improve the work efficiency of the data analysis (Figure 1). About 5000 images of the modern Harbin architectural facades and details were taken. The architectural facades were inevitably shielded by plants, wires, and other architecture. In addition, for the architecture with high height, we failed to find the appropriate shooting point. In the process of the image processing of this architecture, the more complete images of the architectural facades were picked out first, and then, some images were manually processed with Photoshop software so that they would look completely available (Figure 2).

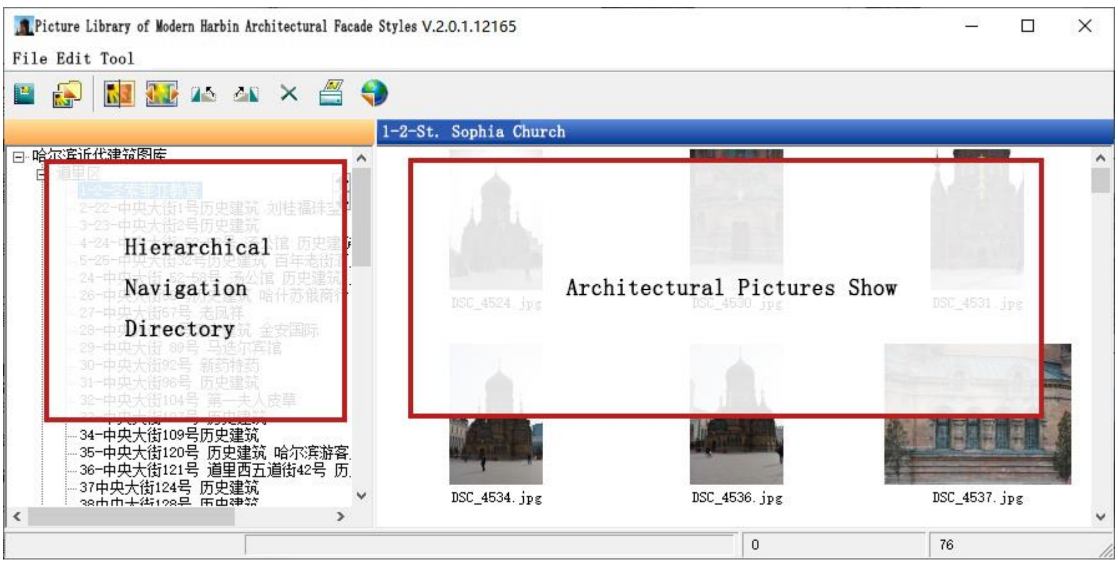


Figure 1. Platform for the picture library of modern Harbin architectural facade styles.

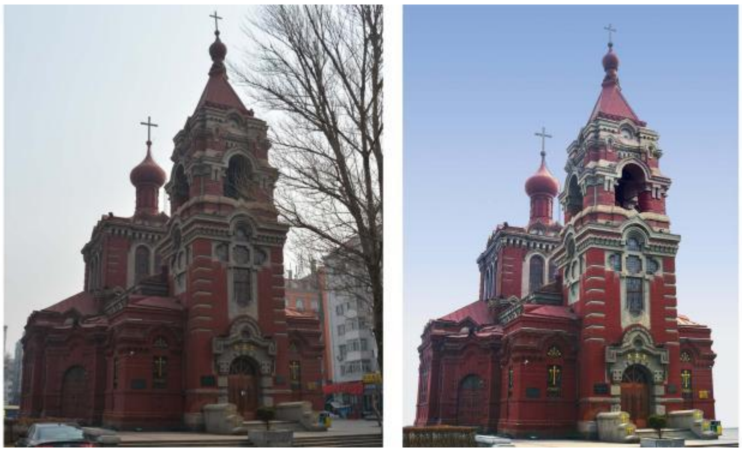


Figure 2. Image preprocessing of St. Alekseyev Church.

3.2.2. Data Analysis and Decorative Pattern Extraction

In these images, the decorative patterns that could distinguish the different historical styles, rather than the standard pattern of a certain architectural style, were mainly extracted. For example, in the facade of the St. Alekseyev church at No. 47, Shike Street, Harbin, the flame pattern was extracted from the decorative pattern of the facade, and the decoration of the main entrance of the church showed a huge and distinct flame symbol. The flame symbol of the main wall of the church was combined with the window and the window frame, which controlled the whole facade. A total of 21 valuable morphological elements were extracted from the collation of the decorative symbols of the church facade to correspond with the image recognition of the Russian style architecture (Figure 3). In the external facade processing of the Harbin Tang Gong Pavilion, the facade restoration was carried out according to the style elements of the new art movement architecture and the historical architectural photos (Table 2). A large number of CAD drawings were drawn for this purpose and were most likely to present the architectural facade and architectural details of the new art movement architecture because there was little modern architecture of this style in China. Moreover, some typical representative architectures were accurately processed in the same way.

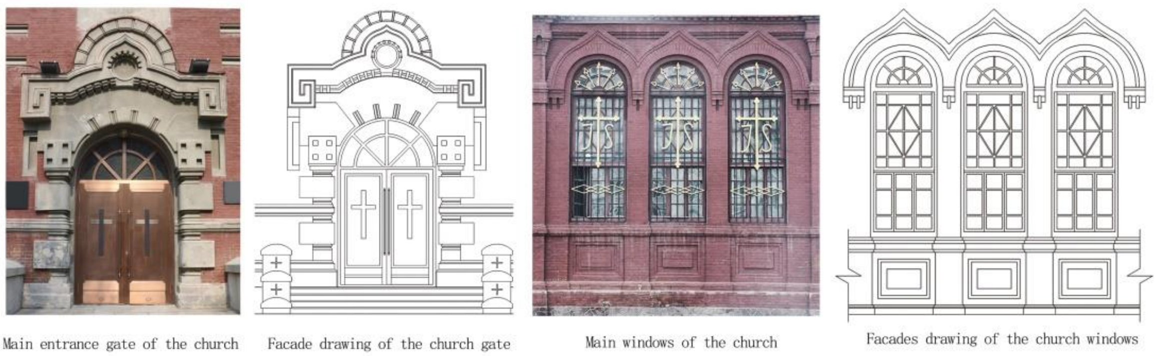
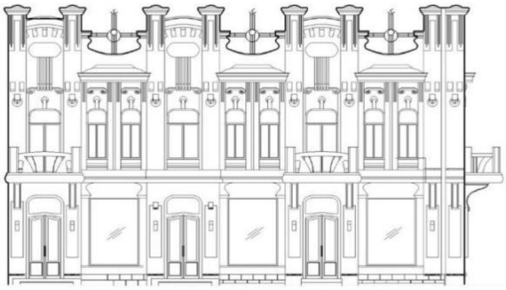


Figure 3. Facade decoration of St. Alekseyev Church.

In the process of artificial data collection, we carried out a large number of the above image data analyses, followed by classification. The pattern arrangement and the restoration of some of the representative external facades and fine decorative patterns were conducted using computer-assisted design software, as well as decorative symbol deformation and variation, in order to facilitate the further accurate identification by the network neural-based computer deep learning algorithm.

Table 2. Design drawings and photos of restoration and protection of Harbin Tang Gong Pavilion.

	Harbin Tang Gong Pavilion		
Restored facade drawing			
	Old photo	Photo before restoration	Photo after restoration
Reference photos			

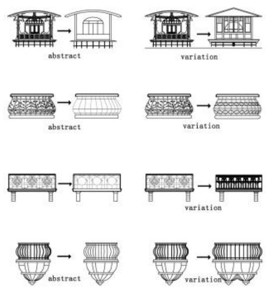
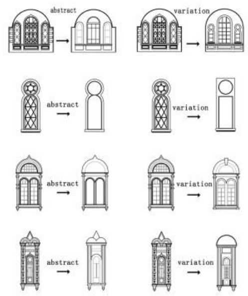
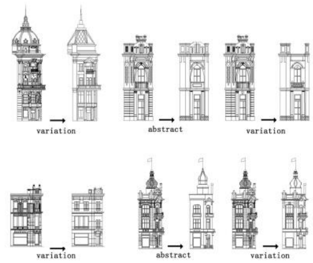
3.2.3. Deformation and Variation of Image Data

Through the analysis of the relevant literature, the judgment of architectural history professionals, and the assistance of historians, we summarized the circumstances in which the deformation and the variations of facades and detailed decorative symbols appeared in the Harbin modern architecture, according to the experience of our team. In the determination of the modern Harbin architectural facade style, we often found that the decorative symbols of the facades in some of the architectures did not possess the typical form of a certain style. However, the simplification of form and the increase or decrease in lines appeared, while the architectural style still belonged to a certain system. For example, the decorative symbols of windows and rails are usually more complex and classical in the architectural dominant position, but the symbols will be deformed in the secondary facade and large-area secondary wall. The way of deformation is generally to simplify the lines and preserve the decorative form. Therefore, in terms of this characteristic, the artificial deformation and variation of the modern Harbin architectural facade styles were conducted. Through the data and the images in the platform for the picture library of the modern Harbin architectural facade styles, 20 of the architecturs were randomly selected as the research objects for manual drawing and labeling. This work was conducted by six trained graduate students majoring in art design. The doors and windows, rails, parapet walls, domes, and facade forms were mainly collated and deformed to remove some specific elements, such as the contemporary renovation and restoration of historical architecture; finally, their deformation and variation forms were summarized. (Table 3)

3.3. Classification Method of Architectural Facade Styles Based on Deep Learning

Deep learning, as the mainstream method for current image classification tasks, has been developed for several years and has produced classic neural network models such as AlexNet [28], VGG [29], GoogLeNet [30], and ResNet [31]. These models mainly use the advantages of CNN to process high-dimensional data without pressure by sharing the convolution kernel, automatically completing feature extraction, learning high-level features layer by layer from low-level features, and ignoring details unrelated to the target [32–36]. Meanwhile, CNN is robust for light, clutter wave, rigid transformation, scale transformation, etc. [37].

Table 3. Deformation and variation forms of architectural decoration symbols.

Deformation of Windows	Deformation of Rails	Deformation of Facades
		

In order to further evaluate the applicability of these deep learning methods in the automatic classification task of the Harbin modern architectural facade styles, we decided to use different types of neural network models for testing and to improve the model according to the characteristics of the classification task in order to find the best solution for this specific classification task. At present, ResNet is the most popular network model, and it had obvious advantages in previous image classification tasks.

3.3.1. Residual Networks

ResNet (Residual Networks) was proposed by H K.M. in 2015 [31]. Before that, people believed that as long as the convolution layer was continuously added, better performance could be obtained, but the reality was not satisfactory. People found that if the model reached more than 30 layers by simply stacking convolution layers, even when the regularization method was used, the accuracy did not increase; rather, it decreased. This was because, with the deepening of the model, the parameters of the back propagation would be infinitely small after multiple propagations, resulting in many parameters almost no longer changing; this meant that the convolution layers in front of the networks were almost no longer learning. It is the so-called degradation problem. Therefore, He K.M. proposed that by using a shortcut the gradient of all the layers could be jumped to the first layer without reduction. The main idea of the shortcut theory was to enlarge the change of feature in each residual block in the form of a residual and to reduce the degradation of the network.

The comparison between the conventional convolution block and the residual block of ResNet is shown in Figure 4. For the conventional convolution block (on the left of Figure 4), we can use a nonlinear function H to describe the input and output of a block; that is, the input is X, and the output is H(x). H usually includes convolution, activation, or other operations. For the residual block of ResNet (on the right of Figure 4), through the shortcut connection, the output of the block is expressed as the linear superposition of a nonlinear transformation of input and input, where F(x) + x, F(x) is the residual. When F(x) = 0, this is an identity mapping transformation. Usually, F(x) is not equal to 0. To compare the left subgraph of Figure 4, F(x) = H(x) − x. Obviously, F(x) is more sensitive to the change of network output than H(x). The use of shortcuts is equivalent to amplifying the features of the changes, which is conducive to the learning of model weights; thereby, the network can be deeper. The proposed shortcut enables the convolution network model to reach thousands of layers, which effectively improves the performance of the model.

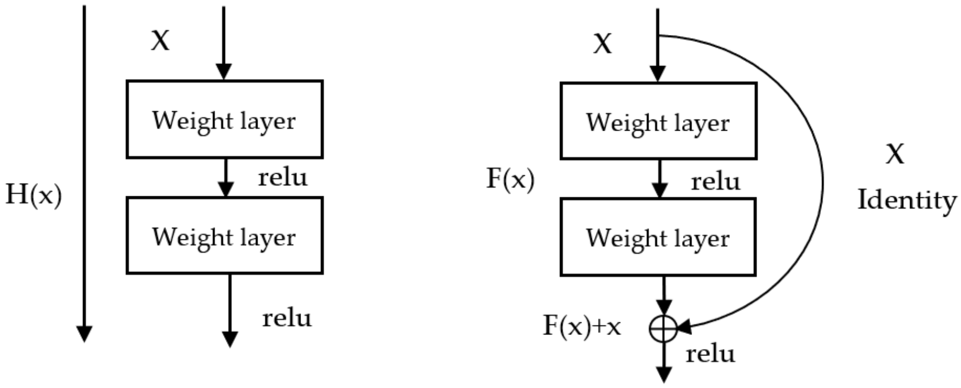


Figure 4. Conventional convolution block (**left**) and residual block of ResNet (**right**).

3.3.2. Multi-Scale Residual Network Based on Channel Attention

Based on the application characteristics of the image classification task of the modern Harbin architectural facade styles, architectural style should not only pay attention to the overall external morphology of the architecture, but also to distinguishing the characteristics of the architectural detail elements. However, for the overall extraction of the external morphology of the image and the accurate extraction of the discriminative regional characteristics of the image, the conventional ResNet model is still insufficient in feature extraction. In order to better capture the overall structural information of the architectural image and the local style features of architectural components, this paper proposes a multi-scale residual network based on the channel attention (CA-MSResNet) model, as shown in Figure 5.

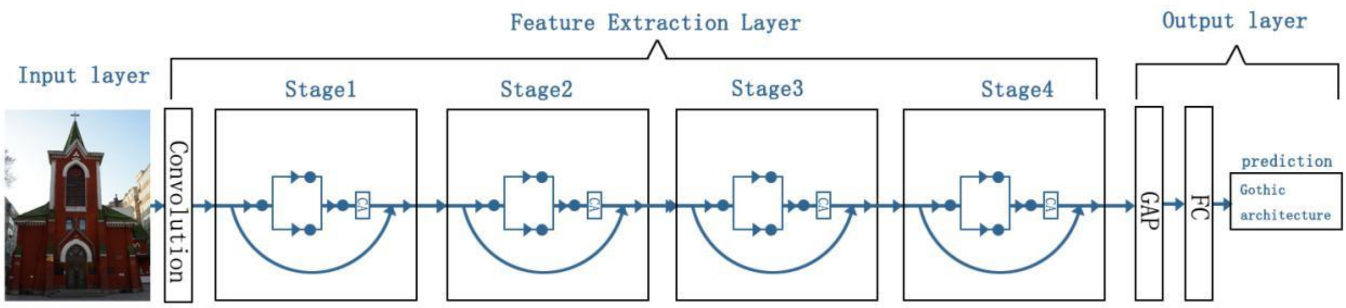


Figure 5. Structure diagram of the CA-MSResNet model.

The main network of the CA-MSResNet model is the same as that of the ResNet model, which is composed of an input layer, a feature extraction layer, and an output layer. As the bottom layer of the CA-MSResNet model, the input layer directly reads the images and preprocesses them. The feature extraction layer first uses a convolution layer to initialize the input images. After that, the model is divided into four stages, and each stage contains multiple identical residual blocks. In order to capture more abundant overall structural information, each residual block uses both group convolution and dilated convolution for feature extraction at the same time. In order to accurately extract discriminative regional features in the image, the channel attention (CA) module is used after each residual block. In the first residual block of Stage 2, Stage 3, and Stage 4, the down-sampling with a stride of 2 is used to obtain the feature map with half the length and width of the original for further extraction of the higher-level semantic information. The output layer is mainly composed of the global average pooling, the fully connected layer, and the Softmax function. The global average pooling (GAP) and the fully connected (FC) layer quantize the feature map output by the feature extraction layer into a one-dimensional vector. Finally, the Softmax function is used to obtain the prediction probability of the different categories in the one-dimensional vector, and the maximum value is taken as the prediction classification of the model.

Dilated convolution is widely used in the field of image segmentation and can expand the receptive field without changing the amount of calculation [38]. As shown in Figure 6, the left figure is the ordinary convolution, and the right figure is the dilated convolution with a dilated rate of 2. As they are both 3 × 3 convolutions, the amount of calculation is the same. However, it can be seen from the right figure that the receptive field of dilated convolution with the dilated rate of 2 is equivalent to the ordinary convolution of 7 × 7, which is more conducive to obtaining the overall features of the image. Although the 3 × 3 convolution kernel can extract effective feature information while maintaining a small amount of computation, the ability of the single-scale convolution kernel to extract features is limited. When the size of the image is relatively large, with the deepening of the number of network layers, a larger receptive field is needed to obtain the effective features of the image.

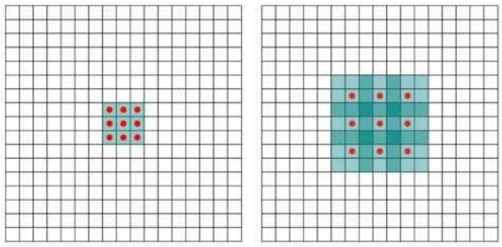


Figure 6. Ordinary convolution (**left**) and dilated convolution with dilated rate of 2 (**right**).

In order to obtain the multi-scale features of the target image, the CA-MSResNet model uses a residual block with group convolution and dilated convolution to enhance the extraction of context information features. Each residual block is shown in Figure 7, where the left is the ordinary residual block, and the right is the residual block of the CA-MSResNet model. Firstly, 1 × 1 convolution is performed to obtain the new channel dimension, and then, two operations are performed on these channels; the first is to use the group convolution to increase the diagonal correlation between the convolution kernels; the second is to use the dilated convolution to provide different scale features for the network to obtain a larger receptive field. Then, a 1 × 1 convolution is used for dimension reduction, and the outputs of the two are connected to obtain the enhanced context information feature; finally, the dimension is reduced by a 1 × 1 convolution layer. The residual block of the CA-MSResNet model uses group convolution to reduce the risk of over fitting, which greatly reduces the number of parameters, and uses dilated convolution to expand the receptive field and obtain multi-scale fusion features and richer context information. In Figure 7, the digits 16, 64, and 256 represent the number of channels, g represents the number of group convolutions, and r represents the dilated rate of dilated convolution.

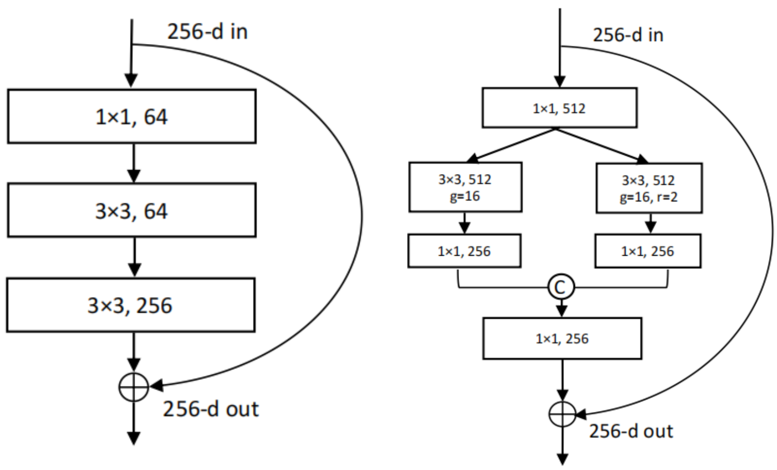


Figure 7. Ordinary residual block (**left**) and residual block of CA-MSResNet (**right**).

In order to gain an accurate extraction of the discriminative regional features of the target image, channel attention is added to each residual block of the CA-MSResNet model. After the channel attention, the feature maps can focus on different feature information, emphasize key local features, suppress unnecessary features, and greatly enhance the expression ability of the features. We improved the classic SE (squeeze-and-excitation) module [39,40] and proposed the CA (channel attention) module. Unlike the SE module which only uses the global average pooling, the CA module performs the global average pooling and the global maximum pooling for the feature map $F \in R^{C \times H \times W}$ at the same time. The value of the global average pooling represents the perceptual domain of this channel to a certain extent, and the value of the global maximum pooling represents the position where this channel has the greatest impact as a supplement to the feature information of this channel. Then, the correlation between each channel is constructed through the fully connected layer, and a group of weights with the same number of channels are the output. The value of this group of weights represents the importance of each channel. Finally, this group of weights is normalized and multiplied with the feature map F. It is shown in Formula (1).

$$F' = \left(\sigma\left(W_1\left(W_0\left(F_{avg}\right)\right) + W_1\left(W_0\left(F_{max}\right)\right)\right)\right) \otimes F \tag{1}$$

where σ represents the Sigmoid function, F_{avg} represents the global average pooling, F_{max} represents the global maximum pooling, W_1 and W_0 represent the parameters of the two hidden layers, respectively, and F' represents the feature map after CA processing.

4. Experiment and Result Analysis

4.1. Experimental Datasets

In the above practical work, we collected and classified a large number of facade and detail images of the Harbin modern architecture and carried out two-dimensional CAD drawings on some of the important representative facades and details. For the convenience of the research, we archived them into three datasets: MCAP-façade-2210, MCAP-Details-2555, and MCAP-CAD-660. The architectural images in each dataset are divided into 12 categories according to their styles, including Chinese–Western integration architecture, imitation Chinese traditional architecture, Russian style architecture, classical architecture, Gothic architecture, Baroque architecture, eclectic architecture, Byzantine architecture, Renaissance architecture, the new art movement architecture, Japanese style architecture, and decorative art architecture. Table 4 shows some example images of our datasets.

The MCAP-Facade-2210 dataset was randomly divided into two parts according to the ratio of 7:3, namely the MCAP-Facade-Training-1570 dataset and the MCAP-Facade-Test-640 dataset. The MCAP-Facade-Training-1570 was the training dataset, and the MCAP-Facade-Test-640 was the test dataset. We all know that the deep learning model requires a large amount of data for model training; even though many pre-training models can be used now, fewer datasets can be used [41]. In order to improve the accuracy of the architectural facade style classification task, we also used the architectural details dataset, MCAP-Details-2555, and the architectural two-dimensional CAD drawing dataset, MCAP-CAD-660, as enhanced training datasets to obtain the improved models.

4.2. Experimental Environment and Configuration

The programming language and the environment used in the experiment were Python 3.10 and Pytorch 1.0, respectively. The hardware configuration was a server with an E5-2620 processor, 120 G memory, and four GPU (GTX1080 Ti).

The CA-MSResNet model uses the Nesterov's accelerated momentum gradient descent method [42]. The experiments are set with 300 epochs, and the initial learning rate is 0.1. The learning rate is adjusted to one tenth of the original at the 150th and 225th epoch. The batch size for training is set to 16 and the batch size for testing is set to 128.

We trained on the MCAP-Facade-Training-1570 dataset and verified on the MCAP-Facade-Test-640 dataset. Table 5 shows the model structure comparison between the

CA-MSResNet model and the ResNet model. The input image size is 224 × 224 × 3 in the CA-MSResNet model. After a 7 × 7 convolution, it is initialized as 112 × 112 × 64 feature map and then passes through four stages. Each stage contains multiple identical residual blocks. In each block, G represents group convolution and D represents dilated convolution. The number of the group is 16 and the dilated rate is 2. In the first block of Stage 2, Stage 3, and Stage 4, the feature maps of 28 × 28 × 256, 14 × 14 × 512, and 7 × 7 × 1024 are down-sampled witha stride of 2. Then, the feature maps are transformed into 1 × 1 × 2048 one-dimensional vector by using 7 × 7 global average pooling. Finally, the Softmax classifier is used to classify.

Table 4. Image datasets of Harbin modern architecture (part).

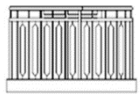
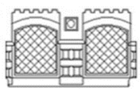
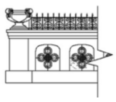
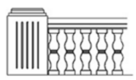
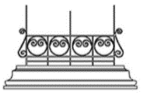
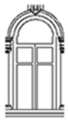
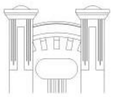
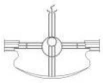
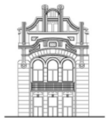
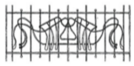
Category		Example			
Chinese–Western integration architecture	Facade				
	Details				
	CAD drawing				
Russian style architecture	Facade				
	Details				
	CAD drawing				
Eclectic architecture	Facade				
	Details				
	CAD drawing				

Table 4. *Cont.*

Category		Example			
Renaissance architecture	Facade				
	Details				
	CAD drawing				
New art movement architecture	Facade				
	Details				
	CAD drawing				

4.3. Experimental Results and Analysis

We constructed, respectively, the AlexNet, VggNet, GoogLeNet, ResNet, and CA-MSResNet to classify the architectural facade styles. According to the configuration in Section 4.2, we trained on the MCAP-Facade-Training-1570 dataset and verified on the MCAP-Facade-Test-640 dataset. The classification results of each model are shown in Table 6.

It can be seen that our CA-MSResNet-M1 model performs better in the actual architectural facade style classification task. The accuracy on the test dataset is 83.6%, which is 4.76% higher than that of the classic ResNet model, and the F1 index is 0.822, which is increased by 5.79% when compared with the ResNet model. It shows that our CA-MSResNet-M1 model can better capture the overall structural information of the architectural image and its detailed style features by using group convolution, dilated convolution, and the improved channel attention mechanism, and then, it can improve the classification performance of the model. At the same time, in order to observe the difference in classification accuracy of each architectural facade style in detail, we provide the error confusion matrix of the CA-MSResNet-M1 model on the test dataset, as shown in Figure 8.

It can be seen from Figure 8 that the accuracy of each style in the test dataset is different. The highest accuracies are 100% for the Japanese style, 100% for the Baroque, 100% for the Byzantine, 100% for the Chinese and Western integration, and 93.75% for the new art movement. The low accuracies are mainly 40% for decorative art, 61.9% for the eclectic, 80% for the classical, and 80% for the Russian style. The heat map is an effective visualization method for the interpretability of the neural network model. In this paper, the class activation mapping method proposed in [43] is used to generate the heat map. Through the heat map analysis, it was found that the images that were incorrectly classified had similar characteristics in the architectural facades to the images that were correctly classified. For example, in Figure 9, the left image that is correctly classified is eclectic architecture, and its hotspot is concentrated in the eaves, but the right images are incorrectly classified are classical architecture and Renaissance architecture, both of which have similar eaves, which leads to the CA-MSResNet-M1 model identifying them as eclectic architecture. It is difficult to improve the prediction accuracy of the model only trained by the architectural facade images.

Table 5. Network structure comparison of the CA-MSResNet model and the ResNet model.

Layers	Output Size	ResNet (50 Layer)	CA-MSResNet (50 Layer)
Conv 1	112 × 112	7 × 7, 64, stride 2	
		3 × 3 max pool, stride 2	
Stage 1	56 × 56	$\begin{bmatrix} 1\times1,\ 64 \\ 3\times3,\ 64 \\ 1\times1,\ 256 \end{bmatrix}\times 3$	$\begin{bmatrix} \multicolumn{2}{c}{1\times1,\ 128} \\ 3\times3,\ G & 3\times3,\ G,\ D \\ 1\times1 & 1\times1 \\ \multicolumn{2}{c}{1\times1,\ 256} \\ \multicolumn{2}{c}{CA} \end{bmatrix}\times 3$
Stage 2	28 × 28	$\begin{bmatrix} 1\times1,\ 128 \\ 3\times3,\ 128 \\ 1\times1,\ 512 \end{bmatrix}\times 4$	$\begin{bmatrix} \multicolumn{2}{c}{1\times1,\ 256} \\ 3\times3,\ G & 3\times3,\ G,\ D \\ 1\times1 & 1\times1 \\ \multicolumn{2}{c}{1\times1,\ 512} \\ \multicolumn{2}{c}{CA} \end{bmatrix}\times 4$
Stage 3	14 × 14	$\begin{bmatrix} 1\times1,\ 256 \\ 3\times3,\ 256 \\ 1\times1,\ 1024 \end{bmatrix}\times 6$	$\begin{bmatrix} \multicolumn{2}{c}{1\times1,\ 512} \\ 3\times3,\ G & 3\times3,\ G,\ D \\ 1\times1 & 1\times1 \\ \multicolumn{2}{c}{1\times1,\ 1024} \\ \multicolumn{2}{c}{CA} \end{bmatrix}\times 6$
Stage 4	7 × 7	$\begin{bmatrix} 1\times1,\ 512 \\ 3\times3,\ 512 \\ 1\times1,\ 2048 \end{bmatrix}\times 3$	$\begin{bmatrix} \multicolumn{2}{c}{1\times1,\ 1024} \\ 3\times3,\ G & 3\times3,\ G,\ D \\ 1\times1 & 1\times1 \\ \multicolumn{2}{c}{1\times1,\ 2048} \\ \multicolumn{2}{c}{CA} \end{bmatrix}\times 3$
Classification On Layer	1 × 1	7 × 7 global average pool 12-d FC, Softmax	

Table 6. Comparison of Classification results of different models.

Models	Accuracy	F1-Score
AlexNet [28]	0.703	0.655
VGG [29]	0.773	0.728
GoogLeNet [30]	0.786	0.753
ResNet [31]	0.798	0.777
CA-MSResNet-M1 (only trained by architectural facades dataset)	0.836	0.822
CA-MSResNet-M2 (Enhanced training model using architectural details dataset)	0.859	0.862
CA-MSResNet-M3 (Enhanced training model using architectural details dataset and CAD drawing dataset)	0.875	0.902

Therefore, we used the architectural details dataset, MCAP-Details-2555, to retrain the CA-MSResNet-M1 model, and we obtained the CA-MSResNet-M2 model. The classification accuracy of the CA-MSResNet-M2 model is 85.9%, and the F1 index is 0.862. At the same time, in order to observe the difference in the classification accuracy of each architectural facade style in detail, we provide the error confusion matrix of the CA-MSResNet-M2 model on the test dataset, as shown in Figure 10.

Figure 8. Error confusion matrix of CA-MSResNet-M1 model (trained by architectural facades dataset).

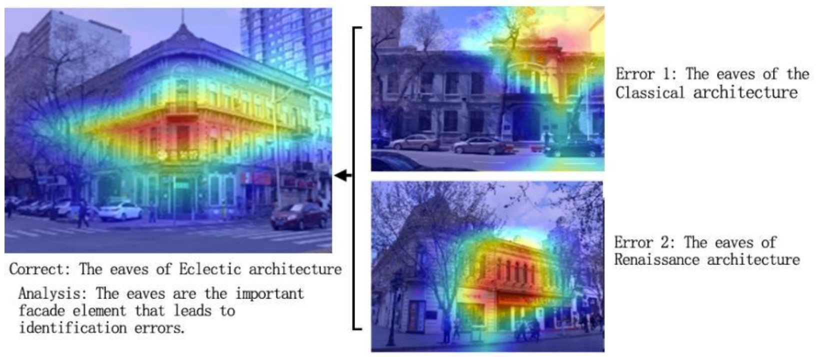


Figure 9. Analysis of heat maps of architectural facades that are incorrectly classified.

Figure 10. Error confusion matrix of CA-MSResNet-M2 model (enhanced training model using architectural details dataset).

As can be seen from Figure 10, in terms of the accuracy, the decorative art has increased to 60%, the classical has increased to 88%, the imitation Chinese traditional has increased to 92.86%, and the eclectic has increased to 80.95%. However, there is also a decline in the classification accuracy of some styles, such as the Russian style from 80% to 60% and the Byzantine from 100% to 75%. We tried to analyze the reasons for the rise through the heat map and found that the CA-MSResNet-M2 model was able to learn the details of the architecture, which helps it to improve the accuracy of its classification of architectural facade styles (Figure 11). The first line of Figure 11 shows that the model correctly learned the parapet wall of Chinese–Western integration architecture, the window frame of classical architecture, and the eaves of eclectic architecture, and the second line shows that the model correctly learned the window of the Renaissance architecture, the entrance of classical architecture, and the eaves of Russian style architecture, which are helpful in improving the accuracy and interpretability of the model in identifying the corresponding architectural facade styles. At the same time, we found that the model has higher accuracy for images with obvious decorative symbols in detail than images without decorative symbols. However, the above two models failed to predict the similarities of some decorative symbols. One of the reasons why neither model has found a suitable decorative style may be the existence of facade decoration symbol deformation and variables, which make it difficult for the models to recognize these features.

In the test, we also found that some errors were due to the inconsistency between the decorative patterns at a specific position of the architectural details and the correct style patterns, and there were deformation and quantitative changes in the pattern. So, we retrained the CA-MSResNet-M2 model using the MCAP-CAD-660 as the training dataset and obtained the CA-MSResNet-M3 model. The classification accuracy of the CA-MSResNet-M3 model is 87.5%, and the F1 index is 0.902. At the same time, in order to observe the difference of classification accuracy of each architectural facade style in detail, we provide the error confusion matrix of the CA-MSResNet-M3 model on the test dataset, as shown in Figure 12. The Byzantine comes back to 100%, the Russian style comes back to 80%, the eclectic rises to 90.48%, and the Gothic rises to 100%. Through the analysis of the heat map, we can see that the CA-MSResNet-M3 model can better learn the key features of the pattern symbols contained in the CAD drawings, such as the decorative part of the Baroque, the flame pattern window frame of the Byzantine, the spire of the Russian style, etc. (Figure 13).

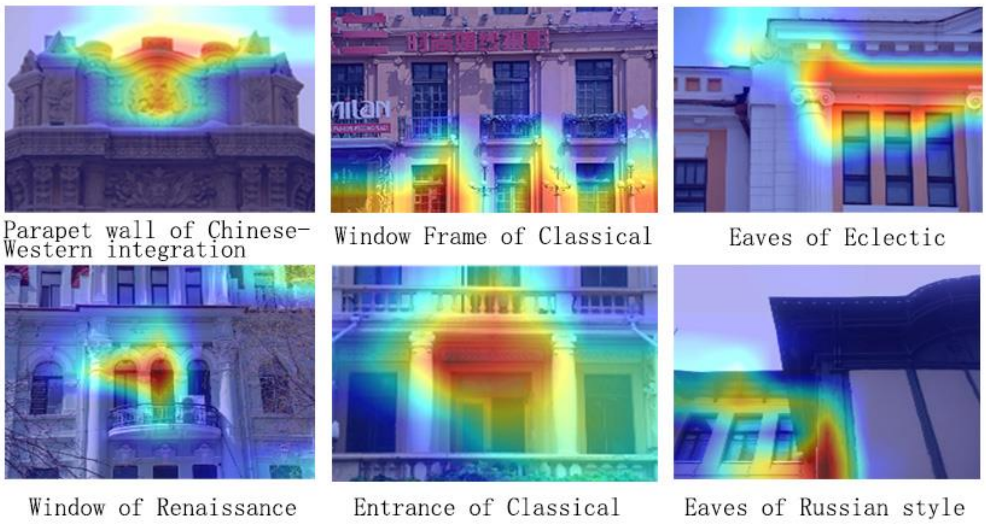


Figure 11. Heat maps of architectural details images.

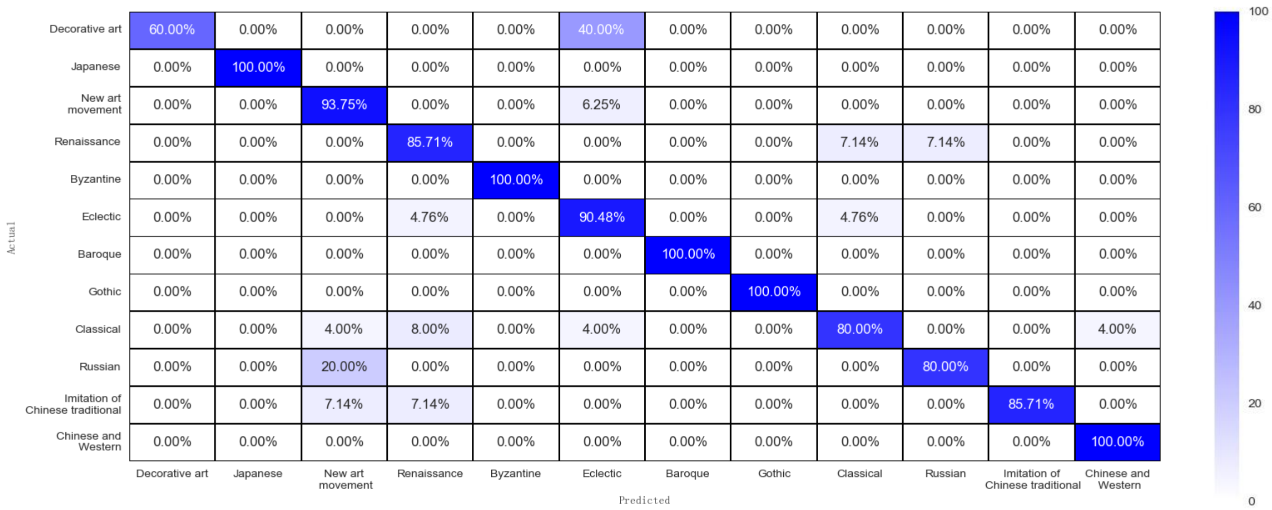


Figure 12. Error confusion matrix of CA-MSResNet-M3 model (enhanced training model using architectural details dataset and CAD drawing dataset).

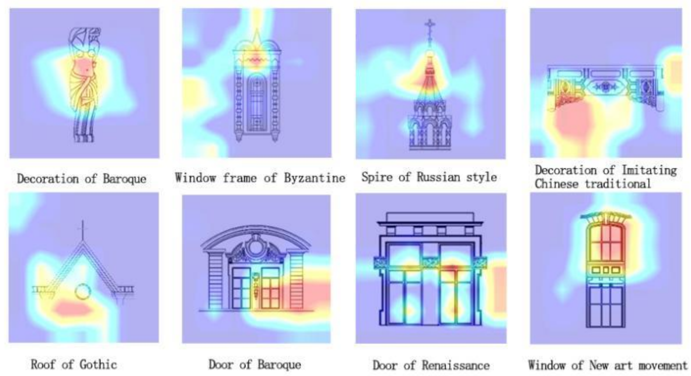


Figure 13. Heat maps of architectural CAD drawings.

5. Conclusions

The classification of architectural facade styles has always been the focus and the difficulty of architectural design. Architectural historians play a key role in the classification of architectural style. At present, the style identification of professional scholars is still the mainstream in the research of architectural design and theory, but it also limits the cognition of architectural style by the public. So, we tried to use computer deep learning and image recognition algorithms to judge the modern Harbin architectural facade style so that the public and architectural designers can quickly and accurately understand the style of the architecture.

Through the study of the definition standard of the architectural style, we found that the facade morphological characteristics, detail structure, and decorative patterns play an important role in the determination of the architectural style. Therefore, we collected and collated a large number of modern Harbin architectural facade images and built a platform for the picture library of the modern Harbin architectural facade styles for effective management, data analysis, and sustainable utilization of renovation and restoration. The facade styles of Harbin modern architecture are divided into 12 categories according to the data analysis and the conventional classification methods in architectural history, including 10 kinds of Western styles and 2 kinds of Chinese traditional revival styles. In order to better capture the overall structure information of the architectural images and the local style features of the architectural components, we constructed the CA-MSResNet-M1 model based on the classical ResNet model by adding group convolution and dilated convolution and introducing the improved channel attention. The accuracy of the MCAP-Facade-Test-640 dataset is 83.6%, which is significantly improved compared with the standard ResNet model. In order to improve the accuracy of the model, we used the architectural details dataset MCAP-Details-2555 as the enhanced training dataset to retrain the model, and the accuracy of the improved CA-MSResNet-M2 model is 85.9%. In order to further improve the accuracy, we used the architectural CAD drawing dataset, MCAP-CAD-660, as the enhanced training dataset to train the model again, and the accuracy of the new improved CA-MSResNet-M3 model was 87.5%. By adding MCAP-Details-2555 and MCAP-CAD-660 to enhance the training of the model, the improved model can learn more features of architectural details and architectural decorative patterns, which can effectively improve the performance of the model and prove the effectiveness of our technical route for the classification and identification of the Harbin modern architectural facade styles. At the same time, we found that the root affecting the accuracy of the identification of the Harbin modern architectural facade styles is that some architectures have some imitation components, and the style characteristics are not prominent. We classified them as eclectic architecture, which reduces the accuracy of recognition. Therefore, the research on the correction of the architectures with imitation components that are incorrectly classified is the basis for the future improvement of this study.

To sum up, the goal of this study is to provide a useful tool for the public and architectural designers, in order to improve the classification efficiency and accuracy of architectural facade styles, to promote the sustainable utilization of architectural heritage in urban planning, and to assist the style positioning of the facades of the old urban architecture. This study will continue as a part of the intelligent style design for the renovation and restoration of the old urban architecture. The aim of further study is to conduct a variety of style design attempts on old architectures at the same time in the future urban design and to identify and evaluate the design effect and the degree of integration with the surrounding environment in order to provide decision makers and users with better design styles. The use of computer recognition technology will quickly carry out the preliminary style design to reduce the cost of the design.

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

Funding: This research was financially supported by the National Social Science Foundation of China (No. 15BG103) and the Tianjin Natural Science Foundation (No. 20JCZDJC00400).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors acknowledge the support provided by the College of Fine Arts and Design of Tianjin Normal University.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hou, Y.B. *Overview of Modern Chinese Architectural History (Harbin)*; China Architecture & Building Press: Beijing, China, 1992.
2. Pan, G.X. *History of Chinese Architecture*, 5th ed.; China Architecture & Building Press: Beijing, China, 2014.
3. Notice on Further Strengthening the Management of Urban and Architectural Styles. Available online: https://zjj.sxxz.gov.cn/xydt/hyxw/202005/t20200514_3522656.html (accessed on 22 February 2022).
4. Xia, B.; Li, X.; Shi, H. Style Classification and Prediction of Residential Buildings Based on Machine Learning. *J. Asian Archit. Build. Eng.* **2020**, *19*, 714–730. [CrossRef]
5. Leland, M.R.; Amanda, C.R. *Understanding Architecture: Its Elements, History, and Meaning*; Routledge: London, UK, 2013.

6. Xu, S.B. *Modern Chinese Architectural History—Westernization and Modernization of Chinese Cities and Buildings*; China Architecture & Building Press: Beijing, China, 2016; Volume 1.
7. China National Knowledge Internet (CNKI). Available online: https://www.cnki.net/ (accessed on 22 February 2022).
8. Feng, L. *National Forms of Modern Chinese Architecture*; China Architecture & Building Press: Beijing, China, 2021.
9. Lai, D.L. *Study on Chinese Modern Architectural History*; China Architecture & Building Press: Beijing, China, 2007.
10. Zhang, F.H. Modern Architectural History Research in China and Modern Architectural Heritage Protection. *J. HIT Soc. Sci. Ed.* **2008**, *10*, 12–25.
11. Jin, L. The Protection and Inheritance of Architecture in New China Needs Masterpieces—Thoughts on reading Zheng Shiling' Modern Architectural Style of Shanghai. *Constr. Archit.* **2020**, *19*, 62–64.
12. Liu, S.F. *Modern Transformation and Mode of Urban Architecture in Harbin*; China Architecture & Building Press: Beijing, China, 2003.
13. Krajin, H.N. *Harbin—An Ideal City in the Minds of Russian People*; Harbin Publishing House: Harbin, China, 2007.
14. Chang, H.S. *Harbin Architectural Art*; Heilongjiang Science and Technology Press: Harbin, China, 1990.
15. Nie, Y.L. *Harbin Historic Building (I)*; Harbin Urban and Rural Planning Bureau: Harbin, China, 2005.
16. Owen, H. *Architectural Styles: A Visual Guide*; Beijing Arts and Photography Publishing House: Beijing, China, 2017.
17. Shalunts, G.; Haxhimusa, Y.; Sablatni, R. Architectural Style Classifification of Building Facade Windows. In *Advances in Visual Computing 6939*; Springer: Las Vegas, NV, USA, 2011; pp. 280–289.
18. Goel, A.; Juneja, M.; Jawahar, C.V. Are buildings only instances? Exploration in architectural style categories. In Proceedings of the Eighth Indian Conference on Computer Vision, Graphics and Image Processing, Mumbai, India, 16–19 December 2012.
19. Zhao, P.; Miao, Q.; Song, J. Architectural style classification based on feature extraction module. *IEEE Access* **2018**, *6*, 52598–52606. [CrossRef]
20. Mathias, M.; Martinovic, A.; Weissenberg, J.; Haegler, S.; Van Gool, L. Automatic Architectural Style Recognition. In Proceedings of the 4th ISPRS International Workshop 3D-ARCH 2011, Trento, Italy, 2–4 March 2011; Volume XXXVIII-5/W16, pp. 171–176.
21. Llamas, J.; Lerones, P.; Zalama, E.; Gómez García-Bermejo, J. Applying Deep Learning Techniques to Cultural Heritage Images within the INCEPTION Project. In *EuroMed 2016: Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection. Part I, Nicosia, Cyprus, 31 October–5 November 2016*; Springer: Cham, Switzerland, 2016; Volume 10059, pp. 25–32.
22. Llamas, J.; Lerones, P.M.; Medina, R.; Zalama, E.; Gómez-García-Bermejo, J. Classification of architectural heritage images using deep learning techniques. *Appl. Sci.* **2017**, *7*, 992. [CrossRef]
23. Wei, W.X.; Wang, N. The Application Status and Prospect of Artificial Intelligence in the Management of Smart Scenic Spots. *J. Hebei Youth Adm. Cadres Coll.* **2019**, *31*, 54–59.
24. Zhang, R.; Zhao, Y.; Kong, J. Intelligent Recognition Method of Decorative Openwork Windows with Sustainable Application for Suzhou Traditional Private Gardens in China. *Sustainability* **2021**, *13*, 8439. [CrossRef]
25. Xu, H.; Wang, L.B.; Fang, Z.X. Street-Facing Architectural Image Mapping and Architectural Style Map Generation Method Using Street View Images. *Geomat. Inf. Sci. Wuhan Univ.* **2021**, *46*, 659–671.
26. Yi, Y.K.; Zhang, Y.; Myung, J. House Style Recognition Using Deep Convolutional Neural Network. *Autom. Constr.* **2020**, *118*, 103307. [CrossRef]
27. Yang, S.L. Retrospection of Chinese Modern Architecture. *Archit. J.* **1987**, *3*, 59–63.
28. Krizhevsky, A.; Sutskever, I.; Hinton, G. ImageNet Classification with Deep Convolutional Neural Networks Alex. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 1097–1105.
29. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-scale Image Recognition. *arXiv* **2014**, arXiv:1409.1556.
30. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going Deeper with Convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.
31. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.
32. Liu, L.; Wang, H.; Wu, C. A Machine Learning Method for the Large-scale Evaluation of Urban Visual Environment. *arXiv* **2016**, arXiv:1608.03396.
33. Sa, I.; Ge, Z.; Dayoub, F.; Upcroft, B.; Perez, T.; McCool, C. DeepFruits: A Fruit Detection System Using Deep Neural Networks. *Sensors* **2016**, *16*, 1222. [CrossRef] [PubMed]
34. Rawat, W.; Wang, Z. Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review. *Neural Comput.* **2017**, *29*, 2352–2449. [CrossRef] [PubMed]
35. Chu, W.-T.; Tsai, M.-H. Visual Pattern Discovery for Architecture Image Classification and Product Image Search. In Proceedings of the 2nd ACM International Conference on Multimedia Retrieval, Hong Kong, China, 5–8 June 2012.
36. Zhang, L.; Song, M.; Liu, X.; Sun, L.; Chen, C.; Bu, J. Recognizing Architecture Styles by Hierarchical Sparse Coding of Blocklets. *Inf. Sci.* **2014**, *254*, 141–154. [CrossRef]
37. Xu, Z.; Tao, D.; Zhang, Y.; Wu, J.; Tsoi, A.C. Architectural Style Classification Using Multinomial Latent Logistic Regression. In *Computer Vision—ECCV 2014*; Springer: Cham, Switzerland, 2014; Volume 8689, pp. 600–615.
38. Yu, F.; Koltun, V. Multi-scale Context Aggregation by Dilated Convolutions. *arXiv* **2015**, arXiv:1511.07122.
39. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 7132–7141.

40. Park, Y.J.; Tuxworth, G.; Zhou, J. Insect Classification Using Squeeze-and-excitation and Attention Modules-a Benchmark Study. In Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 22–29 September 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 3437–3441.
41. Bengio, Y. Deep Learning of Representations for Unsupervised and Transfer Learning. In Proceedings of the ICML Workshop on Unsupervised and Transfer Learning, Bellevue, WA, USA, 2 July 2011; Volume 27, pp. 17–36.
42. Ninomiya, H. Neural Network Training Based on Quasi-Newton Method Using Nesterov's Accelerated Gradient. In Proceedings of the IEEE Region 10 Conference, Singapore, 22–25 November 2016; pp. 51–54.
43. Zhou, B.L.; Khosla, A.; Lapedriza, A.; Oliva, A.; Torralba, A. Learning deep features for discriminative localization. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, Nevada, USA, 26–30 June 2016; pp. 2921–2929.