ISSN 1330-3651 (Print), ISSN 1848-6339 (Online)
https://doi.org/10.17559/TV-20220427035848
Original scientific paper

Circuit Breaker Fault Diagnosis Method Based on Improved One-Dimensional Convolutional Neural Network

Jie SHI, Guoqing DU*, Haifeng SHEN, Fei DING, Weixiang KONG

Abstract: Aiming at the problems of manual feature extraction and poor generalization ability of model in traditional circuit breaker fault diagnosis technology, a circuit breaker fault diagnosis method based on improved one-dimensional convolutional neural network is proposed. Firstly, the input feature sequence is adaptively weighted by self-attention mechanism to highlight the weight of important information; Secondly, 1 × 1 convolution layer and global average pooling layer are used to replace the full connection layer, which reduces the model training parameters, improves the training efficiency and prevents the phenomenon of over-fitting. Aiming at the problem of small number of data samples, the data is enhanced by Generative Adversarial Network. After adding the generated data to the original data, the accuracy of fault identification is further improved. The experimental results show that this method can effectively and accurately identify different fault types of circuit breaker, and verify the feasibility of its engineering application.

Keywords: circuit breaker; data enhancement; global mean pooling; one-dimensional convolution; self-attention

1 INTRODUCTION

Low voltage circuit breaker is the key equipment to ensure the safety of low-voltage distribution system, and its health state affects the performance and stability of distribution system [1]. Therefore, it is of great significance to study an efficient and accurate fault diagnosis method of low-voltage circuit breaker.

In recent years, deep learning has been widely used in the field of fault diagnosis because of its powerful feature extraction ability. Huang et al. [2] manually extracted the current characteristics of opening and closing coils, and used convolutional neural network to diagnose the fault of circuit breaker; Xing et al. [3] extracted the characteristics of circuit breaker opening and closing action signal by improving multi-scale basic entropy algorithm, and used deep belief network for fault diagnosis; Ling et al. [4] used bispectrum analysis and wavelet analysis for feature extraction, and used dual stream convolution neural network for fault diagnosis; Sun et al. [5] used the one-dimensional convolutional neural network with the first convolution kernel as the wide convolution kernel to diagnose the circuit breaker fault. Yan et al. [6] denoised the current signal by ensemble empirical mode decomposition, extracted the key features by time-domain extremum method, and then continued fault diagnosis by convolutional neural network; Yang et al. [7] used Hilbert Huang transform to convert one-dimensional vibration signal into three-dimensional time-frequency image, and used convolutional neural network to diagnose circuit breaker fault; Cao et al. [8] proposed a hybrid depth network based on convolutional neural network and long short-term memory to diagnose the fault of circuit breaker. Although the above research has achieved some research results, there are still the following problems: 1) The opening and closing coil current of the circuit breaker is a one-dimensional time sequence signal, which adopts manual feature extraction without considering the sequence relationship of each time; 2) The traditional convolutional neural network gives the same weight to each feature of the input timing signal, and does not consider which features in the original signal can represent the fault information and which may cause interference; 3) Due to the existence of full connection layer, the traditional convolutional neural network occupies a lot of computing resources, which is not conducive to the scene of real-time online monitoring.

To solve the above problems, this paper proposes a circuit breaker fault diagnosis method based on improved one-dimensional convolutional neural network (ICNN-1D). Firstly, the self-attention mechanism is introduced to adaptively weight the input sequence features to highlight the useful fault feature information. Secondly, one-dimensional convolution neural network is used to extract the features of the original current signal directly, so as to avoid the subjectivity of manual feature extraction. Finally, 1 × 1 convolution layer and global average pooling layer are used to replace the full connection layer, which reduces the amount of parameters and calculation of the model and is conducive to rapid fault diagnosis. Aiming at the problem of less data samples, the original data is enhanced by Generative Adversarial Network (GAN), and the generated data samples are added to the original data samples for training. Finally, the fault diagnosis accuracy of the trained model is improved.

2 THEORETICAL BASIS

2.1 One Dimensional Convolutional Neural Network

Convolution neural network (CNN) is a feedforward neural network including convolution calculation, which has good representation and learning ability [9]. As a special form of CNN, one-dimensional convolutional neural network can directly analyze and process one-dimensional signals. The input layer of one-dimensional convolution neural network receives one-dimensional data, and uses one-dimensional convolution kernel for convolution operation. After convolution and pooling, the final output result is also one-dimensional data.

A typical CNN model mainly includes two stages: convolution and classification. The convolution stage includes convolution layer and pooling layer, which is mainly used for data feature extraction. The classification stage includes full connection layer and classifier, which is mainly used for classification tasks.

2.1.1 Convolution Layer

The convolution layer performs convolution operation by convolution checking the local receptive domain of the input data, and each convolution kernel extracts the local features of the local receptive domain of the input data. The mathematical model can be expressed as:

$$x_i^l = \sum_{c=1}^{C} x_{i-1}^c \otimes w_i^{c,l} + b_i^l \tag{1}$$

where x_i^l is the l-th output characteristic diagram of the i-th convolution layer; C represents the number of channels of the input characteristic diagram; $w_i^{c,l}$ represents the weight of the l-th convolution kernel of the c-th channel of layer l; $\otimes$ represents convolution operation; b_i^l represents the offset of the l-th convolution kernel of the i-th convolution layer.

After the convolution operation is completed, the output as the convolution layer is activated by the nonlinear activation function [10]. Among them, relu activation function is the most commonly used activation function, and its formula is:

$$p_i^l = f(x_i^l) = \max\{0, x_i^l\} \tag{2}$$

where, x_i^l is the output characteristic diagram obtained after convolution operation of Eq. (1); p_i^l is the output value of x_i^l after relu activation.

2.1.2 Pooling Layer

Pooling layer, also known as the down-sampling layer, is used to reduce the dimension of data and extract important features at the same time [11]. Pooling operation mainly includes maximum pooling and average pooling, in which maximum pooling outputs the maximum value in the pool nuclear receptive domain and average pooling outputs the average value of elements in the pool nuclear receptive domain. In this paper, maximum pooling is selected in the network structure, and its formula is as follows:

$$q_i^l = \max_{(j-1)S+1<t<jS} \{p_i^l(t)\} \tag{3}$$

where: $p_i^l(t)$ is the output of the t-th neuron of the l-th output characteristic diagram of layer l; S is the size of pooled core; j is the step length; q_i^l is the output characteristic diagram after pooling operation.

2.1.3 Full Connection Layer

The function of full connection layer is to further extract the features extracted from convolution layer and pool layer [12]. Specifically, the output of the last layer of convolution layer and pooling layer is transformed into a one-dimensional array through flattening operation and input to the full connection layer. The formula is as follows:

$$y = f(w \cdot x + b) \tag{4}$$

where x is the input of the full connection layer, y is the output of the full connection layer, w and b are the weight and bias matrix of the neurons in the full connection layer, and $f(\cdot)$ is the activation function. Finally, the output of full connection layer is classified by softmax classifier.

2.1.4 Global Average Pooling Layer

The parameters of the whole connection layer account for 80% ~ 90% of the total parameters of the whole convolutional neural network, which greatly increases the computational resources of model training and is not conducive to real-time online monitoring [13]. Global average pooling (GAP) is the spatial average operation of the original fully connected inputs, which enhances the correlation between features and target categories. Without training parameters, it can greatly reduce the computing resources of the model and has stronger robustness. The formula is as follows:

$$y^c = \frac{1}{n}\sum_{i=1}^{n} x_i^c \tag{5}$$

where: x_i^c is the i-th eigenvalue of the c-th channel, n is the number of characteristics of each channel, y^c is the average value of the c-th channel of the input characteristic diagram.

2.2 Self Attention

The purpose of attention is to screen out a small amount of important information from the sequence, focus on these important information, give greater weight to important information, and ignore most unimportant information [14, 15]. Self-attention mechanism is a special case of attention mechanism, that is, the input of sequence feature extraction and state feature extraction is the same sequence. Firstly, assume that the feature sequence $x = [x_1, x_2, \ldots, x_n]$ of the input CNN, n is the sequence length, and then obtain the autocorrelation weight matrix $w = [w_1, w_2, \ldots, w_n]$ of the sequence. Finally, multiply the original feature sequence by the autocorrelation weight matrix to obtain the feature sequence x' filtered by self-attention. The formula is as follows:

$$w = \mathrm{softmax}(x \cdot x^T \cdot x) = [w_1, w_2, \cdots, w_n] \tag{6}$$

$$x' = x \cdot w \tag{7}$$

where: x' is the output sequence after self attention mechanism, x is the input sequence signal, and w is the autocorrelation weight matrix of the sequence signal.

2.3 Generative Adversarial Network

GAN mainly includes generator G and discriminator D, and its network structure is shown in Fig. 1.

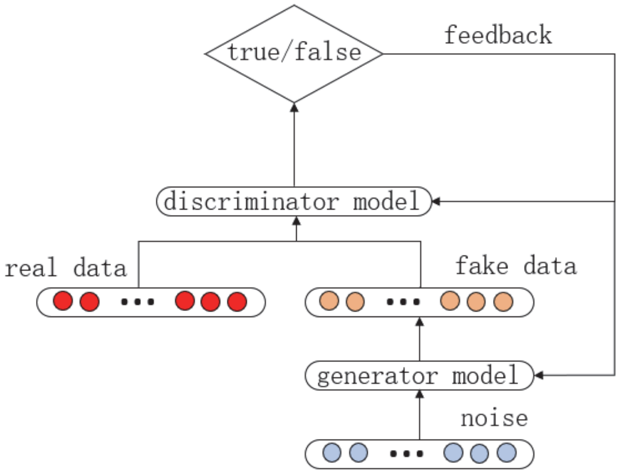

Figure 1 Schematic diagram of generation adversarial network structure

The task of the generator is to learn the distribution characteristics of real data and generate as real data as possible; the task of the discriminator is to distinguish the real data from the generated data as much as possible [16]. The generator and discriminator continuously optimize their performance through confrontation training, and finally achieve Nash balance [17]. The objective function formula is as follows:

$$\min_G \max_D V(G,D) = E_{x\sim P_r}[\ln D(x)] + E_{z\sim P_z}[\ln(1 - D(G(z)))] \tag{8}$$

where: Pr is the real sample distribution, Pz is the random noise distribution, $E(\cdot)$ represents the calculation expectation, $G(z)$ represents the sample generated by the generator, and $D(\cdot)$ represents the result output by the discriminator.

3 FAULT DIAGNOSIS METHOD BASED ON ICNN-1D

3.1 ICNN-1D Model Structure

Based on CNN, this paper proposes a circuit breaker fault diagnosis model (ICNN-1D) based on improved one-dimensional convolutional neural network. The specific model structure is shown in Fig. 2. Compared with traditional CNN, ICNN-1D mainly makes the following improvements:

(1) ICNN-1D directly takes the original opening and closing current signal of the circuit breaker as the input, and both convolution core and pool core adopt one-dimensional structure.

(2) ICNN-1D uses the self-attention mechanism layer to adaptively weight the input signal, highlight the weight of important information and suppress the weight of unimportant information, so that the network can obtain more important discrimination features.

(3) In the last convolution layer, 1 × 1 convolution kernel is used to compress the feature channel and retain the significant features, and gap layer is used to replace the full connection layer, which greatly reduces the parameters of the model and improves the training efficiency of the model. Since the gap layer output is connected with the softmax classifier, the number of convolution cores of the last layer 1 × 1 convolution layer is set as the number of fault categories.

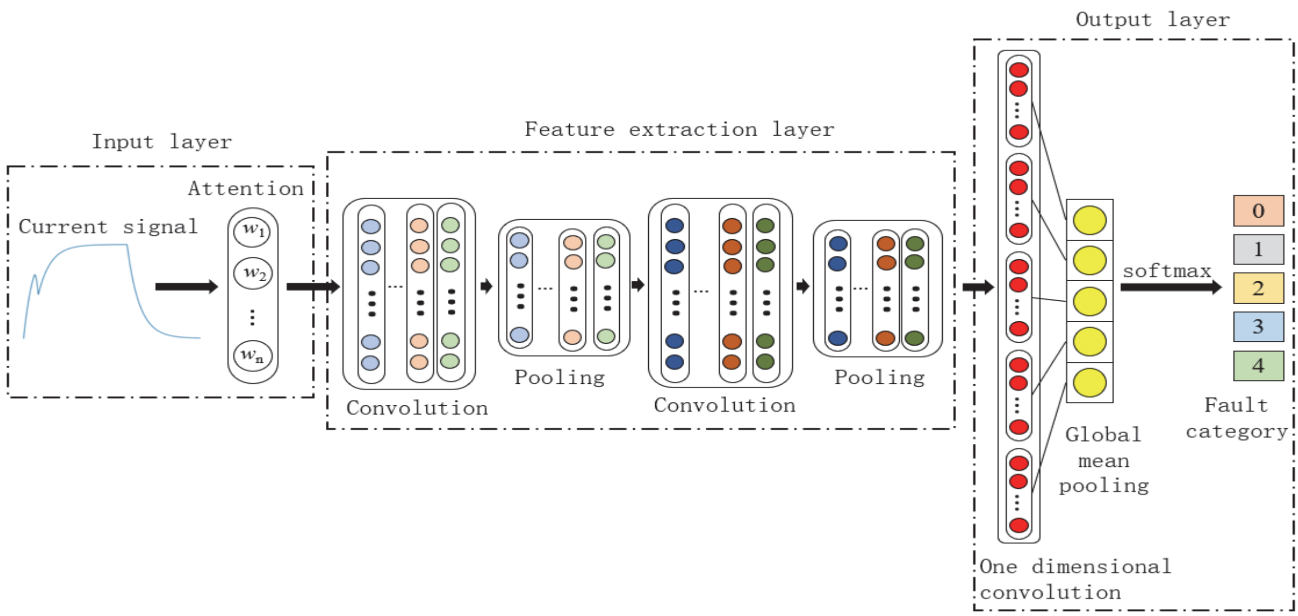

Figure 2 ICNN-1D network structure

3.2 Process Design of ICNN-1D Method

The fault diagnosis process based on ICNN-1D is shown in Fig. 3, which is mainly divided into three links: 1) data set preparation; 2) Model training; 3) Model testing. The specific steps are described as follows:

Step 1: The collected current signal of circuit breaker opening and closing coil is randomly divided into training set and verification set.

Step 2: The network structure of the model is designed and the model parameters are initialized. Taking the cross entropy as the loss function, the parameters of the model are updated through repeated forward propagation and back propagation until the training requirements are met and the optimal model is saved.

Step 3: The trained model is used to test the test set to verify its diagnostic performance.

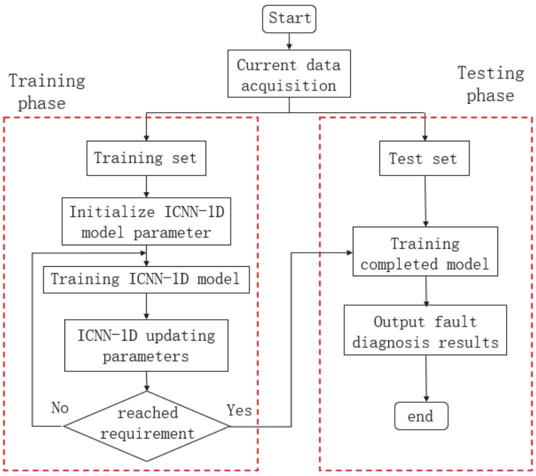

Figure 3 ICNN-1D fault diagnosis process

4 EXPERIMENTAL VERIFICATION

4.1 Data Set Introduction

Relevant research shows that during the opening and closing process of low-voltage circuit breaker, the opening and closing coil current is not only easy to detect, but also can well reflect the change of mechanical state of low-voltage circuit breaker [18]. Therefore, the opening and closing coil current can be used as the basis for fault diagnosis of low-voltage circuit breaker.

Taking "CW1-1600" low-voltage universal circuit breaker as the experimental object, this paper uses NI USB-6002 data acquisition card to add sample at 10 kHz, and uses "CMS050NPT" Hall current sensor to collect the opening and closing coil current. 100 groups (500 groups in total) of opening and closing coil current data under 5 states are collected, and each waveform contains 2048 sampling points. The typical current waveform of each state is shown in Fig. 4 (0-normal operation; 1-Low action voltage; 2-aging of closing coil; 3-excessive empty stroke of closing iron core; 4-jamming of iron core).

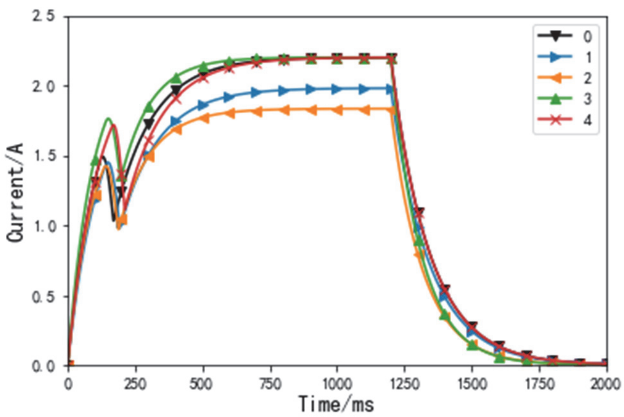

Figure 4 Typical current diagram of each state of opening and closing coil

4.2 Data Enhancement

In this paper, the original data samples are enhanced by generating countermeasure network. Its network structure includes four full connection layers, the activation function is relu activation function, and the loss function adopts MSE loss function. In the training process, the generator will gradually generate realistic sample data according to the feedback of the discriminator, and the discriminator will gradually be unable to judge whether the data is real data or generated data, so the discriminator will output the discrimination result of 0.5. Because we use the mean square error, the final loss will gradually converge to 0.25. The training loss curve of the discriminator is shown in Fig. 5.

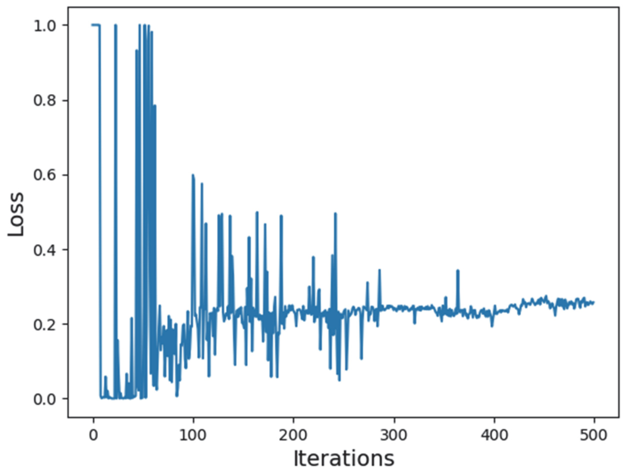

Figure 5 Discriminator training loss curve

After the training generator is obtained, we will continuously generate realistic data by continuously inputting random noise. By adding these data to the original data set, we can train the CNN model. Its network structure includes two layers of convolution layer and pooling layer, and the training iteration is 100 times. We use CNN to train the original data set and the data set with 50%, 100%, 200%, 300% and 400% data enhancement. The training diagnosis accuracy and training time are shown in Tab.1. The fault diagnosis accuracy of CNN model trained by the original data set is 84.6%. On this basis, by adding 50%, 100% and 200%, 300%, 400% data enhanced samples for training, the accuracy of CNN model fault diagnosis has been improved by 0.5%, 0.9%, 2.1%, 2.4% and 2.6% respectively. It can be found that with the increase of data samples, CNN can learn features from more samples, better extract data features, and the accuracy of model fault diagnosis continues to improve. However, with the increase of samples, the training time will be greatly increased, and the accuracy will not be significantly improved. Therefore, 200% data enhancement is the most appropriate choice in this paper.

Table 1 Diagnostic accuracy and training time of CNN in different data sets

Data sets	Diagnostic accuracy / %	Training time / ms
Original	84.6	231
+50%	85.1	278
+100%	85.5	321
+200%	86.7	369
+300%	87.0	512

4.3 ICNN-1D Model Parameter

In this paper, the input signal of the model is a one-dimensional current signal with low dimension. Considering that the deep convolution network will lead to overfitting of the model, this paper adopts a two-layer convolution network structure. Firstly, the network adaptively weights the current signal through the self-attention layer, and the process does not need training parameters; then, the features are extracted through the alternating training of two-layer convolution layer and pool layer. Finally, the full connection layer is replaced by 1 × 1 convolution layer and gap layer. Finally, the classification results are output through softmax classifier. The specific network parameters are shown in Tab. 2:

Table 2 Diagnostic accuracy and training time of CNN in different data sets

Layer	Size × step size	Number of channels	Output dimension
Input layer	-	-	2048 × 1
Self-attention	2048 × 1	-	2048 × 1
Convolution layer 1	8 × 3	8	683 × 8
Pooling layer 1	3 × 3	8	227 × 8
Convolution layer 2	3 × 3	16	76 × 16
Pooling layer 2	3 × 3	16	25 × 16
1 × 1 Convolution layer	1 × 1	5	25 × 5
GAP layer	-	-	5
Softmax	-	-	5

4.4 Comparative Analysis of Models

This experiment is based on the network model built by keras development environment. The batch processing capacity is set to 64, the learning rate is set to 0.00001, and the number of iterations is set to 100. Adam optimization algorithm is used for optimization. The proportion of training set and verification set is divided into 7:3. In order to verify the fault diagnosis accuracy and efficiency of ICNN-1D, the number of network training parameters is analyzed and compared with the traditional CNN method, as shown in Tab. 3:

Table 3 Comparison of trainable parameters between traditional CNN and ICNN-1D network models

Layer	Traditional CNN	ICNN-1D
Input layer	-	-
Self-attention	-	0
Convolution layer 1	72	72
Pooling layer 1	0	0
Convolution layer 2	400	400
Pooling layer 2	0	0
Convolution layer 3	-	85
Full connection layer	2005	-
GAP layer	-	0
Total parameters	2477	557

It can be seen from Tab. 3 that the total training parameters of traditional CNN using full connection layer network are 2477, while the total training parameters of this method using gap instead of full connection layer network are 557, which reduces 77% of the training parameters compared with traditional CNN, greatly saves computing resources and improves the efficiency of model training.

Fig. 6 shows the change of the accuracy of training set and verification set with the number of iterations in the training process of traditional CNN and ICNN-1D. As can be seen from Fig. 6: when the iteration cycle of traditional CNN is only 50 times, the training set and verification set are not fully convergent, and the accuracy of the training set is only 86.7%, which does not reach the best. And the accuracy of the validation set does not converge to the test set, indicating that the model may have poor generalization ability or over fitting problems. After 12 cycles of iteration, the accuracy of ICNN-1D training set has reached the peak of 99.2%, and the accuracy of training set and test set also tends to converge, indicating that the model has good fast learning ability and good generalization ability.

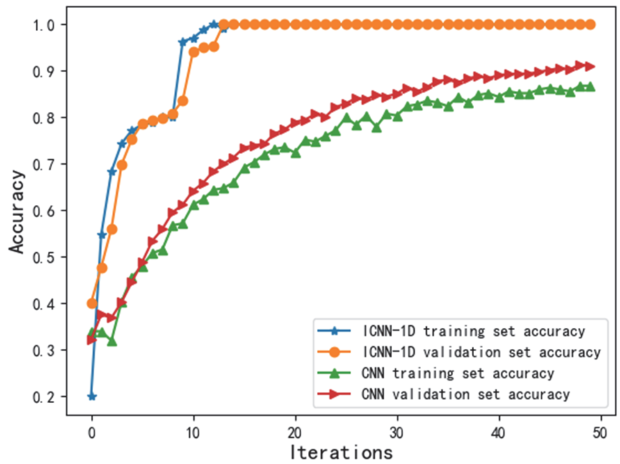

Figure 6 ICNN-1D and CNN training process

In order to more intuitively reflect the classification of each type of fault by icdd-1d, the diagnosis results and corresponding misclassification rate of each type of fault are displayed by introducing confusion matrix. The definition of confusion matrix of classification results is shown in Tab. 4.

Table 4 Confusion matrix of classification results

confusion matrix		Predict	
		Positive	Negative
Real	Positive	TP (True Positive)	FN (False Negative)
	Negative	FP (False Positive)	TN (True Negative)

The confusion matrix of traditional CNN and ICNN-1D is shown in Fig. 7. The left figure is the confusion matrix of traditional CNN diagnosis results, and the right figure is the confusion matrix of ICNN-1D diagnosis results. The horizontal axis represents the predicted fault type number, the vertical axis represents the real fault type number, and the diagonal represents the classification accuracy, and outside the diagonal represents the misclassification rate. It can be seen from the figure that the traditional CNN and ICNN-1D have reached 100% recognition accuracy for fault 0, fault 1 and fault 2, but the traditional CNN has a high misjudgment rate for fault 3 and fault 4, which are 14% and 12% respectively; ICNN-1D can achieve 100% recognition accuracy for fault 4, and the misjudgment rate of fault 3 is only 4%. It can be seen that ICNN-1D has higher fault identification accuracy.

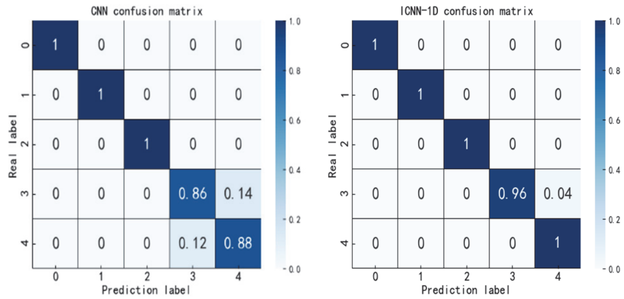

Figure 7 Confusion matrix of CNN and ICNN-1D

In order to further verify the advantages of this method, this paper compares the proposed method with traditional CNN, support vector machine (SVM) and BP neural network. The parameters of ICNN-1D network model are shown in Tab. 1. Traditional CNN includes two convolution layers, two pooling layers and one full connection layer. The size of convolution kernel is 1 × 8 and 1 × 3 respectively, the depth is 8 and 16 respectively, the size of pooling kernel is 1 × 2, the step size is 3, and the iteration cycle is 100 times. SVM adopts Gaussian kernel function. The number of units in the input layer, hidden layer and output layer of BP neural network is 500, 100 and 5 respectively, and the activation function is sigmoid function.

In this paper, the diagnostic performance of each model is evaluated by Accuracy, Recall and F1 score. Its definitions are as follows:

$$Accurancy = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

$$Recall = \frac{TP}{TP + FN} \tag{10}$$

$$F1\ score = \frac{2 * TP}{2TP + FP + FN} \tag{11}$$

Table 5 Comparison between ICNN-1D and other models

Model	*Accuracy* / %	*Recall* / %	*F1 score* / %
ICNN-1D	99.2	99.4	99.2
CNN	86.7	86.8	86.3
SVM	82.1	85.4	83.4
BP	79.5	84.3	80.1

The diagnosis comparison results of each model are shown in Tab. 5. It can be seen from Tab. 5 that the fault diagnosis method based on deep learning has obvious advantages over the traditional fault diagnosis methods (SVM, BP). Because SVM and BP are shallow network models and only learn shallow features, their accuracy is slightly insufficient, which are 82.1% and 79.5% respectively. The method proposed in this paper is a deep learning model, which can learn deep features. Because the self-attention mechanism is added to give higher weight to the important features of the original signal, the feature extraction speed of convolution layer is faster and the accuracy is higher. Therefore, compared with the traditional CNN, the fault diagnosis rate is improved by nearly 12.5%. In terms of Recall, compared with other models, ICNN-1D also performs best, which shows that the model has satisfactory accuracy in predicting a few types of faults. It can also be seen from F1 score that the performance effect of ICNN-1D is also the best, which shows that the model has good robustness. Overall, the proposed method has good fault diagnosis accuracy and robustness.

5 CONCLUSION

In this paper, a circuit breaker fault diagnosis method based on improved one-dimensional convolution neural network is proposed. The method improves one-dimensional convolution by using self-attention and using 1 × 1 convolution and global mean pooling layer instead of full connection layer, and enhances data through GAN. Through the research of this paper, the following conclusions can be obtained:

1) One dimensional convolution can directly process one-dimensional current signal of circuit breaker without manual feature extraction.
2) Through the self-attention, the original feature sequence is adaptively weighted to highlight important information, so that the training convergence is faster and the accuracy is higher.
3) Using 1 × 1 convolution and global mean pooling layer instead of full connection layer greatly reduces the parameters and calculation of the model and prevents over fitting.
4) Data enhancement through GAN enables the model to learn more features and further improves the accuracy of model diagnosis.

The experimental results show that the proposed method has faster convergence speed and higher fault recognition rate than other intelligent fault diagnosis models, and has certain engineering application value.

6 REFERENCES

[1] Shuguang, S., Wei, Z., Jingqin, W., Taihang, D., & Hui, G. (2020). Mechanical life prediction of low-voltage circuit breaker based on vibration detection during operation. *Chinese Journal of Scientific Instrument*, *41*(12), 146-157. https://doi.org/10.19650/j.cnki.cjsi.J2006940
[2] Xinbo, H., Xiaowen, H., & Yongcan, Z. (2018). Fault diagnosis of high-voltage circuit breaker based on convolution neural network. *Electric Power Automation Equipment*, *38*(05), 136-140+147. https://doi.org/10.16081/j.issn.1006-6047.2018.05.020
[3] Wei, X., Bo, Y., & Yu, L. (2020). Fault Diagnosis of High Voltage Circuit Breaker by Fusing IMBSE with AFS-DBN Model. *Modular Machine Tool & Automatic Manufacturing Technique*, (11), 68-73+81. https://doi.org/10.13462/j.cnki.mmtamt.2020.11.016
[4] Chuan, L. & Qifeng, X. (2021). Fault diagnosis of circuit breaker based on bispectrum and two-stream convolutional neural network. *Electronic Measurment Technology*, *44*(23), 165-172. https://doi.org/10.19651/j.cnki.emt.2107997
[5] Shuguang, S., Li, Q., & Taihang, D. (2020). Fault Diagnosis of Accessories for the Low Voltage Conventional Circuit Breaker Based on One-Dimensional Convolutional Neural Network. *Transactions of China Electrotechnical Society*, *35*(12), 2562-2573.

https://doi.org/10.19595/j.cnki.1000-6753.tces.190453

[6] Renwu, Y., Chuan L., & Weilang S. (2022). Research on Circuit Breaker Fault Diagnosis Based on EEMD and Convolutional Neural Network. *High Voltage Apparatus*, *58*(04), 213-220. https://doi.org/10.13296/j.1001-1609.hva.2022.04.029

[7] Qiuyu, Y., Jiangjun R., Zhijian, Z., & Daochun, H. (2019). Condition Evaluation for Opening Damper of Spring Operated High-Voltage Circuit Breaker Using Vibration Time-Frequency Image. *IEEE Sensors Journal*, *19*(18), 8116-8126. https://doi.org/10.1109/jsen.2019.2918335

[8] Yupeng, C., Lin, L., Qiao, W., & Jianliang, Z. (2021). Fault diagnosis of high-voltage vacuum circuit breaker with a convolutional deep network. *Power System Protection and Control*, *49*(03), 39-47. https://doi.org/10.19783/j.cnki.pspc.200335

[9] Qi, W., Linfeng, D., & Rongzhen, Z. (2022). Fault recognition of rolling bearing based on improved 1D convolutional neural network. *Journal of Vibration and Shock*, *41*(03), 216-223. https://doi.org/10.13465/j.cnki.jvs.2022.03.026

[10] Yan, X., Ning, S., & Dongyang, D. (2021). Fault Degree Diagnosis of Rolling Bearings Based on One-Dimensional Convolutional Neural Network. *Bearing*, (04), 48-54. https://doi.org/10.19533/j.issn1000-3762.2021.04.011

[11] Liling, M., Xiaoran, L., Wei, S., & Junzheng, W. (2020). Motor Fault Diagnosis Method Based on an Improved One-Dimensional Convolutional Neural Network. *Transactions of Beijing Institute of Technology*, *40*(10), 1088-1093. https://doi.org/10.15918/j.tbit1001-0645.2019.201

[12] Dengping, T., Wenjia, C., & Li, Z. (2021). Application of Improved Convolution Neural Network in Fault Diagnosis of Transformer. *Computer Engineering and Applications*, *57*(11), 239-247.

[13] Wenfeng, G., Hui, C., & Zehui, Z. (2020). Intelligent fault diagnosis for rolling bearing based on improved convolutional neural network. *Journal of Vibration Engineering*, *33*(02), 400-413. https://doi.org/10.16385/j.cnki.issn.1004-4523.2020.02.021

[14] Liang, M., Chao, Y., & Xixiang, Z. (2020). Non-intrusive load disaggregation based on ID convolutional neural network and self-attention mechanism. *Power Systems and Big Data*, *23*(10), 1-8. https://doi.org/10.19317/j.cnki.1008-083x.2020.10.001

[15] Yongcan, Y., Tao, L., & Xiaoqin, L. (2021).Fault Diagnosis of Planetary Gearbox Based on One dimensional Convolutional Neural Network with Attention Mechanism. *Machinery & Electronics*, *39*(10), 3-8.

[16] Dongdong, L., Yuhang, L., & Yao, Z. (2021). Fault Diagnosis Method of Wind Turbine Planetary Gearbox Based on Improved Generative Adversarial Network. *Proceedings of the CSEE*, *41*(21), 7496-7507. https://doi.org/10.13334/j.0258-8013.pcsee.202223

[17] Yunpeng, L., Ziqiang, X., & Jiahui, H. (2020). Data Augmentation Method for Power Transformer Fault Diagnosis Based on Conditional Wasserstein Generative Adversarial Network. *Power System Technology*, *44*(04), 1505-1513. https://doi.org/10.13335/j.1000-3673.pst.2019.2416

[18] Tianze, W., Erdong, C., & Xiaojun, J. (2021). Design of Circuit Breaker Opening and Closing Coil Current Signal and Characteristic Point Acquisition System. *Instrument Technique and Sensor*, (10), 47-51+64.

Contact information:

Jie SHI, Research Fellow
School of Electronics and Information Engineering,
Suzhou University of Science and Technology,
No. 1, Kerui Road, Huqiu District, Suzhou 215000, China
E-mail: 17751455752@163.com

Guoqing DU, Master
(Corresponding author)
School of Electronics and Information Engineering,
Suzhou University of Science and Technology,
No. 1, Kerui Road, Huqiu District, Suzhou 215000, China
E-mail: 1103403082@qq.com

Haifeng SHENG, Master
School of Electronics and Information Engineering,
Suzhou University of Science and Technology,
No. 1, Kerui Road, Huqiu District, Suzhou 215000, China
E-mail: 1195534012@qq.com

Fei DING, Master
School of Electronics and Information Engineering,
Suzhou University of Science and Technology,
No. 1, Kerui Road, Huqiu District, Suzhou 215000, China
E-mail: 953687854@qq.com

Weixiang KONG, Master
School of Electronics and Information Engineering,
Suzhou University of Science and Technology,
No. 1, Kerui Road, Huqiu District, Suzhou 215000, China
E-mail: 1229682452@qq.com