2022 21st International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES)

# Fault Diagnosis of Power Grid Based on Convolutional Neural Network

1st Liping QU, 2nd Jie Zhang
School of Electrical Engineering and Information,
Beihua University
Jilin, China
Email:*goudan1990@126.com*
*Email:1784318406@qq.com*

3rd Tailu Gao
School of Electrical Engineering and Information,
Beihua University
Jilin, China
Email:*Gaotailu135102@163.com*

***Abstract*—Due to the operation of regional networking, the scale of the power grid is becoming larger and larger, and a fault in the power grid needs to be located in the fault area timely and accurately. The models and structures of BP neural network and convolution neural network are analyzed. The training and test samples are constructed for a power grid model, and the BP neural network and convolution neural network are used for simulation verification respectively. The simulation results show that the convolutional neural network based grid fault diagnosis method has higher accuracy and fault tolerance.**

***Keywords-BP Neural Network ; Convolutional Neural Network; Fault Diagnosis; Power Grid***

## I. INTRODUCTION

When the grid fault occurs, if it cannot be accurately located to the fault location and repaired, it is likely to cause large-scale cascading failures and bring huge losses to the power supply department and users. The effective diagnosis of grid fault is an important means to ensure the safe and reliable operation of the grid [1]. The research on power grid fault diagnosis can be traced back to the 1940s. With the continuous development of power technology, many new methods have emerged in the field of power grid fault diagnosis. For example, expert system, information theory[2] , artificial neural network, Bayesian network and Petri net theory. Artificial neural network can simulate the information transmission between neurons and has strong learning ability, so it is widely used in the field of power grid fault diagnosis. However, the neural network is easy to fall into local minima in the process of learning the model, which affects the accuracy of fault diagnosis. In order to solve the problem that the neural network is easy to fall into the minimum value, usually based on the neural network, various optimization algorithms are introduced, such as optimizing the weights and thresholds of BP neural network by genetic algorithm[3] , and optimizing the network parameters by least square method[4] .

These optimization methods improve the fault diagnosis ability of BP neural network to some extent, but also make the optimization and fault diagnosis process too complex. How to find a simple but effective method has become a research hotspot based on neural network fault diagnosis. Based on this, this paper introduces convolutional neural network which has achieved great success in the field of image recognition into power grid fault diagnosis, constructs a power grid fault diagnosis method based on convolutional neural network, and compares it with the classical BP neural network.

## II. BASIC PRINCIPLES AND LEARNING ALGORITHMS OF BP NEURAL NETWORKS

BP neural network is widely used in the neural network. As a feedforward neural network, it mainly includes the input layer, the hidden layer and the output layer . There are corresponding weights between the input layer and the hidden layer, as well as between the hidden layer and the output layer. The model is shown in Figure 1.

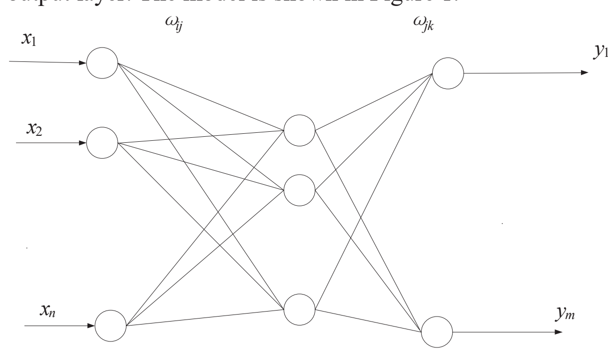


Figure 1. BP neural network topology

The learning algorithms of BP neural network include forward propagation and error back propagation.

In forward propagation, the signal is input from the input layer and transmitted to the output layer through the hidden layer. If the actual output of the output layer is equal to the expected output, the algorithm ends, or it turns to reverse propagation.

Back propagation through the sample expected output and the actual output of the error, reverse transmission calculation, constantly correct the weights and thresholds of neural network, reduce the error.

BP learning algorithm steps are as follows :

① Set the initial weights and thresholds ;

② The input and output samples of the network are given to determine the structure of the network ;

③ Calculate the values of hidden layer and output layer ;

$$H_j = f(\sum_{i=1}^{n} \omega_{ij} x_i - a_j) \tag{1}$$

In the formula, $i = 1, \dots, n$, the number of nodes in the input layer ; $j = 1, \dots, l$, is the number of hidden layer nodes ; $f$ Activation function ; $\omega_{ij}$ is the weight input to the hidden layer ; $x_i$ is the input information ; $a_j$ is the hidden layer threshold.

2473-3636/22/$31.00 ©2022 IEEE
DOI 10.1109/DCABES57229.2022.00033

$$O_k = g(\sum_{j=1}^{l} \omega_{jk} H_j - b_k) \quad (2)$$

In the formula, $k = 1, \dots, m$, the number of nodes in the output layer ; $g$ is the output layer activation function ; $\omega_{jk}$ is the weight between the hidden layer and the output layer ; $b_k$ is the output layer threshold.

④ Calculation error and objective function ;

$$e_k = Y_k - O_k \quad (3)$$

$$E_p = \frac{1}{2} \sum_k e_k^2 \quad (4)$$

$$J = \sum_p E_p \quad (5)$$

⑤ Compare total error with expected error ;

If $J \leq \varepsilon$, the algorithm ends ; otherwise turn to step ⑥.

⑥ Back propagation;

Update of weights and thresholds from input layer to hidden layer :

$$\omega_{ij}(t+1) = \omega_{ij}(t) - \eta \frac{\partial J(t)}{\partial \omega_{ij}(t)} \quad (6)$$

$$a_j(t+1) = a_j(t) - \eta \frac{\partial J(t)}{\partial a_j(t)} \quad (7)$$

Weight and threshold update from hidden layer to output layer :

$$\omega_{jk}(t+1) = \omega_{jk}(t) - \eta \frac{\partial J(t)}{\partial \omega_{jk}(t)} \quad (8)$$

$$b_k(t+1) = b_k(t) - \eta \frac{\partial J(t)}{\partial b_k(t)} \quad (9)$$

The neural network needs to activate the function in the learning process. Usually, Sigmoid function and Tanh function can be selected.

## III. STRUCTURE AND LEARNING ALGORITHM OF CONVOLUTIONAL NEURAL NETWORKS

### A. *Structure of convolutional neural networks*

Convolutional neural network is one of the important models in deep learning. In recent years, convolution neural network has achieved great success in image classification and speech recognition. In the calculation of traditional BP neural network, all data are calculated. This calculation model lacks attention to the essential characteristics of data, and it will also cause large amount of calculation and even over-fitting, resulting in poor effect. Convolutional neural network is different, its advantage is to extract features before computing, and then calculate the eigenvalues rather than directly calculate the original value. Based on the powerful classification ability of convolutional neural network, this paper studies its application in the field of power grid fault diagnosis, and uses it to realize the fault diagnosis and location of power grid.

As one of the representative algorithms of deep learning, convolutional neural network has achieved great development in many fields. Similar to other feedforward neural networks, convolutional neural networks generally include input layer, hidden layer and output layer. Unlike other neural networks, the hidden layer generally includes three structures : convolution layer, pooling layer and fully connected layer. When the sample is input into the model, the convolution layer will process and extract its characteristics[6]. Then the activation layer transforms the results of the convolution layer once, and then connects with the pooling layer[7-10]. The main function of pooling layer is to extract and reduce features. After the pooling layer is the full connection layer, which is equivalent to a classifier that classifies the original samples. Figure 2 is a typical LeNet-5 convolutional neural network.

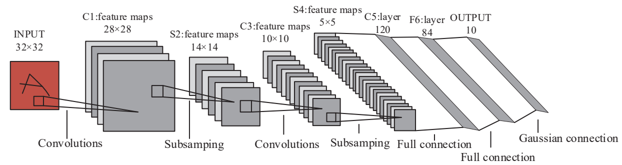


Figure 2. LeNet-5 convolutional neural network

The core of convolutional neural network is convolution operation, and the convolution layer completes convolution operation. The convolution kernel operates with the input matrix and is connected to the pooling layer. The essence of the convolution kernel is equivalent to the filter. When performing the inner product operation, the convolution kernel will calculate the input data according to the set step size. During the operation, the convolution kernel moves one step at a rate from left to right, from top to bottom. In the convolution layer, the size of the convolution kernel is fixed, but its number can change. The above operation of convolution kernel is actually a process of feature extraction of input data, and the output matrix can be obtained output matrix can be obtained. Figure 3 is a simple schematic of convolution operation. When the step size is 1, the output is a 2 × 2 matrix. If the step size is changed, the corresponding output matrix will also change. The convolution formula is shown below.

$$I^{l+1}(i,j) = \sum_{K=1}^{K_l} \sum_{x=1}^{f} \sum_{y=1}^{f} \left[ I_K^l (s_0 i + x, s_0 j + y) w_k^{l+1}(x,y) \right] + b_{l+1} \quad (10)$$

$$L_{l+1} = \frac{L_l + 2p - f}{s_0} + 1 \quad (11)$$

In the formula, $i$、$j \in \{0,1,\dots,L_{l+1}\}$, $I^l$ is the input of the $l+1$ layer, $I^{l+1}$ is the output, $w$is the weight, $b_{l+1}$ is the offset of the $l+1$ layer, $L_{l+1}$ is the size of the output matrix, $K$ is the number of channels, $p$ is the number of filled, $s_0$ is the set step size.

$$\begin{array}{|c|c|c|}\hline 1 & 1 & 1 \\ \hline 0 & 1 & 1 \\ \hline 1 & 1 & 0 \\ \hline \end{array} \quad * \quad \begin{array}{|c|c|}\hline 0 & 1 \\ \hline 2 & 3 \\ \hline \end{array} \quad = \quad \begin{array}{|c|c|}\hline 4 & 6 \\ \hline 6 & 3 \\ \hline \end{array}$$

Figure 3. Convolution operation

In convolution operation, the number of times that the edge information participates in the calculation is relatively small, and the calculation stops when the convolution kernel moves to the edge. However, the information located in the middle can be repeatedly involved in the calculation. It is precisely because of this feature that the edge information is likely to be lost after operation. In order to solve this problem, Padding can be introduced, that is, some information is added to the periphery, usually 0. By introducing filling, the stability of boundary information can be maintained, and some irregular matrices can also be filled to make them regular matrices.

Usually there will be activation function after the convolution layer. If the activation function is not used in the operation, the input and output will become a simple linear relationship. For a practical classification problem, if the classification problem is relatively simple, the linear output may be able to solve the problem. When faced with complex and linear inseparable problems, the linear output is difficult to meet the requirements, so it is necessary to introduce nonlinear factors into the activation function. Common activation functions include Sigmoid, Tanh, ReLU, etc. Here are only ReLU functions.

One of the advantages of ReLU function is that it does not appear the saturation phenomenon of Sigmoid function and Tanh function, and its convergence speed is faster in training. Therefore, ReLU function is generally selected as the activation function of convolutional neural networks, as shown in formula ( 12 ).

$$ReLU(x) = \max(x, 0) \tag{12}$$

The input enters the pooling layer after the convolution layer output. The purpose of pooling operation is to sample, process the feature matrix, and extract the most important features. At the same time, the feature matrix will also be reduced after pooling operation, which is equivalent to the dimension reduction operation and simplifies the computational complexity. The two most commonly used pooling operations in pooling layer are maximum pooling and mean pooling. The maximum pooling is equivalent to the maximum in the data area, and the average pooling is similar to the operation of taking the average.

The maximum pooling is shown in Figure 4, and the step size is set to 2.

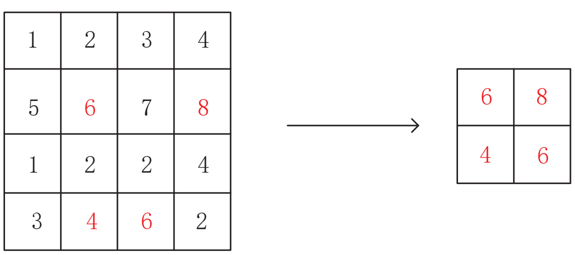


Figure 4 Maximum pooling diagram

The full connection layer is the connection of two neurons, and all the features are combined. Full connection layer realizes classification and transformation from matrix data to one-dimensional data through feature mapping. The output results can be transformed into the probability values between 0 and 1, and the classification results are output in the form of probability. The function of the output layer is usually SoftMax function.

### *B. learning algorithm of convolutional neural networks*

Convolution neural network is also a feedforward neural network in essence, and its model training is similar to that of BP neural network. The update of parameters also relies on the error back propagation algorithm, and the weights and biases are gradually adjusted according to the cost function. The ultimate goal is to minimize the cost function, so that the actual output approaches the expected output. The cost function has various forms. The mean square error between the expected output and the actual output is used as the representation of the cost function as follows :

$$E(w, b) = \frac{1}{2}\sum_{n=1}^{N}(t_n - y_n)^2 \tag{13}$$

In the formula, $N$ is the number of training samples, $t_n$ is the expected value, $y_n$ is the actual output value.

Learning process of convolutional neural network

1) Network initialization, get the weight of each layer w and offset b ;

2) The input data are transmitted forward through each layer until the output layer ;

Convolution layer calculations : $x_1^l = f(w^l x_1^{l-1} + b^l)$

Pooling layer calculation : $x_2^l = pool(x_2^{l-1})$

Full connection layer calculation :

$x_3^l = f(w^l x_3^{l-1} + b^l)$

Output layer calculation :

$x_4^l = SoftMax(w^l x_4^{l-1} + b^l)$

3) According to the results of the output layer and the expected output error ;

4) To judge whether it meets the requirements or not, if it fails to meet the expected error requirements, it begins to propagate back into 5 ) ; if the error meets the requirements, the training is completed ;

5) Full connection layer error transfer ;

6) Pooling layer error transfer ;

7) Error transfer of convolution layer : firstly, the error is filled, and then the convolution layer is rotated by 180° to calculate the error.

8) Parameter update ;

full connection layer :

$w^l = w^l - \alpha \sum_{i=1}^{m} \delta^{i,l} (x_3^{i,l-1})^T$ ,

$b^l = b^l - \alpha \sum_{i=1}^{m} \delta^{i,l}$;

convolution layer :

$w^l = w^l - \alpha \sum_{i=1}^{m} \delta^{i,l} rot180(x_1^{i,l-1})$,

$b^l = b^l - \alpha \sum_{i=1}^{m} (\delta_j^{i,l})_{u,v}$;

9) Transfer to 2).

## IV. CASE ANALYSIS

In this paper, four-bus system as the simulation object, and the topological structure of the system is shown in Figure 5.

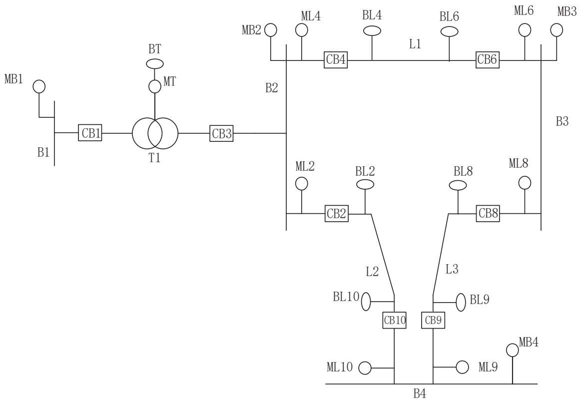


Figure 5. System wiring diagram

The system model consists of four buses B1, B2, B3, B4, a transformer T and three transmission lines L1, L2, L3. MB1 ~ MB4 is the main protection of bus B1 ~ B4, ML represents the main protection of the line, BL represents the backup protection of the line, and also the backup protection of the bus. A total of 20 training samples are constructed, including the normal action of the main protection and circuit breaker, the action of the main protection refusing to operate backup protection, and the refusing action of circuit breaker.

In order to compare the diagnostic effect of convolutional neural network and BP neural network, the fault test samples shown in Table 1 are used.

Table 1 Fault test samples

| Testing samples | Protection or circuit breaker action | Faulty section |
|---|---|---|
| 1 | MB1、MB2 | B1 |
| 2 | MB2、CB3 | B2 |
| 3 | MB3、CB8 | B3 |
| 4 | MB4、CB10 | B4 |
| 5 | MT、CB3 | T |
| 6 | ML4、ML6、CB6 | L1 |
| 7 | ML2、ML10、CB10 | L2 |
| 8 | ML8、ML9、CB9 | L3 |
| 9 | MB1、BL4、CB1 | B1 |
| 10 | MB2、ML8、CB2、CB3、CB4 | B2 |
| 11 | MB3、BL2、CB6 | B3 |
| 12 | MB4、BL2、CB9、CB10 | B4 |
| 13 | MT、BT | T |
| 14 | ML4、ML6、CB6、CB8 | L1 |
| 15 | ML2、ML10、CB2、CB6、CB10 | L2 |
| 16 | ML8、ML9、CB8、CB9、CB3 | L3 |
| 17 | CB6 | NO |

Convolution neural network and BP neural network are used for simulation respectively. BP neural network uses Tanh function and L-M algorithm.

A convolutional neural network is constructed, which includes input layer, two convolution layers, two pooling layers, full connection layer and output layer. The convolution layer is 2 × 2, the step length is 1, the pooling layer adopts the maximum pooling, the first pooling step is 2, the second layer is set to 1, and the activation function is ReLU function.

The constructed network is used to train the samples in the fault diagnosis decision table, and 17 fault samples in table 3.1 are used to test. In order to output the diagnosis intuitively, the confusion matrix is used to represent the fault diagnosis results of the test set. The number 1 ~ 9 represents 9 areas such as B1, B2, B3, B4, T, L1, L2, L3, NO ( no fault ), and the abscissa represents the target area, and the ordinate represents the area where the actual diagnosis is located. The classification diagnosis of 9 regions is shown in Figure 6 and Figure 7.

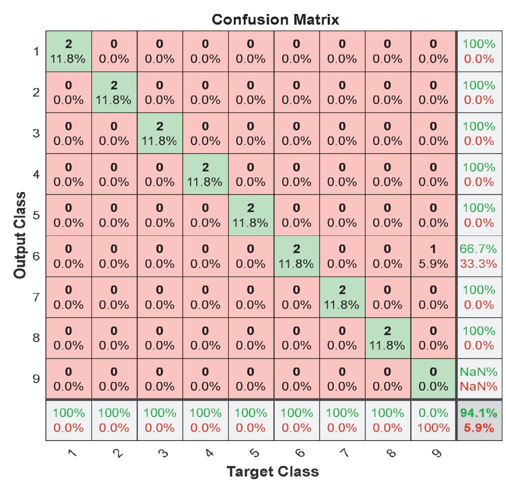


Figure 6. Fault diagnosis results of convolutional neural network

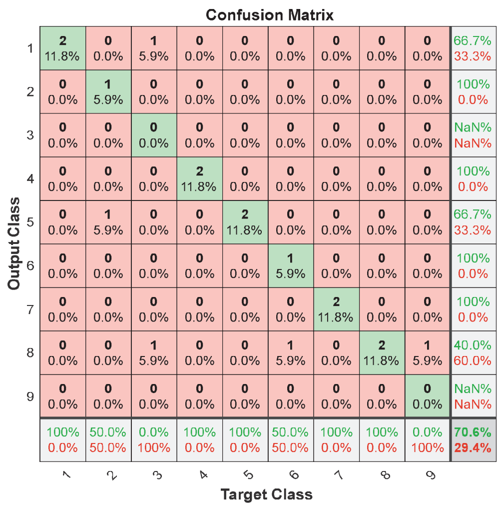


Figure 7. Fault diagnosis results of BP neural network

Compared with the fault diagnosis results of convolutional neural network and BP neural network, the total correct rate of convolutional neural network is 94.1 %, and only one of the 17 fault test samples is not correctly diagnosed. For 17 test samples of BP neural network, 5 samples of BP neural network are not correctly located in the fault area, and 5 fault diagnosis samples are bus B2 fault, bus B3 fault, line L1 fault and fault-free area ( NO ). For the test sample of bus B3, the fault BP neural network of the two test samples is not diagnosed, and the overall accuracy is only 70.6 %, and its generalization ability and fault tolerance are worse than those of convolution neural network. The reason for this difference is that the learning ability of BP neural network is weaker than that of convolution neural network. There are feature extraction operations in the learning process of convolution neural network, and the deep-seated rules of learning samples are continuously strengthened in the learning process. While BP neural network learns all samples at the same time, and there is no feature extraction process. Therefore, its learning ability is poorer than that of convolution neural network, and the accuracy of diagnosis is also relatively low.

The following table is the performance comparison table of convolutional neural network and BP neural network.

Table 2 Performance comparison between CNN and BP

| Network | Accuracy | Training and testing time |
|---|---|---|
| BP | 70.6% | 2.7s |
| CNN | 94.1% | 6s |

It can be seen from Table 2 that the time of BP neural network is less than that of convolution neural network in training and testing. This is because the structure of BP neural network is relatively simple, and the structure of convolution neural network is relatively complex. In particular, the full connection layer contains a large number of neurons, which increases the calculation amount.

## V. Conclusion

By comparing the diagnostic results of convolutional neural network and BP neural network, this paper verifies the effectiveness of the convolutional neural network power grid fault diagnosis method. Using the deep learning ability of convolutional neural network, the fault training samples are deeply studied, and the most essential features are extracted. The accuracy of fault test sample diagnosis proves that the convolutional neural network has a unique effect in power grid fault diagnosis. However, the full connection mode of convolutional neural network is too redundant. How to further improve the computational efficiency of convolutional neural network needs further study.

## References


[1] Yuan Pu, Mao Jianlin, Xiang Fenghong , et al. Grid Fault Diagnosis Based on Improved Genetic Optimization BP Neural Network[J]. Proceedings of the CSU-EPSA, 2017,29(01) : 118-122.

[2] Zhang Xu, Wei Juan, Zhao Dong Mei, et al. Research Course and Prospects of Power Grid Fault Diagnosis[J]. Power System Technology, 2013,37(10):2745-2753.

[3] Zhang Xi, Guo Xiang, Yu Yunhao , et al. A fault diagnosis method of power grid based on BP neural network optimized by genetic algorithm[J]. Technology Innovation and Application,2019(29):15-17.

[4] Xiong Guojiaqng . Power Grid Fault Diagnosis Based on Radial Basis Function Neural Network[D]. Hubei: Huazhong University of Science and Technology,2012.

[5] Jiang Bo, Li Zhenjia. Research on power grid fault diagnosis based on BP neural network [J]. China Energy and Environmental Protection,2017,39(11):246-249

[6] Zhai Haibao, Wang Xingzhi, Ge Minhui, et al. Fault Diagnosis Method for Transmission Lines Based on the Improved Convolutional Neural Network [J]. Electrical automation, 2021,43 ( 02 ) : 105-108.

[7] Xiao Xiong, Wang Jianxiang, Zhang Yongjun, et al. A Two-dimensional Convolutional Neural Network Optimization Method for Bearing Fault Diagnosis[J]. Proceedings of the CSEE, 2019,39(15):4558-4568.

[8] Zhao Danyang, Dong Weiguang, Gao Fengyang. Improved Inverter Fault Diagnosis Based on Convolutional Neural Network[J]. Journal of Power Supply, 2020,18(03):124-132.

[9] Zhang Ke, Feng Xiaohan, Guo Yurong, et al. Overview of deep convolutional neural network for image classification [J].Chinese Journal of Image Graphics, 2021,26 ( 10 ) : 2305-2325.

[10] Wang Ting, Zhu Xiaofei, Tang Gu. Knowledge-enhanced graph convolutional neural networks for text classification [J]. Journal of Zhejiang University ( Engineering Edition ), 2022,56 ( 02 ) : 322-328.