ISSN: 2664-9640

Wind power prediction based on LSTM-CNN optimization

Yang Liu, Liqun Liu*

School of Electronic Information Engineering, Taiyuan University of Science and Technology, Taiyuan 030024, China

Abstract

The irregularity and unpredictability of wind power are the main obstacles to the integration of wind power in the grid. In order to improve the accuracy of wind power forecasting, genetic algorithm (GA) is used to optimize the model based on convolutional neural network and long short-term memory network. Effective feature extraction is performed on the data. The data is divided into training data set and The backup data set is based on the backup data. The test data is collected after collecting the samples, and the other samples become the error test set to improve the wind power prediction, perform error correction, and improve the instability of LSTM-CNN's data prediction phenomenon. By analyzing the model prediction results of LSTM-CNN and GA-LSTM-CNN, and comparing them with the actual power, the results show that the GA-LSTM-CNN prediction model has higher accuracy.

Keywords

LSTM network; wind power prediction; genetic algorithm; deep learning.

1. Introduction

Due to the random regularity of wind energy, the power generated by wind energy will be affected by factors such as high-altitude temperature, humidity and air pressure in the weather, and it is difficult to predict the power of wind power. Currently, physical models, traditional statistics, and artificial intelligence methods are used to predict wind power. The advantage of the physical model is that even if there is no data, it can still predict wind power, but the error of the physical model is high; traditional statistical methods use data as research, which is relatively simple compared to the physical model, but is affected by nonlinear data. Predictive ability is limited; artificial intelligence methods mainly have many improvements in algorithms, which are extensions of traditional statistical methods. Literature [1] uses Elman neural network to decompose the wavelet packet, maintains mutual integration, and predicts wind speed, wind power and other parameters of wind farms. It is a brand-new detection method, and it also gives detailed use steps of the new method; Literature [2] extends the echo state network and uses a mathematical method supplementing the leakage integral ESN for power prediction; Literature [3] is based on the mapping and noise reduction autoencoder (SDAE) in the system network in many-to-many directions A multi-scale wind power prediction method is proposed; Literature [4] proposes to improve the fruit fly optimization algorithm (FOA), and on this basis, optimize the wind speed sub-sequence reconstruction parameters, the least square support vector machine parameters, and build Prediction model to achieve convergence speed and prediction accuracy improvement; Literature [5] uses cascaded convolutional neural network-gated loop prediction model, and uses wavelet decomposition on the basis of IMF1 (high frequency eigenmode function), from wind The mathematical characteristics of wind power, wind speed and wind direction are extracted from each subset of the data, and the mutual influence relationship between the data is studied; Literature [6] proposed a combination of multiple clustering and hierarchical clustering for the accuracy of wind power forecasting. This method classifies the historical power and historical weather in the training

ISSN: 2664-9640

samples, establishes multiple particle swarm optimization-backpropagation prediction models for different time periods to be tested corresponding to different data, and calls and calls the best time period to be tested. A similar prediction model improves the prediction accuracy; literature [7] proposed a probability model based on Markov chain for the impact of prediction uncertainty on the power system to predict the input wind energy flowing into the grid, and the experimental results verify the method's effectiveness Feasibility; Literature [8] proposed a low-complex pseudo-inverse matrix Legendre neural network (PILNNR), the radial basis function (RBF) unit exists in its hidden layer, which can be performed in a short period of 10 minutes to 60 minutes Interval wind power forecasting; Literature [9] uses two kinds of hybrid recursive dynamic neural networks to make short-term forecasts of wind power. This method uses wavelet decomposition of wind energy and wind speed time series, uses NAR and NARX recursive dynamic neural networks to regress the decomposed sub-sequences, and then gathers the individual output data characteristics of each sub-sequence and model construction, and obtains the final wind power prediction result.

Combining existing prediction methods, artificial intelligence methods have better prediction accuracy than physical models and statistical methods. Among all the current algorithms, Deep learning neural networks have broad prospects. The data extraction capabilities and descriptions shown by CNN and LSTM The relationship of time series is widely used in wind power prediction. This article optimizes and improves the wind power prediction model based on LSTM-CNN, and adds genetic algorithm (GA). Experiments verify that LSTM-CNN with genetic algorithm is added. The prediction effect of LSTM-CNN is better than that of LSTM-CNN.

2. LSTM-CNN neural network

2.1. Convolutional Neural Network (CNN)

The standard model of the convolutional neural network is composed of different functional layers. Among them, there are various hidden layers between the input and output layers, which can realize different feature extraction functions, as shown in Figure 1. This paper realizes the acquisition of wind power data based on the convolutional layer and pooling layer in this network.

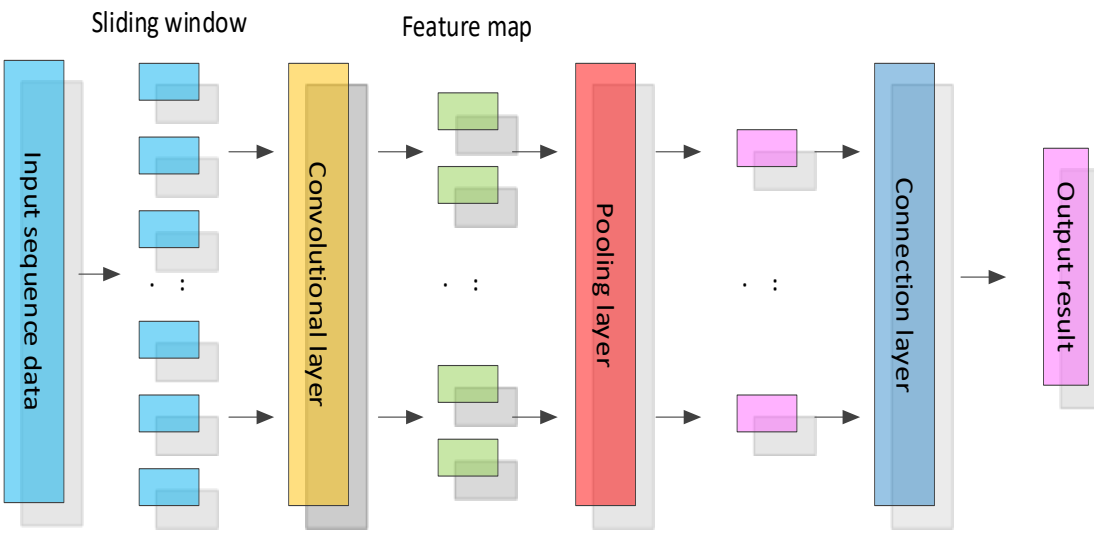

Fig. 1 Structure of convolution neural network

2.2. Long short-term memory network (LSTM)

The structure diagram of LSTM is shown as in Fig. 2. The key to LSTM is the cell state, which is the top transmission line in the picture, which is the location of information memory, which

varies with time. The control gate controls the change of information. The sigmoid function and the dot multiplication operation constitute the control gate. The sigmoid function ranges from 0 to 1. The dot multiplication operation controls how much information is transmitted. When the dot multiplication value is 0, no information transmission occurs. , When the value is 1, all information is transmitted. The memory gate, input gate, and forget gate in the memory network are used to control the opening and closing of the gate, so as to better extract the time relationship in the data and save it [11].

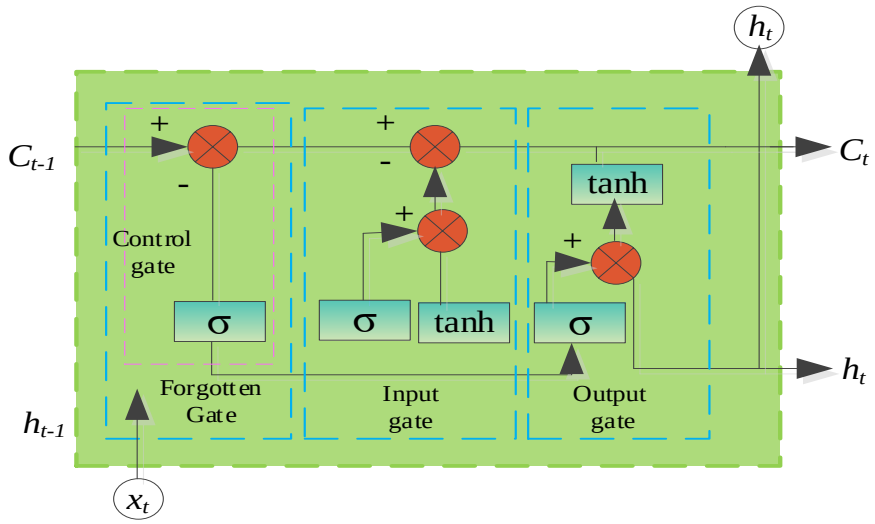

Fig. 2 LSTM structure

The working principle of LSTM is divided into the following steps:

Forgetting door: Forgetting some information in the past (the past information uses sigmoid(x,h) to control the door and forgetting some cell state), the calculation formula is:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{1}$$

Where: σ is the sigmoid activation function, W_f is the weight matrix multiplied by the input x_t and the hidden state h_{t-1} of the previous layer, b_f is the corresponding bias, and f_t is the output result of the forgetting gate.

Input gate: The input gate unit is stored in the input gate unit through the current input data characteristics (the information is extracted with the tanh(x,h) function, and the sigmoid(x,h) function is used for information selection. It is not used as the part to be stored. Delete), the calculation process is shown in formula 2:

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) \tag{2}$$

$$\tilde{C}_t = tanh\left(W_C \cdot [h_{t-1}, x_t] + b_C\right) \tag{3}$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{4}$$

In the formula: σ is the sigmoid activation function, W_i and W_C are the weight matrix of the input gate and input node and input x_t multiplied by the hidden state h_{t-1} of the previous layer, b_i and b_C are the corresponding biases, i_t, $\tilde{C}_t$, C_t are the output results of the state of the input gate and the input node memory unit respectively.

Output gate: output the processed information

$$o_t = \sigma\left(W_o [h_{t-1}, x_t] + b_o\right) \tag{5}$$

$$h_t = o_t * \tanh\left(C_t\right) \tag{6}$$

ISSN: 2664-9640

Where: σ is the sigmoid activation function, W_o is the weight matrix of the output gate and input x_t and the hidden state h_{t-1} of the previous layer, b_o is the corresponding bias, o_t and h_t are the output of the output gate and the hidden state result.

3. Wind power prediction model optimized based on LSTM-CNN

3.1. LSTM-CNN prediction model

The prediction process of the LSTM-CNN model is shown in Figure 3. First, combine the feature performance in the LSTM extraction process to extract the corresponding data feature information, and then use CNN for information extraction. According to the content of Figure 3, it can be judged that the historical data needs to enter the LSTM neural network first, collect time information, and then use the LSTM neural network for information transfer. The output information is the input of CNN and the data is extracted again to obtain the output result.

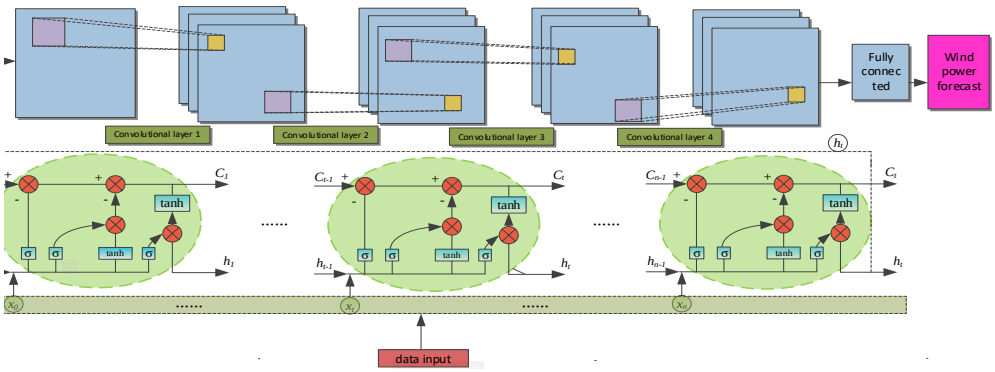

Fig. 3 Prediction process of LSTM-CNN model

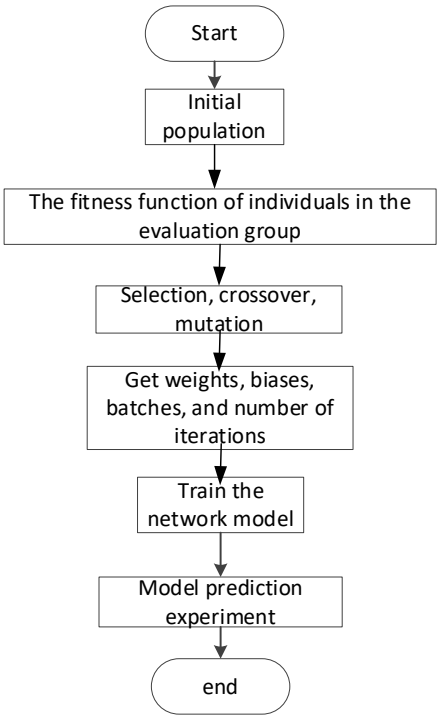

Fig. 4 Flow of genetic algorithm

3.2. GA-LSTM-CNN's wind power prediction model

3.2.1 Genetic optimization algorithm

Genetic Algorithm (GA) is the genetic algorithm. The process is shown in Figure 4. This method is based on the biological calculation method of sample extraction, feature fusion, and genetic evolution. Its core technology is to select a certain group to be optimized at first, establish a

ISSN: 2664-9640

calculation function suitable for this group, and finally evaluate and improve it. And with the eigenvalue calculation algorithm, the optimal solution is obtained through the final search.

(1) Initialize the population

After the genetic algorithm is started, it is necessary to detect the independent initial population through randomization. Under normal circumstances, in the early stage of initialization, the model scene will be established to estimate the scope

(2) Fitness function

The genetic algorithm appears after initializing the parent individual, and selects excellent individuals for inheritance based on the individual. Generally, fitness function is used to screen individuals, and the selection and setting of fitness function will have a direct impact on the quality of offspring.

(3) Selection

After initializing the population and setting the fitness function, the genetic algorithm mainly selects outstanding individuals. When making a selection, it is usually performed by the fitness value of the individual.

(4) Cross

After selecting high-quality individuals, the genes between chromosomes are combined with a crossover operator for fusion and conversion. After gene interaction and fusion, the algorithm will be more efficient in searching for effective information.

(5) Variation

Simultaneous processing of biological genetic thinking patterns, due to the selection and crossover of chromosomes, may lead to chromosome mutations, but the overall probability is not high.

(6) Circular evaluation

After the genetic algorithm is executed, Individual implementation of individual verification. Start to re-operate the cycle process. Check whether the calculated population meets the standard. If it meets the criteria, output the optimal solution; if it does not meet the criteria, continue to loop the above operations; if the number of loops meets the preset value standard, stop looping.

This paper chooses to apply genetic algorithm, optimize LSTM-CNN, and intelligently select the most suitable parameters to adaptively adjust the model parameters. The specific flow chart is shown in Figure 2.3 below:

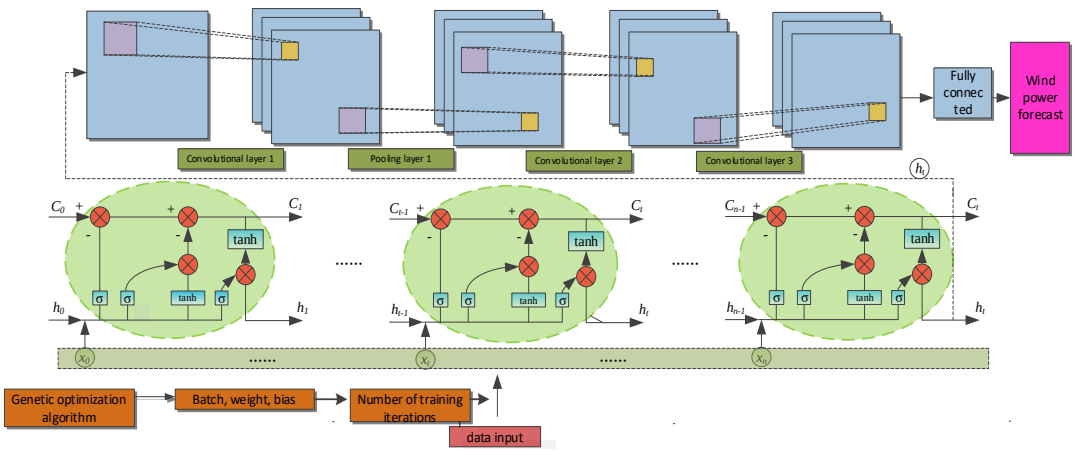

Fig. 5 Flow chart of lstm-cnn optimized by genetic algorithm

ISSN: 2664-9640

3.3. Evaluation Index

In order to judge the quality of the evaluation model, it is necessary to use the model error as the criterion. The error can be used to analyze the predictive performance of the model. Keeping the lower error indicates the higher the accuracy of the forecast.

Combine MAE, RMSE, MAPE and coefficient of determination (R^2) as the final standard for error evaluation. The average value of absolute error is MAE; RMSE can reflect the degree of data concentration or dispersion; MAPE emphasizes the error between the real value and the actual value, and reflects the closeness between the predicted value and the actual value, R^2 represents a variable The degree of interpretation of another variable. The calculation formula of each parameter is as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|z_i - x_i\right| \tag{7}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(z_i - x_i\right)^2} \tag{8}$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{z_i - x_i}{x_i}\right| \tag{9}$$

$$R^2 = 1 - \frac{\sum\left(x - \hat{x}\right)^2}{\sum\left(x - \bar{x}\right)^2} \tag{10}$$

Where: $\hat{x}_i$ means the estimated value of the i-th sample; x_i is the true value of the i-th sample; n is the number of samples.

4. Example analysis

4.1. Model parameter selection

In order to verify the reliability of the algorithm model, this paper selects the data of a wind farm in Shanxi Province, uses 80% of the data as the trainable data set, and the remaining 20% of the untrained data as the test set. The program code is python3.6.5 version, CPU The configuration is Intel Core i5-6200U, the frequency is 2.4GHz, the graphics card configuration is Graphics 520, the framework is Tensorflow, and the compilation environment is Pycharm.

The LSTM network step size is set according to 100, the time data dimension is set to 7, and the system's corresponding hidden layer neuron is set to 64, then the LSTM terminal time output is kept the same as the CNN input, and then CNN is used 3 times in the convolutional layer Get the data, get the data in the maximum pooling layer, get the convolution kernel as (3, 3), step 1, and its depth is 8, 16, 64, and then the calculated output is 1024, and finally the regression layer is fully connected The layer realizes the mapping of wind power.

Get the prediction result as shown in the figure below.

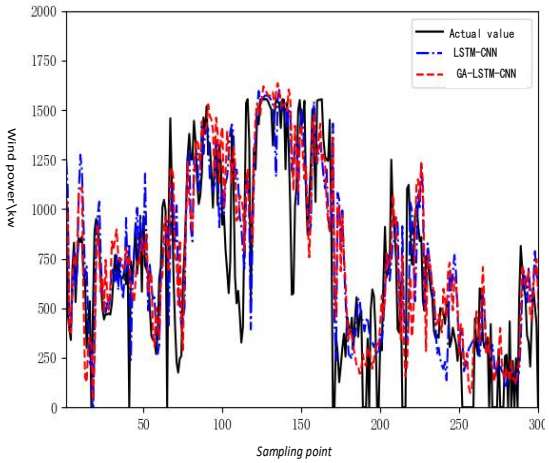

Fig. 6 GA-LSTM-CNN and LSTM-CNN prediction results

Intercept the data from the 130th to the 150th sampling points and zoom in and compare them as shown in Figure 7.

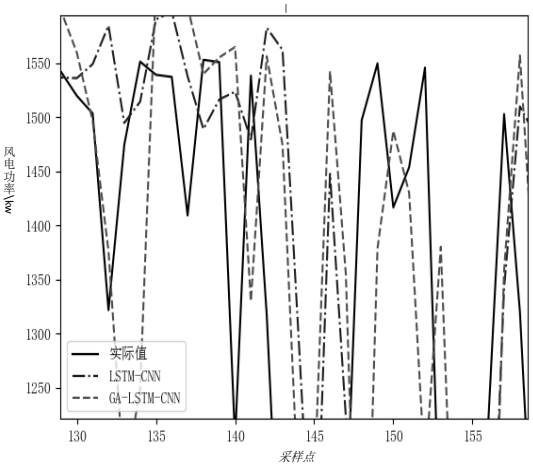

Fig.7 Enlarged figure of prediction results

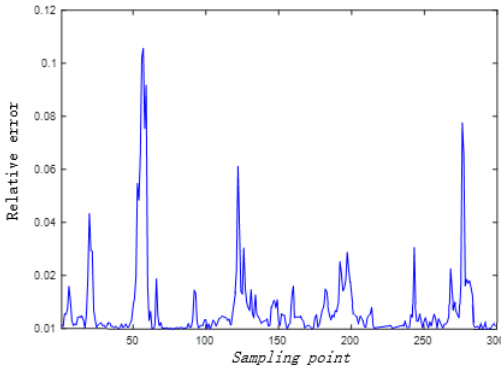

Fig. 8 Relative error

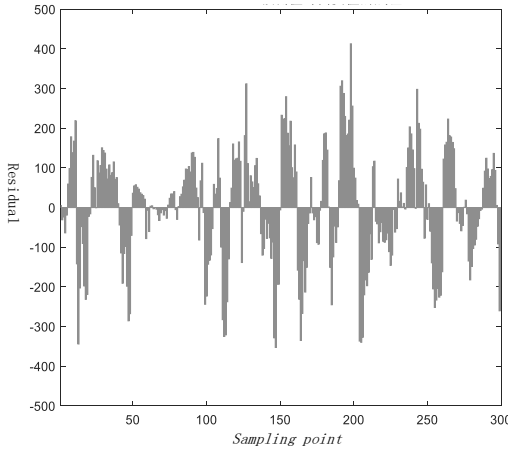

Fig. 9 Residual diagram

From the prediction results, it can be seen that the prediction accuracy of LSTM-CNN is higher than that of LSTM. At the same time, the evaluation indicators of the two methods can be seen in Table 1:

Table 1 Model prediction error analysis

model	RMSE	MAPE	MAE	R2
LSTM-CNN	0.8324	0.8884	0.6655	0.696
GA-LSTM-CNN	0.4816	0.5110	0.3643	0.926

It can be seen from Table 3-1: Compared with LSTM-CNN, the improved LSTM-CNN model has reduced RMSE by 42.14%, MAPE by 42.4%, MAE by 45.2%, and R2 by 33%.

5. Conclusion

Aiming at wind power prediction, this paper optimizes the prediction of the algorithm on the basis of LSTM-CNN, and adds a genetic algorithm to further optimize the settings of the parameters, which improves the prediction accuracy. The genetic algorithm is introduced to optimize the parameters of the LSTM, and then the feature information is extracted by applying the LSTM, and then the spatial feature information is extracted by using CNN to obtain the wind power prediction results. Through simulation, the improved GA-LSTM-CNN model compared with LSTM-CNN has reduced RMSE by 42.14%, MAPE by 42.4%, MAE by 45.2%, and R2 by 33%. The improved GA-LSTM-CNN model has a better prediction effect than the combined model of long and short-term memory and convolutional neural network.

Acknowledgements

Fund Project: Key Research Project of Shanxi Province (201803D121103).

References

[1] Ye Ruili, Guo Zhizhong, Liu Ruiye, et al. Wind farm wind speed and wind power prediction based on wavelet packet decomposition and improved Elman neural network[J]. Transactions of China Electrotechnical Society, 2017, 32(21): 103-111.

[2] Wang Xinyou, Wang Chenhua, Zhang Yi, Lu Jiang. Short-term wind power prediction based on improved ESN[J]. Automation and Instrumentation, 2017(02): 29-32.SSS K,AYUSH K,BABU R V. DeepFix: A Fully convolutional neural network for predicting human eye fixations[J].IEEE Translation on Image Processing,2017,26(9):4446-4456

[3] Cheng Qiming, Chen Lu, Cheng Yinman, et al. Wind power short-term prediction method based on EEMD and LS-SVM model[J]. Electric Power Automation Equipment, 2018, 38(05): 27-35.

[4] Yin Hao, Ou Zuhong, Chen De, Meng Anbo. Ultra-short-term wind power prediction based on quadratic mode decomposition and cascaded deep learning [J]. Power System Technology, 2020, 44(02): 445-453.

[5] Peng Chenyu, Chen Ning, Gao Bingtuan. Ultra-short-term wind power forecasting method combining multiple clustering and hierarchical clustering[J]. Automation of Electric Power Systems, 2020, 44(02): 173-181.D.C.F , ALFREDO V, Domenico V. A Markov chain-based model for wind power prediction in congested electrical grids[J]. The Journal of Engineering. 2019(18):4961-4964.

[6] S. P. MISHRA, P. K. DASH. Short-term prediction of wind power using a hybrid pseudo-inverse Legendre neural network and adaptive firefly algorithm[J]. Neural Computing and Applications. 2019, 31(7):2243-2268.

[7] P. K. SINGH, N. SINGH, R NEGI. Short-Term Wind Power Forecasting using Wavelet-based Hybrid Recurrent Dynamic Neural Networks. International Journal of Performability Engineering.2019, 15(7):1772-1782.

[8] [10] Zhang Lin, Yuan Feiniu, Zhang Wenrui, Zeng Xialing. A review of research on fully convolutional neural networks [J]. Computer Engineering and Applications, 2020, 56(01): 25-37.