Research Article

Data Model of Key Indicators of Urban Architectural Design Based on Long- and Short-Term Convolutional Memory Network

Xiyin Ma,[1] **Jian Li** [ID],[2] **and Xu Zhang**[3]

[1]*Shanghai Technical Institute of Electronics & Information, Shanghai 201411, China*
[2]*Shanghai Institute of Technology, Shanghai 201400, China*
[3]*Southeast University, Jiangsu, Nanjing 210000, China*

Correspondence should be addressed to Jian Li; lijian@sit.edu.cn

The analysis and design of key indicators of urban architecture has always been a frontier subject in the field of urban architecture. Building key indicator network area flow is challenging due to its high degree of nonlinearity and randomness. Based on the extraction method of key indicators of urban architectural design based on long and short-term convolutional memory network, this paper designs a data model of key indicators of urban architectural design. The model can effectively simulate the urban architectural design system to obtain the key index information of urban architectural design of objects step by step. The architecture of cascaded full convolutional long and short-term convolutional memory network is designed and improved. The experiment adopts the block method to ensure the output result with the same resolution as the input urban building image, which solves the problems of the traditional fully convolutional network such as small local receptive field, low output resolution. The simulation results show that the data prediction accuracy of the key indicators proposed in this paper reaches 92.3%, which is 16% and 11.5% higher than the prediction accuracy of the other two depth algorithms (76.3% and 80.8%) respectively, which promotes the key indicators of information flow within the network.

1. Introduction

As one of the most important components of surface information (accounting for more than 80% of urban surface information), buildings have been at the forefront of science and technology such as urban mapping, urban infrastructure design and planning, land use coverage type investigation, and 3D digital city construction. The field has been widely studied and applied [1]. In addition, relying on traditional methods may waste time and money, and these two defects restrict the large-scale application of high-resolution satellite remote sensing images in the contribution of urban construction, and also cause unnecessary waste of resources [2–6]. Therefore, how to use the inherent key indicators of urban architectural design of buildings to obtain accurate information of buildings from satellite remote sensing images quickly. In order to solve the above problems, this paper is oriented to RGB analysis of key indicators of urban architectural design a deep network of structured coding of key indicators of 3D urban architectural design is proposed. The embedded structured learning layer organically combines the conditional random field and the structured coding algorithm of key indicators of 3D urban architectural design. Comprehensively and accurately we learn the object distribution of the key indicators of 3D urban architectural design where objects are located and the positional relationship of the key indicators of 3D urban architectural design between objects; on this basis, the fusion layer of key indicators of urban architectural design of the network skillfully utilizes deep confidence network to realize the fusion of RGB and deep urban architectural image key index information of urban architectural design, so as to fully explore the correlation between urban architectural design information provided by RGB urban architectural image and depth information provided by deep urban architectural image [7–9].

The short-term prediction of urban regional flow is a research content in the field of urban computing, which is of great significance to the planning and management of smart cities. In the actual key indicators of urban architectural design, the urban layout is complex and mixed with various factors such as weather conditions and terrain, so the regional flow is highly nonlinear and random, making the study of this problem very challenging [10–13]. Its motivation is to establish and simulate the neural network of human brain for analysis and learning, which is an intelligent learning method that is close to the human brain. Convolutional long short-term convolutional memory network is a very important network model in long and short-term convolutional memory network, especially in the field of urban building image classification has excellent applications [14–16]. For example, the LeNet-5 network model proposed by Professor Chen et al. [17] was successfully applied to the check recognition of Bank of America, and its recognition rate reached the level of naked eye recognition. Today, with the increasing expansion of urban building image resources, there are various complex and changeable factors, such as severe deformation of urban building images, diverse background changes, and low resolution, which will undoubtedly increase the difficulty of urban building image classification. In order to adapt to this complex and changeable data environment, Zhao et al. [18] established a powerful learning model. The biggest difference between the convolutional network model and other deep models is its two-dimensional network model, which can be convolved with two-dimensional convolutional layers. This particularity is especially suitable for urban building image and speech processing. The method proposed by Li et al. [19] to classify urban architectural images through the learned general urban architectural design dictionary is based on the appearance information of urban architectural images, and does not model the key index relationship layout and object shape of local urban architectural design, so that in the classification of urban building images, the computational cost is reduced, but some discriminative abilities are lost. The hierarchical model based on the shape and appearance of urban building images proposed by Song et al. [20] mainly uses the key indicators of local urban architectural design of the urban building image to form a fixed number of urban building image blocks. The hierarchical model based on the appearance and location of urban architectural images proposed by Ai et al. [21], by modeling the location of the key index of local urban architectural design equivalent to the center of the object, this kind of model is often easy to train, but usually needs to find an optimal object in the center of the urban building image when testing. In recent years, in order to improve the classification accuracy of urban architectural images, many researchers and workers have done a lot of articles on mining key indicators of higher-level urban architectural design. The researchers proposed an effective algorithm that can handle the key indicators of urban architectural design [22–24].

Research on automatic extraction of buildings based on cascaded fully convolutional long and short-term convolutional memory networks: training the cascaded fully convolutional long and short-term convolutional memory network proposed in this paper for automatic building extraction, research on long and short-term convolutional memory networks on model accuracy. According to the experimental results on the Massachusetts building dataset, compared with other methods, the overall prediction accuracy of the method proposed in this paper reaches 92.3%, and the prediction accuracy of the other two depth algorithms (76.3%, 80.8%) respectively Increases by 16% and 11.5%; after obtaining the prediction results of the long and short-term convolutional memory network, the fully connected conditional random field algorithm is introduced to postprocess the output prediction result. Application of air city building image classification and multiobject segmentation based on cascaded fully convolutional long and short-term convolutional memory network, introducing atrous convolution to increase the local receptive field of urban building images, key indicators of urban architectural design. The ability to extract the key indicators of architectural design, but the network's ability to learn the structure of key indicators of urban architectural design is weak. The fusion layer of key indicators of urban architectural design of the network skillfully utilizes the depth Belief network and improved conditional random field, this layer can complete deep structured learning based on the comprehensive semantic information of objects and the semantic correlation information between objects generated by the fusion of key indicators of multimodal urban architectural design.

2. Design Indicators of Long- and Short-Term Convolutional Memory Network

2.1. Convolutional Network Activation Function. The activation function of the convolutional network is equivalent to the "axon" of the neurons in the brain, and the calculated information is processed and transmitted to the next computing unit.1) if the activation function becomes a linear function, a two-layer artificial long-term and short-term convolutional memory network is sufficient to simulate almost all mathematical functions. However, when the activation function becomes the identity activation function (that is, $f(x)$), this property cannot be satisfied, and if the multi-layer perceptron uses the identity activation function, then the entire network structure and the single-layer long-term and short-term convolutional memory network becomes an equivalence relationship, and the meaning of multiple layers is lost. (2) the long and short-term convolutional memory network uses a global optimization method based on the gradient descent algorithm.

$$\psi(u)du = \psi'(a)da - \psi'(b)db - \psi'(c)dc. \tag{1}$$

The sigmoid function can force the input value of the long short-term convolutional memory network to "squeeze" to 0 to 1. If the input is a large negative number, the output will be 0; if the input is a large positive number, the output will be 1. In order to enhance the ability of the model, we use the data segmentation mechanism to resegment the dataset, where the training set contains

113,287 urban building images, 5000 urban building images are the validation set, and the final 5000 urban building images for the test set. In order to objectively and quantitatively evaluate the performance of the model after obtaining the test results, we evaluate the model proposed in this paper through five widely used evaluation metrics: BLEU-1, BLEU-4, ROUGEL, METEOR and CIDEr.

$$
\begin{cases}
\psi'(a)\mathrm{d}a - \psi'(b)\mathrm{d}b > 0, \\
\psi'(a)\mathrm{d}a + \psi'(b)\mathrm{d}b < 1.
\end{cases} \tag{2}
$$

Because the functional properties of sigmoid are consistent with the synapses of neuron neurons, and it is easy to obtain derivatives, traditional neural networks take the sigmoid function as the core content of long-term and short-term convolutional memory networks. First, learn a weak prediction model from the initial training set, and then adjust the distribution of data samples according to the prediction performance of the model, so that the subsequent weak prediction model pays more attention to the samples with previous prediction errors or large prediction errors. It can be in the form of giving a larger weight to the prediction error samples, training the model so that it can achieve the optimal prediction result.

$$
\prod \frac{\psi'(a)\mathrm{d}a}{\psi'(b)\mathrm{d}b} - \prod \psi'(a)\mathrm{d}a \times \psi'(b)\mathrm{d}b = 1. \tag{3}
$$

First we set the dictionary size to 10010 words. To test the performance of the model, we evaluate our proposed model on the MSCOCO datasets. The dataset contains 82,783 training and 40,504 validation city building images, and each city building image includes five human-annotated sentences describing city building images. The algorithm regards the decision tree as a base learner, and adopts the idea of model integration. In order to fit the residual of the previous model, a new model is established in the gradient direction of the residual decline each time. Based on the number of learners, the key indicators of urban architectural design of the samples in the data set are input into the model, and the model will generate the first decision regression tree according to the key indicators of urban architectural design and sample labels, and record the predicted value of the first tree and the actual value.

2.2. Convolutional Network Local Perception.

Long short-term memory networks are widely used in language models for urban building image description generation tasks. The LSTM unit corresponding to each time step can generate corresponding words, which are finally combined into a sentence describing the semantic content of urban building images. Since this network has the function of resolving long-term dependencies by storing memory units, it is considered an ideal structure for the task of describing urban architectural design language in urban architectural images. In the constructed language model, the first layer of LSTM is an attention model, and the attention mechanism of urban building images is constructed in this layer. The second layer of LSTM is a language generation model, which is used to generate the language vocabulary vector of the current time step.

The key indicators of external urban building design also include discrete data such as climate type. For discrete data, this paper chooses the one-hot encoding method. The first function of one-hot encoding is to solve the problem that the machine learning model is not easy to handle discrete data, because the model usually determines that the key index values of urban architectural design are continuous and ordered, but the value of discrete data is randomly assigned, each value represents only one category. The second and most critical role is to expand the key indicators of urban architectural design, extending the value of the key indicators of discrete urban architectural design to the key indicators of European-style urban architectural design.

In many cases, it is necessary to calculate the distance or similarity of key indicators of urban architectural design. It is more reasonable to calculate the distance and similarity of key indicators of European-style urban architectural design. The specific operation of one-hot encoding is to expand the dimensions of the key indicators of urban architectural design in Figure 1, and expand to as many dimensions as there are types of key indicators of urban architectural design.

$$
\begin{bmatrix} \psi'(a)\mathrm{d}a & 1 \\ 1 & -\psi'(a)\mathrm{d}a \end{bmatrix} - \begin{bmatrix} \cos a & -\psi(u) \\ \psi(u) & \sin a \end{bmatrix} = 0. \tag{4}
$$

The most commonly used maximum pooling in the pooling operation uses a 2×2 filter to find the maximum value in each area, where the stride $= 2$, and finally extracts it from the key indicators of the original urban architectural design. The main urban architectural design key indicators get the urban architectural design key index matrix on the right. Pooling will compress the input key indicators of urban architectural design. On the one hand, it will make the key indicators of urban architectural design smaller and simplify the network calculation complexity; on the other hand, it will compress the key indicators of urban architectural design and extract the key indicators of urban architectural design. From the prediction result graph, the prediction stability of GRU is also relatively good, and the degree of fitting between the predicted value and the actual value is also relatively high from the early stage of prediction to the later stage of prediction.

2.3. Weight Sharing of Design Indicators.

The design index network parameter settings are shown in the text. The network parameter setting button is located in the lower right corner of the main interface. After clicking, the parameter setting window will pop up. In this window, the user can set the learning rate of the algorithm, the number of calling processes at runtime, the number of input data types, the number of training iterations, and the initial bias term bias, etc. Parameters are set and adjusted. The network parameter setting button is shown in the text. The network parameter setting button is located in the lower right corner of the main interface. In this window, the user can set the learning rate of the algorithm, the number of calling processes at runtime, the number of input data
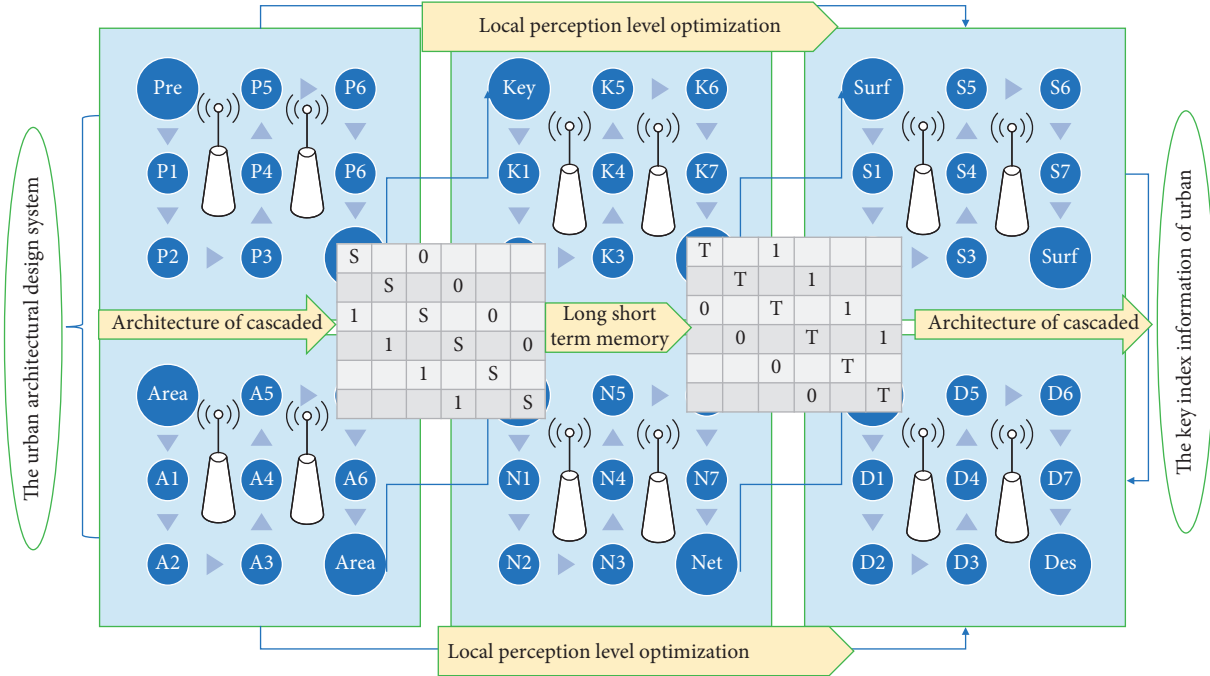
Figure 1: Convolutional network local perception level optimization.

types, the number of training iterations, and the initial bias term bias, etc. Parameters are set and adjusted.

$$\sqrt{\lambda\left(\mu_x + \mu_y + 2\mu_x \times \mu_y\right)} - \sqrt{\lambda\left(\mu_x - \mu_y - 2\mu_x \times \mu_y\right)} > 0. \qquad (5)$$

The function implementation is the focus of this paper. Taking the programming implementation of the automatic extraction program of the convolutional long and short-term convolutional memory network as an example, the network training and testing process is as follows: (1) download the data set, organize the data set, and divide the entire data set into train set (training part), test set (testing part), and validation set (validation part). (2) Creating the layer object of the convolutional long and short-term convolutional memory network, and design the network structure such as network layer type, network layer parameters (convolution kernel size, pooling type, and deconvolution kernel size), number of network layers, weight initialization, and hyperparameters of the training process. (3) By setting the loss function to compare the difference between the final output and the actual output, the back propagation error of the network is calculated.

It can be seen that although the RNN in Table 1 is a classic and mature time series processing algorithm, its accuracy and speed are not ideal. In the 10 validations, the worst MAPE accuracy is 84.77%, the best one result is only 90.26%, and the 10 average MAPE accuracy is 87.52%. In addition, by observing the prediction results and drawing, the fitting degree between the results and the real values in the early stage of prediction is acceptable, but in the later stage of prediction, the gap between the predicted value and the real value is constantly widening, and the deviation is very serious. The above results also support the theoretical defects of RNN, that is, there are serious problems in the processing of large-span time series data. In addition, RNN runs slowly when processing large-scale time series data, and it takes an average of 10.7 seconds to predict 500 data points 10 times, which is difficult to meet the real-time requirements of flight control multi-dimensional time series data prediction tasks.

2.4. Data Subsampling of Key Indicators of Urban Architectural Design. It can be seen from the subsampling results of the key indicators of urban architectural design data that the prediction accuracy of LSTM has been significantly improved compared with RNN. The worst MAPE accuracy in 10 experiments is 89.08%, and the best one MAPE accuracy can reach 94.03%, with an average MAPE accuracy of 90.61%. From the prediction result graph, the performance of the LSTM network is relatively stable no matter in the early stage of prediction or in the later stage of prediction, and its predicted value fits the actual value well from beginning to end. The above results also confirm that the improvement of LSTM compared with RNN is effective, and the introduction of forget gate and memory gate enhances the network's ability to process long-span time series data.

A filter, also known as a convolution kernel, is a neuron with a set of fixed weights, usually a two-dimensional matrix, the size of which can be defined by the user. However, in terms of running speed, the LSTM is not far behind the RNN, with an average time of 10.7 seconds over 10 experiments. This shows that the improvement of the memory gate and forget gate of the LSTM network only improves the prediction accuracy of the network, and does not improve the speed of the network. The LSTM network is still difficult to meet the real-time requirements of aircraft maintenance, that is, flight control multi-dimensional time series data prediction tasks.

TABLE 1: Evaluation criteria analysis network.

Index	Model weight	Layer name	Hard-attention ratio
1	0.116	DropConnect 1	0.884
2	0.119	DropConnect 2	0.915
3	0.122	DropConnect 3	0.946
4	0.125	Dropout 1	0.977
5	0.128	Dropout 2	1.008
6	0.058	Dropout 3	1.039
7	0.061	Dropout 4	0.295
8	0.064	Dropout 5	0.326

3. Construction of Data Model Based on Long- and Short-Term Convolutional Memory Network

3.1. Long Short-Term Convolutional Memory Network Hierarchy. When the input of the long and short-term convolutional memory network is a multi-dimensional urban building image, these advantages are more significant. Therefore, the multi-dimensional urban building image can be directly used as the input data of the long and short-term convolutional memory network, abandoning the traditional extraction algorithm. The selection process of key indicators of urban architectural design and the tedious data reconstruction process; coupled with the subsampling in time and space, and the activation function also brings nonlinear characteristics to the network, and the traditional linear fitting is abandoned in data processing. The nonlinear fitting that is more in line with the actual distribution of the data is adopted, so the convolutional neural network can simulate the real distribution of the data most accurately.

$$\frac{1}{\alpha - r} \times \left[\left(\beta u - \lambda \mu_x - \lambda \mu_y \right) \right] = \begin{cases} \lambda \left(\mu_x + \mu_y \right), & a > 1, \\ \lambda \left(\mu_x - \mu_y \right), & a < 0. \end{cases} \quad (6)$$

The data set contains operational data of 218 parts, which originate from the same type of engine system, such as aircraft engine system. For each part's associated data, the data collection period covers its entire time period after running for multiple cycles until the part fails. The relevant data of each part in each operation cycle includes: part ID code, operation cycle code, 3 kinds of operation setting parameters and 21 kinds of sensor parameters. The total number of data sets is 75738 sets, including 45918 sets of training set data and 29820 sets of test set data. Among them, as shown in the following, for the running cycle, it can be seen that the training set and test set parts run for 357 cycles and 364 cycles at the longest, and after 128 cycles and 15 cycles at the shortest, the parts fail, and the parts run on average respectively 210.63 and 136.79 cycles.

$$\text{if} \{ (x, y) \notin \mathbb{Q} \}, \lim \left[(\beta u + A)^2 - (\beta u - A)^2 \right] - 1 = 0. \quad (7)$$

Using two concatenated 3×3 convolution kernels instead of one 5×5 convolution kernel, the number of parameters is changed from the original 25 (5×5) to 18 ($3 \times 3 \times 2$), and the size of the receptive field is maintained while maintaining the reduced number of parameters.

Similarly, three convolutional layers of size 3×3 can be used in series, and its effect is equivalent to a convolutional layer of size 7×7. In addition, three concatenated 3×3 convolution kernels are more efficient than a 7×3 convolutional layer. The parameters of the $\times 7$ convolution kernel are much less, and the number of parameters of the 3×3 convolution kernel is only 55% of the 7×7 convolution kernel ($3 \times 3 \times 3 / 7 \times 7 = 55\%$).

When there are multiple filters, the stacked structure is the convolution layer. It can be known that the number of parameters in a convolution layer is filter num, where filter is the number of filters (neurons). Most importantly, three 3×3 convolution kernels have more nonlinear transformations than one 7×7 convolution kernel, which makes the convolutional long short-term convolutional memory network more capable of learning key indicators of urban architectural design.

The method in Figure 2 uses the combination of CNN and RNN to mine the correlation of key indicators of urban architectural design in the data, and considers the impact of external events on regional traffic, and realizes the mining of key indicators of external urban architectural design in a fully connected network. The experimental results show that the convolution recursive long short-term convolutional memory network (CRNN) can extract the key indicators of urban architectural design of the data. Compared with the ConvLSTM method, the extraction ability is stronger. This part fully considers various key indicators of urban architectural design that affect regional flow, and still takes the combination of 3, 1, and 1 as the input of key indicators of urban architectural design in time series.

For the selection of surrounding areas of area $3 * 3$, the key indicators of external urban architectural design are treated equally with the key indicators of spatiotemporal urban architectural design, and both are regarded as a common urban architectural design key index of the sample. For the receptive field of neurons, the data stored in the matrix is the coefficient of data processing in the receptive field. The experimental results show that when the key indicators of spatial-temporal urban architectural design and the key indicators of external urban architectural design are taken into account, the prediction effect is improved compared with the previous one. At the same time, the method of gradient boosting tree is more effective than the common single model method.

3.2. Nesting of Key Indicators of Urban Architectural Design. Aiming at the time dependence of urban key indicators of regional flow and the related issues of urban architectural design key indicators, the experiment is based on the time-dependent LSTM network to realize the prediction method of urban regional flow, given different input forms, using (LSTM) temporal dependence of data acquired by long-short-term convolutional memory networks. The experimental results show that the regional flow has the characteristics of long-term time dependence, and it is more reasonable and effective to use the time statistical key indicators of urban architectural design as
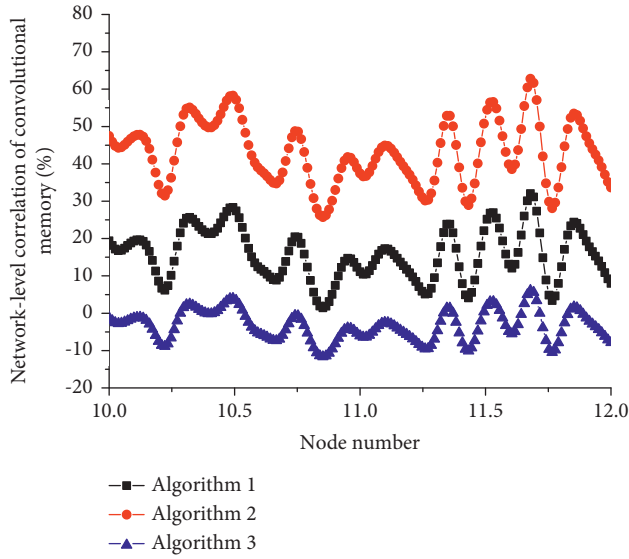
FIGURE 2: Convolutional memory network hierarchical correlation distribution.

FIGURE 3: Distribution of key indicators of urban architectural design.

the input. Then, a method for predicting urban regional flow based on the correlation of key indicators of urban architectural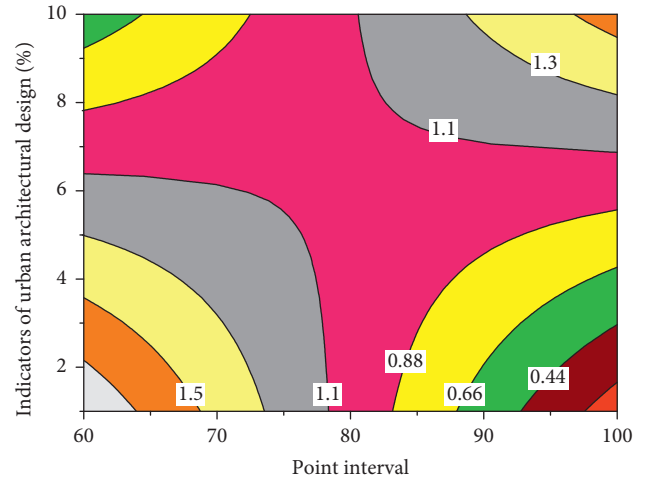 design is proposed. This method uses the time statistics of key indicators of urban architectural design as input, and adds convolution operations on the basis of LSTM network.

$$\int_0^u \exp\left(\frac{\alpha - \beta}{b^2}\right) du - \int_0^u \exp\left(\frac{1}{\pi(\theta)} d\theta\right) du = 0, \quad \text{for} \{a, b, c \subseteq \mathbb{R}(c, t)\}. \tag{8}$$

Different from the traditional training method, the deep belief network has a "pretraining" process before starting training: first, let the weight parameters find a parameter close to the optimal solution, and then train; then fine-tune the entire training process. Using "pretraining" and "fine-tuning" can greatly reduce the time required to train long and short-term convolutional memory networks. However, the input of the deep belief network can only be a one-dimensional vectorized matrix, which does not have much advantages for urban building image processing. The role of the convolution operation is to extract local information from the input data.

When used, filters with different functions can be obtained by changing the weights in the convolution kernel, and different key indicators of urban architectural design can be extracted. Used in the field of urban building images, the convolution kernels of different weight parameters can extract the key indicators of urban architectural design such as color, texture, shape, etc., and integrate these information to identify objects in urban architectural images. In the same way, it is used in the field of time series processing. For the input data of multi-dimensional time series, by adjusting the weight parameters in Figure 3, different convolution kernels can identify different features in the time series data, and these different urban architectural design key. When the indicators are integrated, specific analysis can be made on the aligned time series data.

In the processing of urban architectural images, different weights are assigned to urban architectural images, and the identification of key indicators of urban architectural design in key regions in the image is strengthened. Through continuous training, the neural network can learn the characteristics of each urban architectural image. A large number of studies have shown that attention mechanism can play a role in natural language processing, speech recognition, urban building image processing and other fields. The key indicators of road urban architectural design urban architectural images are generally edited from a continuous video frame. There is a certain time sequence correlation and similarity between urban architectural images, and the objects in each urban architectural image are unevenly distributed.

The research shows that the convolutional long short-term memory network has a good effect in the extraction of time series and key indicators of urban architectural design. Based on the above analysis, this paper introduces the convolutional long short-term memory network to obtain the spatiotemporal urban architectural design key indicators of urban architectural images and fuses the multi-scale urban architectural design key index information, and then combines the attention mechanism to enhance the effective urban architectural design key indicators.

The structure of the internal network in Table 2 is very consistent. Each layer of AlexNet contains only one convolutional layer, and the size of the convolution kernel is

TABLE 2: Clustering of key indicators of urban architectural design.

Number serial	Indicator conv 1	Indicator conv 2	Indicator conv 3	Indicator conv 4	Indicator conv 5
Clustering 1	46.24	11.77	17.42	45.07	−2.36
Clustering 2	52.65	32.08	18.11	45.92	−1.96
Clustering 3	2.26	48.34	18.79	46.73	−1.57
Clustering 4	76.90	39.40	19.31	47.37	−1.26
Clustering 5	10.94	46.61	19.62	47.74	−1.08
Clustering 6	48.93	44.54	19.65	47.77	−1.07
Clustering 7	42.91	28.23	19.25	47.30	−1.30
Clustering 8	62.10	10.82	18.27	46.11	−1.87

7×7; while each layer of VGGNet contains multiple (2~4) convolutional layers, the size of the convolution kernel is 3×3, and the size of the convolution kernel is 3×3 in each layer. VGGNet does not use local response normalization (LRN), because local response normalization cannot improve the performance of the ImageNet Large-scale Urban Building Design. The key indicators of design are related to the key indicators of urban architectural design. On the other hand, the coordinate positions of the superpixels are used to encode the positional relationship of the key indicators of urban architectural design of adjacent superpixel pairs, and the key indicators of urban architectural design are generated. The key indicators of urban architectural design are related to the key indicators of urban architectural design. The distance decay rate ad = 0.5 is set on the PASCAL VOC 2012 dataset, and $d = 0.65$ on the SIFT FLOW dataset.

3.3. Data Strategy for Training Long- and Short-Term Convolutional Memory Networks.

In the process of data strategy

practice of convolutional memory network, the graph structure of CRFs is firstly constructed based on the RGB urban building images segmented by superpixels; then, this paper defines the unary term of CRFs as the key indicators of urban architectural design based on superpixel multidimensional urban architectural design. The softmax classification probability, and the binary term of CRFs is defined as the similarity of adjacent superpixels to the key indicators of LAB color urban architectural design; finally, this paper adopts the cyclic belief propagation algorithm and L. On this basis, this paper uses the SSEA algorithm to learn the structural information of key indicators of urban architectural design of superpixels. On the one hand, SSEA infers the object distribution of key indicators of urban architectural design where superpixels are located according to the CRFs classification probability of superpixels, and generates node urban buildings.

$$\frac{\overbrace{\lambda(\mu_x+\mu_y)}}{(\beta u - \lambda\mu_x), (\beta u - \lambda\mu_y), \ldots, (\beta u - \lambda\mu_x - \lambda\mu_y)} \xrightarrow{i,j>0} \sum \lambda(\mu_x + \mu_y). \tag{9}$$

In the processing of urban architectural images, different weights are assigned to urban architectural images, and the identification of key indicators of urban architectural design in key regions in the image is strengthened. Through continuous training, the neural network can learn the characteristics of each urban architectural image. There is a certain time sequence correlation and similarity between urban architectural images, and the objects in each urban architectural image are unevenly distributed. The research shows that the convolutional long short-term memory network has a good effect in the extraction of time series and key indicators of urban architectural design. Based on the above analysis, this paper introduces the convolutional long short-term memory network to obtain the spatiotemporal urban architectural design key indicators of urban architectural images and fuses the multi-scale urban architectural design key index information and then combines the attention mechanism to enhance the effective urban architectural design key indicators.

Figure 4 contains road scenes from 50 different cities, consisting of a total of 25,000 urban architectural images,

including 20,000 weakly annotated frames and 5,000 high-quality pixel-level annotated urban architectural images. 5000 manually labeled urban building images were randomly divided into three groups, of which 2975 urban building images were used for training, 500 for validation, and 1525 for testing. The dataset has a total of 19 categories, 8 of which are instance-level segmentation. The PASCAL VOC2012 dataset is a comprehensive dataset that can be used for tasks such as classification, semantic segmentation, object detection, and behavior recognition.

$$\frac{\exp\left((\alpha-\beta)/b^2\right)}{\exp\left((\alpha+\beta)/b^2\right)} - \frac{\log\left((\alpha-\beta)/b^2\right)}{\log\left((\alpha+\beta)/b^2\right)} = 0, \quad \text{if} \begin{cases} \alpha + \beta > 0, \\ \alpha > 1. \end{cases} \tag{10}$$

The dataset is used for semantic segmentation of 2,913 urban building images, including 1,464 in the training set and 1,464 in the validation set. 1449 images, in the competition of this dataset, the number of test sets varies. The dataset contains 21 categories, one of which is the background category (except for the specified marked category,
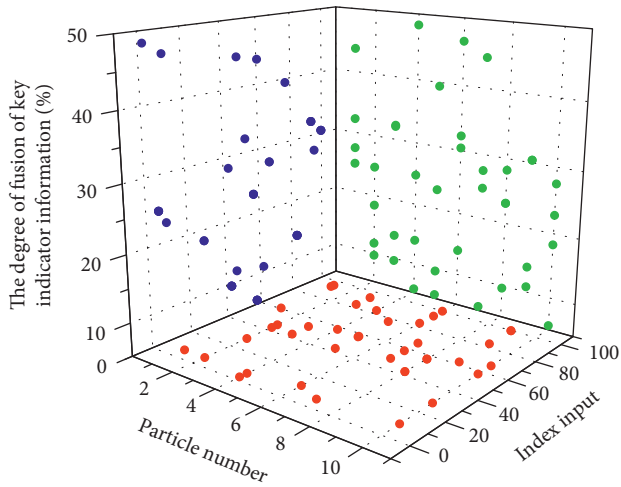
Figure 4: Fusion distribution of design key indicator information.

and the resolution of each urban building image is different. The PASCAL Context dataset and PASCAL Part dataset are extensions of the PASCAL VOC2010 dataset. The former contains 540 categories, of which only 59 common categories are meaningful; the latter subdivides objects on the basis of the original 20 categories, it is divided into front wheel, rear wheel, seat, and handle. These two datasets are more difficult to segment than the original VOC dataset.

4. Application and Analysis of Data Model Based on Long- and Short-Term Convolutional Memory Network

4.1. Data Selection of Long- and Short-Term Convolutional Memory Networks. In the local connection of urban building images, each neuron corresponds to 25 parameters, and there are a total of 100 neurons in the network. If the 25 parameters of these 100 neurons are equal, then the total number of parameters of the network is 25. These 25 parameters are the way that the long-term and short-term convolutional memory network extracts the key indicators of urban architectural design during the training process, and this method is independent of the location. Some key indicators of statistical urban architectural design become the same. The experiment proves that when the residual block exceeds one layer, it will have a promotion effect; otherwise, it will degenerate into linearity. In addition to solving the problem of network degradation, compared with the general network layer, the residual network has many branches to directly connect the input to the following network layer, so that the latter network layer can perform residual learning, which protects the information to a certain extent. Completeness, the entire network only needs to learn the difference between the input and output, which simplifies the learning objective and reduces the learning difficulty.

Since the CNNs in Figure 5 lack the ability to effectively obtain the dependencies between data and their receptive field is small, the key indicators of urban architectural design extracted by CNNs generally lack the global contextual information of objects. To make up for the lack of structured learning ability of CNNs for key indicators of urban architectural design, conditional random fields (CRFs) are added to the backend of CNNs to globally optimize the classification probability. The methods based on CRFs generally optimize the energy function defined on the graph model globally, so that the adjacent objects with similar key indicators (color, texture, and depth) of urban architectural design are of the same category, and the categories with large differences are different, so as to optimize CNNs. The classification probability of urban architectural design realizes the consistency and smoothness optimization of the analysis of key indicators of urban architectural design.

$$\psi(u) = \frac{\int_0^u \exp[f(a) - f(b)]\mathrm{d}u + \int_0^u \exp[g(a) - g(b)]]\mathrm{d}u + \int_0^u \exp[w(a) - w(b)]]\mathrm{d}u}{\int_0^\infty \exp[\pi(u)]\mathrm{d}u}. \tag{11}$$

The difference between the network output and the real result (i.e., the label) is calculated by comparing the loss function, and partial derivatives are obtained for the weights of the network. After training and iteration, the minimum gradient is found. The difference is that the above-mentioned unique structure of the LSTM network enables it to memorize the key indicators of urban architectural design of useful long-term data and forget the key indicators of urban architectural design of useless data, so as to achieve a better time series data processing effect than the RNN network. However, CRFs cannot explicitly infer the object distribution of the key indicators of urban architectural design where objects are located and the dependence of key indicators of urban architectural design between objects, and still lack a strong ability to learn the key indicators of urban architectural design.

4.2. Simulation Realization of Key Indicators of Urban Architectural Design. In order to extract the key indicators of urban architectural design from different urban building images, it is necessary to set filters suitable for different key indicators of urban architectural design, that is, the convolution kernel in the convolutional long and short-term convolutional memory network. Each convolutional layer of the network may have many convolution kernels, such as 100 different convolution kernels, which can learn 100 different key indicators of urban architectural design.
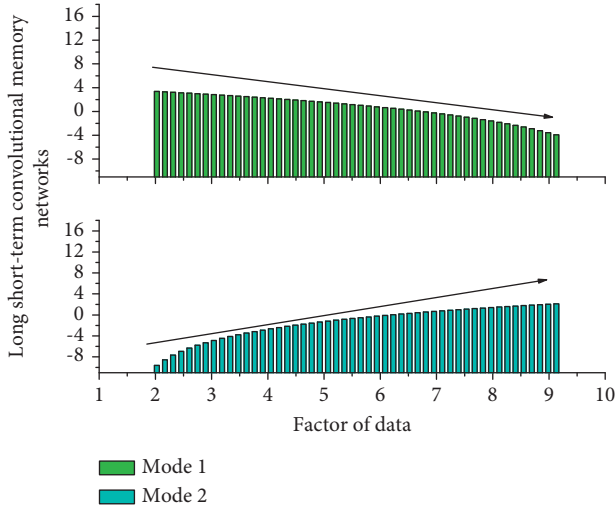
FIGURE 5: Distribution of long and short-term convolutional memory networks.

Before the computational analysis of convolutional layers, several basic concepts need to be introduced first. DCNN uses the method of convolution calculation to extract features from the input data. This paper uses the long and short-term convolutional memory network framework to train a deep convolutional long and short-term convolutional memory network, and completes a complete network training on an NVIDIA GTX 1060 graphics card (4GB video memory) (the number of iterations = 10000) takes 27 hours. Solver is one of the four cores of Caffe (Blob, Layer, Net, Solver), which uniformly distributes the training rhythm of the entire model. Essential to almost all Caffe-based long-short-term convolutional memory networks is the Solver file configuration. There are two basic algorithms in Solver: Forward Compute and Backward Compute. These two calculation methods are alternately performed in the whole network, the purpose is to make the loss function (Loss Function) converge to the global minimum.

$$\int_0^u \frac{1}{\exp\ f(\theta)} \mathrm{d}\theta + \frac{\int_0^u \exp[g(a) - g(b)]\mathrm{d}u}{\int_0^u (1/\exp\ g(\theta))\mathrm{d}\theta} = 1. \tag{12}$$

The calculation process of each Solver is as follows: (1) design a reasonable loss function that can be derived, design a training network for learning parameters and a test network for evaluating the results; (2) through forward computing (forward compute) and Backward compute alternately runs to update the weight parameters; (3) timely and quantitatively analyze the test network; (4) save the model file during network training to display the Solver status. In each network iteration, the process of Solver is as follows: (1) call forward compute to calculate the final output value of the network and the loss value obtained in this iteration; (2) calculate the gradient values of all network layers that need to be derived; (3) select different Slover methods (such as Stochastic gradient descent-stochastic gradient descent) according to actual needs, and use different gradient descent algorithms in the reverse calculation

process to carry out network ownership; (4) save the learning rate used in each iteration, save the model snapshots generated by training, and record the training state of the network at each iteration.

In order to have a more rigorous analysis of the prediction effect of the method in Figure 6, in addition to implementing the regional traffic prediction based on the gradient boosting tree method, several common machine learning methods are also implemented on the same training set and test set to predict regional traffic. The methods of linear regression and simple neural networks, which are widely used in regression-type problems, are included as a contrast to the methods in this chapter. The method in this chapter reconstructs the data set by extracting key indicators of urban building design, and can predict the regional flow of a certain city every time. For a more comprehensive evaluation of the methods in this chapter, it is randomly selected from the city limits are listed in the experimental results table below. Connect one-dimensional TCNs to form a multi-dimensional TCN network, so that the improved network model has the ability to process multi-dimensional time series data, and retains the characteristics of causality and large receptive field of the network model; the activation function of the improved TCN model is replaced by a parameterized ReLU function.

$$\begin{cases} \dfrac{[1 + q(t)]}{F(q,t)} \longrightarrow f[(q,t), (q, t-1)], \\[4mm] \dfrac{F(q,t)}{F(q-1,t)} \longrightarrow 1. \end{cases} \tag{13}$$

The commonly used ReLU function has the ability to derive the input less than 0; the improved TCN model uses a residual connection method with a self-defined span to connect the bottom convolution kernel and the high-level convolution kernel, so that the high-level convolution network can obtain. It retains the details of the underlying network while maintaining a large receptive field; the improved TCN model is a fully convolutional network, that is, the last layer uses the convolutional layer in Table 3 to connect the convolutional layer and the output interface, so that the network can be used for input data of any size can be processed and output in the specified format.

Different from the classification based on the urban architectural design key indicators of the object itself, when two different types of superpixels have similar urban architectural design appearances, SSEDNs can classify objects according to the urban architectural design key indicators of the superpixels. At the same time, when the key indicators of urban architectural design of the two superpixels with the same category are quite different, SSEDNs can also be based on the distribution of the key indicators of urban architectural design of the superpixels. In a word, SSEDNs can not only use the key indicators of visual urban architectural design of superpixels for classification, but also can use the
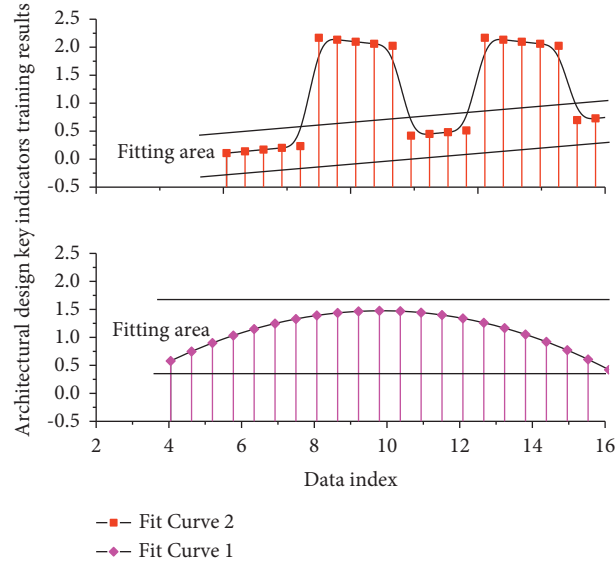
FIGURE 6: Training distribution of key indicators of architectural design.

TABLE 3: Code realization of urban architectural design data.

Number	Architectural design	Code text
1	Large receptive field of $\mu_x\mu_y$	Class javahelloworld {
2	For a comprehensive $\mu_x - \mu_y$	Public static void main(String args[]){
3	Evaluation of the $\alpha + \beta$	System.out.println("hello world\n");
8	Data $\mu_x + \mu_y$ of any size	Import java.util.Scanner;
9	Output in the $\beta u + A$	Public class scannertest {
10	The activation of $t \longrightarrow (1,1)$	Public static void main(String[] args){
11	The layer uses the $\exp[\pi(u)]$	Scanner.scanner = newScanner(System.in);
12	The network $\alpha + \beta > 0$ used	System.out.print("");
13	The network $\sum_{i=1}^{t} Y_i + u$	Int a = scanner.nextint();
14	Connect one-dimensional to it	System.out.printf("%d %d\n",a,a * a);
15	Make the $t + 1\vert t > x$ code	Setlayout(new flowlayout());
16	Specified $\sum mt$ format	Getcontentpane().add(in);
17	The function $f(u + mt)$	Getcontentpane().add(btn);
18	The urban $i = 0, 1, 2, \ldots, n$	Getcontentpane().add(out);
19	Different types of $x, y \notin \mathbb{Q}$	Btn.addactionlistener(new.btnactionadapter());

key indicators of node urban architectural design of superpixels to optimize the classification results of key indicators of urban architectural design.

$$\text{s.t.}\, g(x) \cup g(t) = \varnothing, g(\min.(t-1,x), \min.(x), \min.(t), \min.(x-t)) \quad \underset{t,x \longrightarrow 1,1}{\square} - \frac{\delta[q(t,x-1)]}{\delta[q(t)]} = 0. \tag{14}$$

The process of training or optimizing the long and short-term convolutional memory network is the process of minimizing the loss function. The smaller the loss function value, the closer the value of the corresponding predicted result and the real result is, which reflects the long-term and short-term convolutional memory network. For the convergence characteristics of the loss function, when the value of the loss function is relatively large, its corresponding gradient should also be relatively large, so that the weight

parameters can be updated faster. The "distance" between the actual result and the predicted result can be the Euclidean distance, which corresponds to the quadratic cost function.

4.3. Example Application and Analysis. After the key indicators of urban architectural design are obtained through the convolution operation, the extracted key indicators of urban
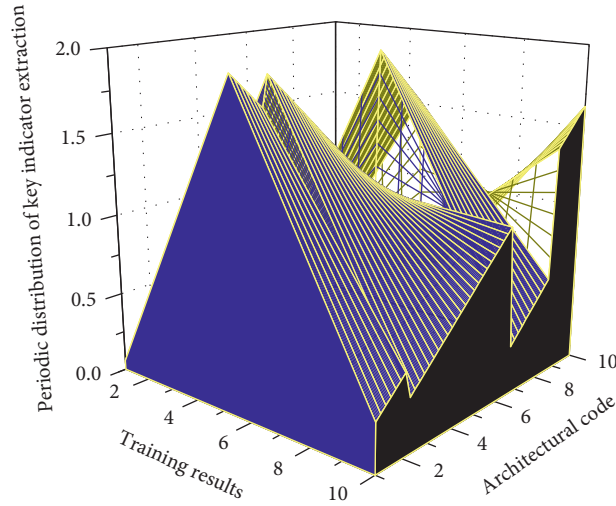
FIGURE 7: Distribution of extraction period of key indicators of architectural design.

architectural design are used for classification or segmentation operations. The common practice is to train a classifier (Softmax multiclassifier) using all the extracted key indicators of urban architectural design. At this time, there is a problem that must be solved, that is, the large amount of calculation. For example, for a 256×256 pixel urban building image, 100 key urban architectural design indicators defined on the 32×32 pixel input have been learned and obtained, and the convolution operation of each urban architectural design key index and the urban building image will be calculated. We get a $(256 - 32 + 1) \times (256 - 32 + 1) = 50625$-dimensional convolutional urban architectural design key indicators, since there are 100 urban architectural design key indicators, each input will get a $50625 \times 100 = 5062500$ dimensional urban building design key indicators vector.

$$\forall m(t, t-1, t+1 | t > x), \exists \lim_{x \longrightarrow \infty} \sum_{i=1}^{t} \left[X_i + Y + \mathrm{cov}(x, y) \right] - 1 = 0. \tag{15}$$

In order to better extract the global information in urban building images and retain the edge information of the target, this paper further proposes a deep segmentation network model combining convolutional long short-term memory network and fully connected conditional random field to improve the effect of semantic segmentation. Considering the fully connected conditional random field as a layer of a long-term and short-term convolutional memory network, and embedding it into the proposed multi-scale residual attention deep network model, a novel end-to-end deep segmentation model is designed; using the fully connected conditional, the random field constrains the pixels with similar key information of urban architectural design such as color and location in the urban architectural image, and makes full use of the local information and global context information of the urban architectural image to further improve the semantic segmentation effect.

$$\left[\sum_{i=1}^{t} Y_i \sum_{i=1}^{t} X_i \right] \times \begin{bmatrix} \sum (u + mt) \\ \sum (u - mt) \end{bmatrix} - \begin{bmatrix} 1 & -\exp[\pi(u)] \\ \exp[\pi(u)] & 1 \end{bmatrix} = 0. \tag{16}$$

When solving the minimum value by stochastic gradient descent, if the learning rate is too large, the value of the loss function may not tend to decrease as the number of iterations increases; if the learning rate is too small, it will find the correct descent direction, the number of iterations has been increasing, but the value of the loss function is almost unchanged. This paper tried different learning rates, including 1e-1, 1e-2, 1e-3, 1e-4, and 1e-5, and kept the learning rate updated during the training process to reduce the error of network training. The training results confirmed that 1e-4 can meet the requirements of the network for the stochastic gradient descent algorithm. The spatiotemporal data set used in urban areas is spatiotemporal sequence data. To predict the flow of a certain area, the key indicators of time-dependent urban building design still occupy the main factor. Therefore, in the experimental process of Figure 7, separate

time-dependent urban buildings are used respectively. The correlation between key design indicators, time-dependent key indicators of urban architectural design and key indicators of urban architectural design.

For the extraction period P and extraction length L of different parts of the time-dependent urban architectural design key indicators, the extraction period is set as the adjacent time stamp for the proximity through the above statistical analysis, the periodicity is set to one day, and the trend is set to 7 days. The extraction length is set based on experience. The proximity is the closest to the prediction target in the time dimension, and is set to 1 or 3, while the periodicity and trend are relatively far away, so this section sets it to 1. The optimal parameters are obtained by experimental verification, and for some hyperparameters of the gradient boosting tree model itself, such as the depth of the tree, the number of decision trees, the learning rate, etc., the optimal parameters are obtained by grid search.

In the original input data, local features are extracted one by one for each small region, and several important concepts are defined in DCNN. For some ultralarge databases such as the ImageNet Large-scale Urban Building Design Recognition Challenge dataset, the dataset contains 1.2 million samples. If the full gradient descent method is used, 1.2 million samples are firstly performed each time. It is an unwise choice in terms of improving time efficiency and reducing memory usage.

$$M(t|c, t \longrightarrow (1,1)) = \begin{cases} f(u+mt) + Z(t), & t = 0,1,2,\ldots,n, \\ \sum_{i=1}^{t} Y_i + u, & i = 0,1,2,\ldots,n. \end{cases}$$
(17)

In the extraction stage of key indicators of urban architectural design, first of all, this paper uses FCN. The framework is used as a fully convolutional long and short-term convolutional memory network, on the one hand, through the public pretraining model VGG network parameters, on the other hand, set the training parameters of FCN-8s: mini-batch size = 50, learning rate = 1.04, momentum = 0.9, weight decay = 54 and epoch = 500, and use stochastic gradient descent algorithm to FCN for training and fine-tuning: then, this paper upsamples the convolution or deconvolution of urban architectural design key indicators at each layer, and cascade all up-sampled urban architectural design key indicators to generate multi-dimensional urban architectural design features, and by calculating the mean value of the key indicators of multi-dimensional urban architectural design of all pixels in the superpixel, the superpixel multidimensional urban architectural design key indicators of urban architectural design are generated.

$$\lim_{x \to \infty} \delta[q(t,x)] - \lim_{x \to \infty} [1 + q(x,t)]h = \begin{cases} q(x,t) - 1, & q > x, \\ 1, & q < x. \end{cases}$$
(18)

On this basis, training DBNs can be divided into two stages: pretraining and fine-tuning. In the pretraining stage, this paper uses an unsupervised greedy algorithm to train RBMs in DBNs layer by layer from the bottom up to obtain the initial DBNs parameters; in the fine-tuning stage, this paper adopts unsupervised wake. In order to make DBNs have classification function, this paper adds a classification and discrimination network (eg softmax function) on top of DBNs. At this time, the structure of DBNs is the same as that of standard feedforward neural network. This paper adopts supervised back-propagation algorithm to further tune DBNs parameters.

5. Conclusion

Considering that the convolutional long short-term memory network has a good effect in the extraction of time-series urban architectural design key indicators and urban architectural design key indicators, this paper designs and proposes a multi-scale urban architectural design key index model to solve urban problems. In the process of index extraction, the key index information of key urban architectural design can be effectively strengthened. Therefore, the convolutional long short-term memory network and attention mechanism are combined to be used in the semantic segmentation task of key index images of urban architectural design to effectively extract spatiotemporal information. At the same time, it strengthens and highlights the key index information of key urban architectural design, and the segmentation performance verifies the effectiveness of the proposed algorithm. Studies have shown that the adversarial training method can not only improve the performance of the generative network through the competition of the discriminative network, but also can effectively reduce the overfitting of the generative network in the training process. Through adversarial training, the structured reasoning embedded adversarial network for key indicators of urban architectural design can not only detect the inconsistency between the analysis results of key indicators of urban architectural design output from the network and the corresponding ones through the analysis and judgment of the discriminant network, but also through the competitive and confrontational network of the discriminant network. The parameters of each layer of the generation network are optimized, structured learning layer and the fusion layer of key indicators of urban architectural design, thereby significantly improving the consistency of the analysis results of key indicators of urban architectural design.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] J. Zhang, F. Chen, Y. Guo, and X. Li, "Multi-graph convolutional network for short-term passenger flow forecasting in urban rail transit," *IET Intelligent Transport Systems*, vol. 14, no. 10, pp. 1210–1217, 2020.

[2] S. Liao, L. Zhou, X. Di, B. Yuan, and J. Xiong, "Large-scale short-term urban taxi demand forecasting using deep learning," *Design Automation Conference (ASP-DAC)*, vol. 2, pp. 428–433, 2018.

[3] W. Jin, Y. Lin, Z. Wu, and H. Wan, "Spatio-temporal recurrent convolutional networks for citywide short-term crowd flows prediction," *Compute and Data Analysis*, vol. 11, pp. 28–35, 2018.

[4] J. H. Tan, Y. Hagiwara, W. Pang et al., "Application of stacked convolutional and long short-term memory network for accurate identification of CAD ECG signals," *Computers in Biology and Medicine*, vol. 94, pp. 19–26, 2018.

[5] Y. Jin, C. Qin, J. Tao, and C. Liu, "An accurate and adaptive cutterhead torque prediction method for shield tunneling machines via adaptive residual long-short term memory network," *Mechanical Systems and Signal Processing*, vol. 165, Article ID 108312, 2022.

[6] J. Zhang, H. Che, F. Chen, W. Ma, and Z. He, "Short-term origin-destination demand prediction in urban rail transit systems: a channel-wise attentive split-convolutional neural network method," *Transportation Research Part C: Emerging Technologies*, vol. 124, Article ID 102928, 2021.

[7] J. K. Das, A. Ghosh, A. K. Pal, S. Dutta, and A. Chakrabarty, "Urban sound classification using convolutional neural network and long short term memory based on multiple features," in *Proceedings of the Intelligent Computing in Data Sciences (ICDS)*, pp. 6–9, Fez, Morocco, October 2020.

[8] H. Yu, "Development of tourism resources based on fpga microprocessor and convolutional neural network," *Microprocessors and Microsystems*, vol. 82, Article ID 103795, 2021.

[9] L. Lin, Z. He, and S. Peeta, "Predicting station-level hourly demand in a large-scale bike-sharing network: a graph convolutional neural network approach," *Transportation Research Part C: Emerging Technologies*, vol. 97, pp. 258–276, 2018.

[10] X. Ma, J. Zhang, B. Du, C. Ding, and L. Sun, "Parallel architecture of convolutional bi-directional LSTM neural networks for network-wide metro ridership prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 6, pp. 2278–2288, 2018.

[11] J. Liu, B. Xiao, Y. Li, X. Wang, Q. Bie, and J. Jiao, "Simulation of dynamic urban expansion under ecological constraints using a long short term memory network model and cellular automata," *Remote Sensing*, vol. 13, no. 8, p. 1499, 2021.

[12] P. Westermann, M. Welzel, and R. Evins, "Using a deep temporal convolutional network as a building energy surrogate model that spans multiple climate zones," *Applied Energy*, vol. 278, Article ID 115563, 2020.

[13] S. Zhang, Y. Chen, and W. Zhang, "Spatiotemporal fuzzy-graph convolutional network model with dynamic feature encoding for traffic forecasting," *Knowledge-Based Systems*, vol. 231, Article ID 107403, 2021.

[14] T. Bogaerts, A. D. Masegosa, J. S. Angarita-Zapata, E. Onieva, and P. Hellinckx, "A graph CNN-LSTM neural network for short and long-term traffic forecasting based on trajectory data," *Transportation Research Part C: Emerging Technologies*, vol. 112, pp. 62–77, 2021.

[15] C. Wen, S. Liu, X. Yao et al., "A novel spatiotemporal convolutional long short-term neural network for air pollution prediction," *The Science of the Total Environment*, vol. 654, pp. 1091–1099, 2019.

[16] Y. Qin, H. Luo, F. Zhao, C. Wang, J. Wang, and Y. Zhang, "Toward transportation mode recognition using deep convolutional and long short-term memory recurrent neural networks," *IEEE Access*, vol. 7, Article ID 142353, 2019.

[17] M. Chen, X. Yu, and Y. Liu, "PCNN: Deep convolutional networks for short-term traffic congestion prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 11, pp. 3550–3559, 2018.

[18] L. Zhao, Y. Song, C. Zhang et al., "T-gcn: a temporal graph convolutional network for traffic prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 9, pp. 3848–3858, 2019.

[19] C. Li, L. Bai, W. Liu, L. Yao, and S. T. Waller, "A multi-task memory network with knowledge adaptation for multimodal demand forecasting," *Transportation Research Part C: Emerging Technologies*, vol. 131, Article ID 103352, 2021.

[20] C. Song, Y. Lin, S. Guo, and H. Wan, "Spatial-temporal synchronous graph convolutional networks: a new framework for spatial-temporal network data forecasting," *Artificial Intelligence*, vol. 34, no. 1, pp. 914–921, 2020.

[21] Y. Ai, Z. Li, M. Gan et al., "A deep learning approach on short-term spatiotemporal distribution forecasting of dockless bike-sharing system," *Neural Computing & Applications*, vol. 31, no. 5, pp. 1665–1677, 2019.

[22] X. Yang, Q. Zhu, P. Li, P. Chen, and Q. Niu, "Fine-grained predicting urban crowd flows with adaptive spatio-temporal graph convolutional network," *Neurocomputing*, vol. 446, pp. 95–105, 2021.

[23] J. Lu, Q. Zhang, Z. Yang, and M. Tu, "A hybrid model based on convolutional neural network and long short-term memory for short-term load forecasting," in *Proceedings of the Power & Energy Society General*, vol. 23, pp. 3–5, Atlanta, GA, USA, October 2019.

[24] K. Wang, C. Ma, Y. Qiao, X. Lu, W. Hao, and S. Dong, "A hybrid deep learning model with 1DCNN-LSTM-Attention networks for short-term traffic flow prediction," *Physica A: Statistical Mechanics and Its Applications*, vol. 583, Article ID 126293, 2021.