*Hydraulic and Civil Engineering Technology VII*
*M. Yang et al. (Eds.)*


*doi:10.3233/ATDE220940*

# The Classification Method of Urban Architectural Styles Based on Deep Learning and Street View Imagery

Haozun SUN [a], Hong XU [a,1] and Quanfeng WEI [a]
[a] *School of Urban Construction, Wuhan University of Science and Technology, Wuhan 430065, China*

**Abstract.** The task of identifying urban architectural styles occupies a very necessary position in the fields of construction of smart cities, sustainable urban development and community regeneration. The research method proposed in this paper can improve on the inconveniences of traditional methods of identifying urban architectural styles, such as: the community building is relatively old, and the integration of more periods of architectural style can significantly affect the test results. It is an established fact that data cannot be collected and processed efficiently by humans alone, and can not enter such qualitative and descriptive research methods into the computer for auxiliary research. This paper is based on the explosion of information data use in the 21st century, and use deep learning technology to process unstructured data with convolutional neural networks as the core to assist in the identification of urban architectural styles. With the rapid development of deep learning technology in recent years, its classification techniques for identification of street images of urban buildings can be used for urban management, and a new strong underpinning for the allocation of urban resources, urban diversification management, and the transformation of old communities in the later period has been provided by the proper classification of urban architectural styles. Notwithstanding its restrictions, the approach presented in this research has shown promise and the valuable value of deep learning-based techniques for the study of architectural styles, and this approach has universal significance.

**Keywords.** Style recognition, urban planning, deep learning

## 1. Introduction

Artificial intelligence has its origins in the 17th century or earlier, when the concept of artificial intelligence was based on reasoning. What does "artificial intelligence" mean? The detection method of artificial intelligence was defined in the paper "Computing and Intelligence" published by computer science and cryptography pioneer Alan Mathison Turing. Let normal people and machines with "human intelligence" communicate with humans normally, during which they can ask any questions. After the test, if they cannot distinguish between humans and machines, then the tested machine is called "artificial intelligence". Behind the development of artificial intelligence is the continuous updating of human's summary and analysis methods of

[1] Hong Xu, Corresponding author, School of Urban Construction, Wuhan University of Science and Technology, Wuhan 430065, China; E-mail: xuhong@wust.edu.cn.

past experiences. Before the advent of machines, people's judgments of things were one-sided and narrow-minded. The human brain's ability to collect, store, and interpret information restricted this type of external cognition. Compared with the human brain, machines are limited by certain conditions. It can continuously collect information, process information and calculate a new model. The most significant thing is the ability to make inferences about the development of things accurately [1]. People process and share the information collected by machines, forming our current Internet.

Through the exploration and discovery of the operation mode and working mechanism of artificial intelligence [2], the core competitiveness of artificial intelligence lies in that it can liberate human beings, realize the independent collection and analysis of the development of things, and summarize its development laws. Going back to the bottom, the core of determining whether a machine is intelligent lies in the machine learning algorithm. Machine learning is classified into shallow and deep learning as a field of artificial intelligence. The conventional machine mode of operation is that the staff manually input a series of instructions to the machine, which is what we call code today, and finally the machine outputs it. As a result, these results are what one would have expected in the first place. Through technical innovations, machine learning changes the original working mode. Machine learning (Machine-learning algorithms) can automatically process relevant data and generate models in self-learning mode by receiving, analyzing, and processing large amounts of data. Analyze and combine historical experience to predict the development trend of things. This process of generating the models used for prediction is very similar to the process of human learning new knowledge, but human learning efficiency is not competitive with computers.

## 2. Related Work

In recent years, the application of artificial neural network in the determination of architectural style has also achieved certain success [3]. The processing of unstructured data information (unstructured data refers to text, images, voice or video clips, etc.) is the specialty of deep learning technology. Convolutional neural networks are used in the majority of contemporary artificial intelligence technologies used in the field of architectural planning. Over the last several years, Street View Imagery (SVI) has acquired a lot of popularity in urban studies. Street View imagery has rapidly risen to become an essential data source for geospatial data collection and urban analysis, gaining insights and supporting informed decision-making. The advantage of street view images is that they have universality, street view images are the use of car-mounted driving recorders and pedestrians to take pictures recorded from the pedestrian's point of view, resulting in a wide coverage area of street view images, vehicle traffic and pedestrians can reach the place can develop street view images, through this way can obtain a large number of data sets. And there are many existing companies that provide free data acquisition services, such as Tencent, Baidu, etc. These free data have a good resolution and can be screened for scientific research. When provided to each study, the number of images used in the data analysis has been extracted, and indeed a large proportion of studies have analyzed thousands of images. In addition, there are more than 12 studies, processing more than 1 million images [4]. Xu et al. took urban street view images as the research object, combined with the spatial geometry and image characteristics of spherical panoramic images, explored the

matching method of building areas with the same name, established the location mapping relationship between single and double image building areas, and actually generated four urban center areas [5]. A convolutional neural network-based technique is proposed by Yue et al. By upgrading VGG convolutional neural networks and their performance is improved, the recognition accuracy of normal-appearing apples, diseased apples and rotten apples was 99.61%, 98.89% and 99.26% respectively [6].

## 3. Method

### *3.1. Theoretical Foundations of Artificial Neural Networks*

#### *3.1.1. Related Technologies*

Deep learning has the most powerful visual recognition accuracy and robustness in the era, and its powerful performance is not only reflected in static pictures, but also has better recognition capabilities for dynamic pictures, videos and real-time monitoring As shown in figure 1). And the rise of technologies such as "end-to-end deep learning" has largely compensated for some of the shortcomings of convolutional neural networks, such as: (1) The need to collect large, high-resolution, diverse types of data sets. (2) It takes a lot of time to manually filter, label, label, and label the dataset. (3) You can save model training time.

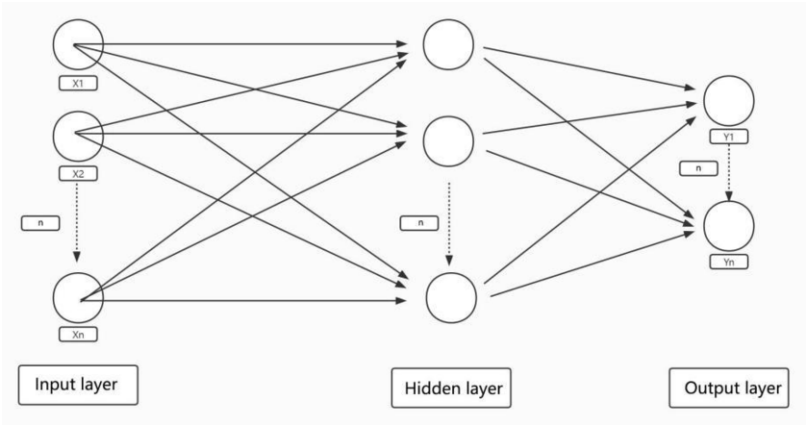


**Figure 1.** Principles of convolutional neural networks.

#### *3.1.2. Image Crawling*

Considering that the training of the neural network requires training the model with a large number of data sets, the acquisition of the data set only relies on human collection or the drone shooting efficiency is low, and the consumption time is too long, so the network picture crawling technology is used to collect the data set (As shown in figure 2). First confirm the destination website of the image and get the URL. Second, change the "user-agent" to get access to the site information. Finally, plug-in technology such as x-path finder is used to complete the crawling of the image.

### *3.2. Classification of Urban Architectural Styles*

The task of identifying urban architectural styles occupies a very necessary position in the fields of construction of smart cities, sustainable urban development and community regeneration. Given that architectural styles evolve and grow steadily through time, with continuity and harmony, architectural styles from comparable times have features of resemblance and correlation, whereas architectural styles from other periods diverge. With a strong color of the times, from imitating and appreciating other decorative styles to the later stage, it gradually forms its own unique and distinctive features. When this feature appears repeatedly in a period of time, it forms its own style. This is a constant process of absorbing, summarizing and summarizing. Taking Wuhan Chuhe Hanjie commercial area as an example, the street buildings are relatively old and incorporate many architectural styles of the period. From another perspective, this forms a great obstacle to urban planning and community image management, the precise positioning of such commercial buildings cannot be given. At present, in most of the urban architectural style identification and classification projects, most of the analysis of architectural styles and building types is based on reviewing the past architectural history and exploring the outstanding characteristics of buildings. Such analysis methods have two shortcomings that are difficult to overcome. (1) The subjectivity is too strong and its accuracy cannot be guaranteed. (2)Under the existing scientific research conditions, such qualitative and descriptive research methods cannot be input into the computer for auxiliary research. For the processing of unstructured data, a deep learning-based convolutional neural network can successfully tackle this problem.

### *3.3. Experimental Steps*

This paper proposes the following ideas for the accurate identification of building styles, and tries to standardize the description and quantitative analysis of urban building types using deep learning technologies based on convolutional neural networks. The main content consists of the following parts:

#### *3.3.1. Conducting Database Creation*

The database should contain a sufficient number of pictures of various architectural styles for training and testing of deep learning models. (The data set can be obtained through network image crawling technology).

#### *3.3.2. Perform Deep Learning Model Building*

Faster region-based CNN (Faster R-CNN) object detection models can extract architectural object regions of various styles [7]. Input the images used for detection in the database into the deep learning model, map the images into 256-dimensional high-dimensional vectors. Finally, using the softmax regression approach, translate these high-dimensional vectors into the likelihood of such architectural styles (As shown in figure 3).

### 3.3.3. Conduct a Man-machine Comparison Test

In order to evaluate the accuracy of convolutional neural network-based deep learning approaches in urban building style categorization tasks to that of humans, researchers used a set of tests, it is necessary to continue to compare the accuracy of human-computer recognition of the same group of urban architectural style pictures in the human-machine comparison experiment. The neural network's identification accuracy was substantially greater than that of the human designer participants, according to the results of the experiment.

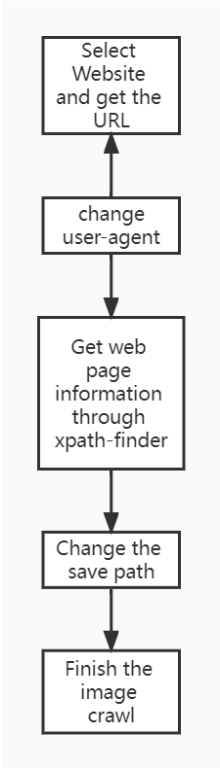


**Figure 2.** Image crawling process.

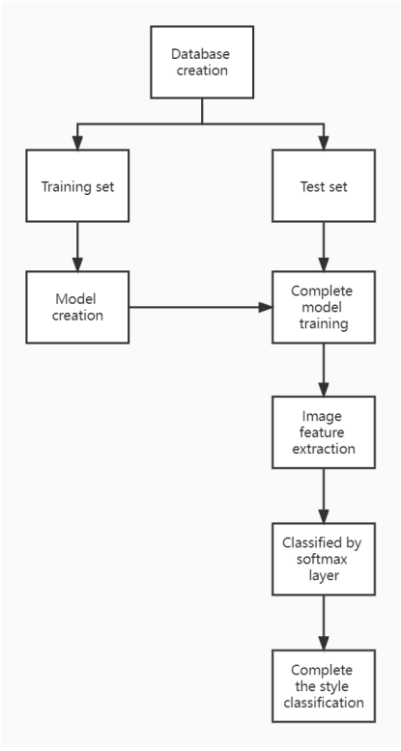


**Figure 3.** Picture style recognition process.

## 3.4. Experiments on Urban Landscape Awareness Based on Deep Learning

### 3.4.1. Data Collection

First of all, a database needs to be established, and a database containing the current status of relevant old community buildings needs to be established initially. The database can be collected and processed on the Internet through the method of web crawling [8], and the real pictures (including unrated ones) of the facades of old community buildings in relevant areas can be collected and processed on the Internet. Search keywords to expand the search scope, respectively locate in Chuhehan Street , Wuhan City and add keywords such as "commercial street", "Commercial pedestrian street", "building facade" (such as searching for "Chuhehan Street", "Pedestrian

Street"), and then automatically crawl the corresponding photos of the building facades in the business district (As shown in figure 4). Secondly, you can also use the image search function of the Baidu app to crawl down the relevant photos posted by individuals on Weibo, Baidu Tieba and other Internet media (such as searching for "Wuchang" and District “Chuhe Han Street Commercial District”).Manual screening is carried out to delete irrelevant and repetitive photos, and the photos crawled on the Internet are divided into two categories. Various sorts of images with important attributes, such as building exterior materials, are utilized as training examples, while the rest photos are used as datasets to be categorized, which are then fed into the model to check its recognition accuracy. In machine learning, there is a rule of thumb that all gathered data is distributed to the training set and the test set in a 7:3 ratio. In the early days of machine learning development, this allocation of data sets was quite reasonable, because the number of previous data sets was much smaller than today, with the development of big data and the Internet as a vehicle for obtaining ever-increasing amounts of data Convenience, so when the database has 1 million training samples, 98% of the data set is used as the training set, the other 1% is used for the development set and 1% is used for the test set to achieve better results.

**Figure 4.** Web image crawl samples.

*3.4.2. Neural Network Model Design and Fine-tuning.*

After the database is established, you need to design an algorithm or use the convolutional neural network model that has been set, and use the convolutional neural network model to map the photo of the building into a 256-dimensional high-dimensional vector, (it is recommended to use ImageNet model to train the crawled pictures [9], so that the size of the pictures in the database are all 224 × 224 pixels) and

then through the fully connected neural network, the 256 high-dimensional vectors converted from photos are converted into architectural styles that are meaningful to humans. Probability, thus completing the recognition of building style by machine learning.

### *3.4.3. Style Recognition*

The training data set produced by the first type of web crawler stated above is used to train the neural network model, and then tested on the second type of test data set. By training the model and fine-tuning the parameters, theoretically the judgment accuracy rate is kept between 80% and 95%, and the fluctuation between them is affected by the number, clarity and accuracy of the training data set. In the man-machine comparison of the three architectural styles of Baroque, Byzantine and SGothic, Sun et al. found that the recognition accuracy of the convolutional neural network has been significantly improved from the inference of the numerical relationship. There is a 9.9% difference in the design subjects in the control group, which is far higher than the average correct rate of all subjects (65.3%), but compared with the best human The average correct rate (90%) of the expert subjects still has a large gap [10], although more human-machine comparison data is needed to verify the correct rate of the convolutional neural network, However, it can be said that the neural network performs better than the average of human subjects in the design profession in the architectural style discrimination task. However, because the model was not properly trained, the neural network identification accuracy was not as good as that of the best-performing design expert subjects.

The experimental results verify the feasibility of deep learning technology for cognition of architectural style. With the results of the experiment and the conclusions drawn, a refined urban architectural style map can be generated more accurately, which proposes a brand-new research for solving problems such as urban planning and design, renovation of old communities, and classification of architectural styles.

## 4. Conclusion and Outlook

In the above cases, the image recognition, modeling training and prediction technologies based on deep learning technology have shown high accuracy, which also indicates that the automatic image recognition model has a high possibility of classifying and evaluating cities. In the future, if further improvements in recognition accuracy are required, researchers will need to tweak the data collection and continue to enhance the neural network model's settings in the following areas. (1) Increase the amount of data; (2) Expand its diversity; (3) Ensure its high resolution. It is believed that with the development of technology and the speed of information dissemination, as well as the emergence of various analytical methods and research tools, there will be an impact on the traditional urban management system, which will be extended to other fields. Residents attach great importance to at the spiritual level today, future study will focus on cross-disciplinary research on the cognitive link between ancient literature and art, music and architecture, and landscape. With the sharp increase in the frequency of data use and the urgent need for data processing in various fields in the 21st century, We as urban governors need to keep up with the times and use technological innovations to solve problems in our cities more efficiently, which will not only raise

the bar for city administration, but will also play a larger role in supporting the betterment of our living environment.

## Acknowledgements

The key projects of philosophy and social sciences of Hubei Provincial Department of Education (Grants#20D025), The National Natural Science Foundation of China (Grant s#41771473).

## References

[1] Jiang JH. Exploring the development trend of machine learning based on the era of big data. Electronic Components and Information Technology. 2021; 12: 32-33
[2] Jiao LC, Yang SY, Liu F, Wang SG and Feng ZX. 2016 Seventy years of neural network: review and prospect. Chinese Journal of Computers. 2016; 39(08):1697-1716
[3] Zhang MD, Ju W and Liu XD. 2019 The application of artificial neural network in the judgment of architectural style-taking modern historical buildings in Dalian as an example. Huazhong Architecture. 2019; 37: 43-46
[4] Filip B and Koichi I. Street view imagery in urban analytics and GIS: A review. Landscape and Urban Planning. 2021: 215
[5] Xu H, Wang LB, Fang ZX, He MH, Hou XX, Zuo L, Guan YL, Xiong C, Gong YY, Pang QL, Zhang H, Sun ST and Nadige A. Street-facing architectural style mapping and map generation method under streetscape image. Journal of Wuhan University (Information Science Edition). 2021; 46(05): 659-671
[6] Yue YJ, Tian BK, Wang HJ and Zhao H. Application of improved VGG model in apple appearance classification. Science, Technology and Engineering. 2020; 20(19): 7787-7792
[7] Wang JJ. The principle of deep neural network for architects. Architecture and Culture. 2022; (12): 186-187
[8] Meng BC. 2022 Application of python web crawler. Radio and Television Information. 2022; 29(03): 108-110
[9] Qin QH, Xiao DW, Luo MN and Tao J. Research on traditional village landscape classification based on convolutional neural network. Urban Planning. 2020; 44(07): 52-58
[10] Sun CY and Hu W. A preliminary study on the research method of architectural style and form characteristics from the perspective of machine learning. Urban Architecture. 2021; 18(07): 104-111