Industrial Engineering Journal
ISSN: 0970-2555
Volume : 51, No. 2, JULY - DEC 2022

superior Inception-ResNet Model for Graph Semantic withdrawal in Power Grid Fault Diagnosis

P.Rajendra[1], D.Aruna Kumari[2], N.Aparna[3], B.Jyothirmai[4],T.Divya[5],M.Srilatha [6]

[1,2,3,4,5,6]Assistant Professor, Department of ECE, Sri Indu Institute of Engineering & Technology, Hyderabad

Abstract: ***Traditional power grid fault diagnosis methods have problems such as large parameters, poor real-time performance, susceptibility to malformed data interference, and low accuracy. In order to solve the above problems, a power grid fault diagnosis method based on improved Inception-ResNet graph semantic extraction is proposed. Real-time monitoring of each node of the power grid based on PMU measurement data. The semantic reconstruction of PMU data is realized by converting data features into image features, thus completing the conversion from quantitative analysis of data form to qualitative analysis of image form, greatly reducing the influence of malformed data on fault diagnosis. The improved Inception-ResNet algorithm is used to extract the semantic features of the fault image, and the corresponding fault type is obtained according to the fault feature and the fault diagnosis result is output. Experiments and simulations show that this method can effectively reduce training parameters and training time, shorten diagnosis time, and improve diagnosis accuracy.***

Keywords: PMU Graph Data, Power Grid Fault Diagnosis, Deep Learning, SE-Inception-ResNet

1. Introduction

With the rapid development of national economy, users have higher and higher requirements for the reliability of power supply, while the scale of distribution network is expanding, and the power grid topology is becoming more and more complex (Cloud Architecture), which puts forward higher requirements for power grid fault diagnosis. The rapid development of artificial intelligence has provided many new and more effective methods for power grid fault diagnosis. How to use deep learning method to improve the speed and accuracy of fault diagnosis has become the focus of many experts and researchers at home and abroad.

Although the construction of power system at home and abroad continues to advance, various faults will still occur in the power system due to the existence of uncertain factors such as power system itself, natural climate change and human factors. When a fault occurs, if the fault information is not processed in time, the faulty equipment is not judged accurately and the decision-making is not timely, it is likely to cause cascading faults, power grid instability or even collapse, resulting in incalculable losses. For example, the power system blackout in the United States and Canada in 2003, and the blizzard disaster in southern China in 2008 also caused large-scale power system failures. These accidents have seriously affected the national industrial development and people's daily life. In addition, with the improvement of power grid informatization, faults caused by network and data attacks sometimes occur [1]. Since the 1970s, the research on related technologies for power grid fault diagnosis has become the focus of many experts and scholars at home and abroad. At present, most of the methods in power grid fault diagnosis use switching value information such as protection action and circuit breaker trip, but the defect of switching value information is that there are inevitable uncertain factors such as refusal of action, misoperation and loss of information under channel interference, as well as multiple protection actions and circuit breaker tripping in case of complex fault. Fault diagnosis based on switching value information is difficult to solve all fault problems [2]. The existing fault diagnosis methods have made a lot of improvements in dealing with inaccurate and incomplete fault information. Although the algorithm has been improved to a certain extent and the influence of uncertain factors on fault diagnosis has been reduced, the accuracy of diagnosis results is still difficult to meet the requirements and needs to be further improved.

In recent years, the Wide Area Measurement System[3-4] (WAMS) based on the synchronous Phasor Measurement Unit (PMU) has gradually developed at home and abroad, and WAMS is more and more widely used in power system [5]. WAMS synchronizes the phase angle and other main data of the whole network by configuring PMU devices at key points in the whole network of the power system. Since the 1990s, the real-time section data of the power grid with a unified time scale provided by PMU has been gradually applied to the power grid fault diagnosis system. At this stage, the fault diagnosis is mainly based on the analysis of action information of protection and switch after the fault occurs by

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 51, No. 2, JULY - DEC 2022

artificial intelligence technology, the analysis of fault recorder data and PMU electrical quantity data by mathematical analysis method, and the integration of various information sources for comprehensive analysis. [6] At present, there are many mature power grid fault diagnosis methods, mainly including expert system, Petri network, genetic algorithm, artificial neural network and fuzzy set theory, etc. [7]

According to the advantages and disadvantages of these diagnostic methods and the shortcomings of existing neural networks in power network fault diagnosis, this paper presents a power network fault diagnosis method based on inception and residual neural networks. Convolution neural network based on deep learning field uses inception and residual neural network, changes the structure of traditional convolution neural network, enables the model to learn more effective data characteristics, and adds a shortcut to the input so that it can be directly transferred to the deep network layer, alleviating the problem of gradient disappearance when the number of network layers increases. Then it extracts features and reduces dimensionality for high-dimensional PMU time series measurement data, converts the obtained time into corresponding pictures, and converts original discrete data into continuous information. Finally, it obtains the fault type with Softmax. The experimental comparison shows that the dimension of the data is reduced, while the time series data features are fully expressed. At the same time, the model has good robustness and anti-noise performance. It can effectively improve the classification efficiency, reduce the complexity of the model, make the classification results more accurate, reasonable and stable, and provide an effective reference for the dispatch center fault diagnosis.

2. Fault Diagnosis method based on CNN Graphic Semantic Extraction

2.1 Inception module

Currently, image classification algorithms based on deep learning mainly include AlexNet, VGG, ResNet, GoogleNet, etc. These models usually have slow convergence speed, too many training parameters, and problems such as gradient dispersion, which are difficult to optimize the model [8].

Inception is a deep learning network based on the CNN network, which has a convolutional layer, a pooling layer, a fully connected layer, a dropout layer and an output layer. Inception mainly solves the problem of excessive stacking parameters of multiple convolutional layers. The network designs a network with excellent local topology, that is, performs multiple convolution operations or pooling operations on the input image in parallel, and stitches all output results. It is a very deep feature map.

Inception v1 stacks the convolution layer whose convolution kernel is 1×1, 3×3, 5×5 and the convolution kernel whose pooling operation is 3 × 3 together. Meanwhile, 1 × 1 convolution kernel is added respectively between the convolution layer whose convolution kernel is 3×3 and the pooling layer whose convolution kernel is 5×5 for dimension reduction [9]. This Inception structure can not only increase the width of the network, but also increase the adaptability of the network to inputs of different scales. Inception v2 highlights the improvement of changing the 5×5 convolution kernel to two 3×3 convolution kernels. On the premise of ensuring the same effect, the use of convolutiondecomposition can reduce the number of parameters and speed up the calculation [10]. Inception v3 puts forward the decomposition idea, divides the n×n convolution kernel into two n×1,1×n convolution kernel, which further deepens the network and increases the network nonlinearity. Inception v4 further improves performance by adding a "residual" structure to the network.

Although the inception structure proposed initially increased the width of the network and reduced the parameters of the network, there were still many parameters in the calculation of 5x5 convolution kernel, which still had a great influence on the calculation. Therefore, a new Inception structure was proposed, as shown in Figure 2.1-2. In the improved structure, 1x1 convolution kernel is added in front of 3x3 and 5x5 convolution kernel to realize the effect of dimension reduction, reduce the calculation parameters, and deepen the depth of the network in another way and improve the performance of the network.

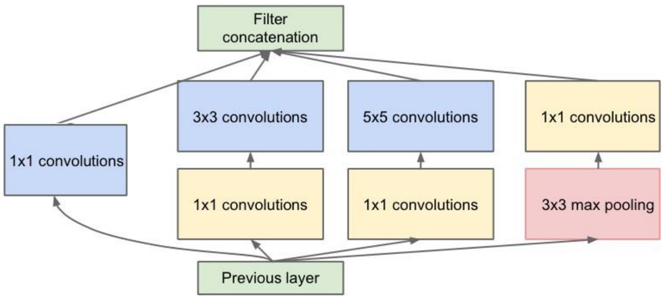


Figure 2.1-2: Inception improvement diagram

2.2 Resnet

The increase in the number of layers of traditional neural networks will cause problems such as overfitting and gradient descent. Using regularization to train a large amount of data can alleviate the problem of overfitting. Using standardization can solve the problem of the disappearance of the gradient of the network with about ten layers, but the number of network layers continues to increase, and the accuracy of the training set will no longer continue to improve. This is because the problem of gradient disappearance has occurred. In response to this problem, He [11] and others applied identity mapping to neural networks to realize layer-jumping connections, thereby alleviating the problem of gradient disappearance. This is the resent network. The main idea of the residual network is to transmit a part of the input directly to the deeper network layer, add it to the output after convolution, and then do ReLU. This alleviates the decrease in accuracy caused by the increase in the number of network layers due to the disappearance of the gradient [12].

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 51, No. 2, JULY - DEC 2022

The most important part of resent is fast residual, which generally has two forms. One is composed of serial connection of two 3x3 convolution layers, and the other is composed of serial connection of three convolution layers of1x1,3x3 and 1x1[13]. The structural diagram is shown

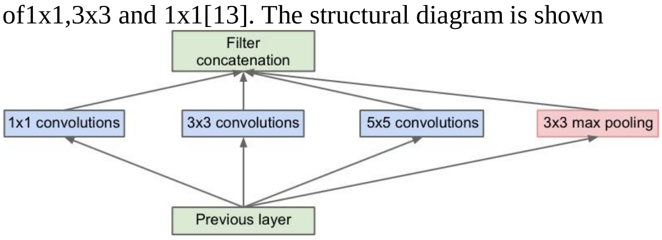


Figure 2.1-1: Inception initial structure

in FIG. 2.2-1 and FIG. 2.2-2. The most important function of the three-layer structure shown in Figure 2.2-2 is to reduce the parameters and reduce the complexity of network layer computing.

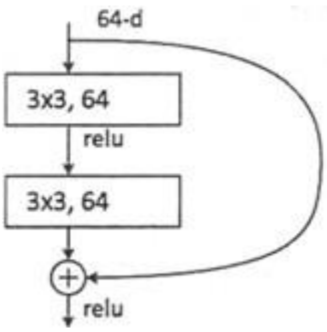


Figure 2.2-1: Two-layer residual block structure

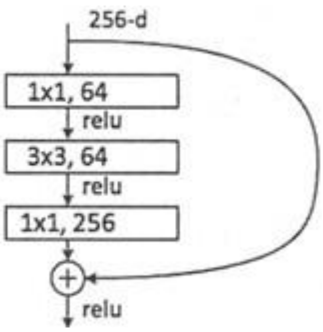


Figure 2.2-2: Three-layer residual block structure

2.3 Power grid fault diagnosis model based on improved Inception-Resent graph semantic extraction

2.3.1 Improved inception-resnet v2 network

In this paper, an improvement is made on the Inception-ResNet V2 network. Based on the original Stem, Inception-ResNet, Reduction, The SE-Block is combined with Inception-ResNet to form the SE-Inception-ResNet layer. For CNN network, convolution operation is the core operator. It learns new feature maps from input feature maps and performs feature fusion for a local region. In order to integrate more effective features in space, the Inception network uses multiple 1× N and N ×1 small-size convolution to replace large-size convolution kernel [14]. Meanwhile, multi-branch structure is used to effectively perform a large number of convolution decomposition in the feature space, which not only deepens the network depth, but also speeds up the network speed. However, as the depth of Inception network deepens, the risk of gradient disappearance and overfitting also increases. Therefore, the residual blocks are added to the Inception network to form Inception ResNet to avoid the phenomenon of network degradation. In addition, all channelsof input features are directly fused in Inception network, so the importance of different channel features cannot be learned.

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} u_c(i, j), z \in R^C \tag{2.3.1-1}$$

To solve this problem, this paper proposes to introduce SE Block into the structure of Inception ResNet. The SE (Sequeeze-and-Congestion) module first performs a Squeeze operation on the feature map obtained by convolution to obtain channel-level global features, then performs an Excitation operation on the global features, learns the relationship between each channel, and finally combines the original feature map to obtain a new characteristic map[15]. In order to balance calculation and efficiency, the original Inception-ResNet layer is deleted and modified, and the Inception-ResNet is modified to a combination of Inception-ResNet and SE-Inception-ResNet. Meanwhile, the network layer number of modules is adjusted appropriately. Figure 2.3.1-1 shows the improved network structure.

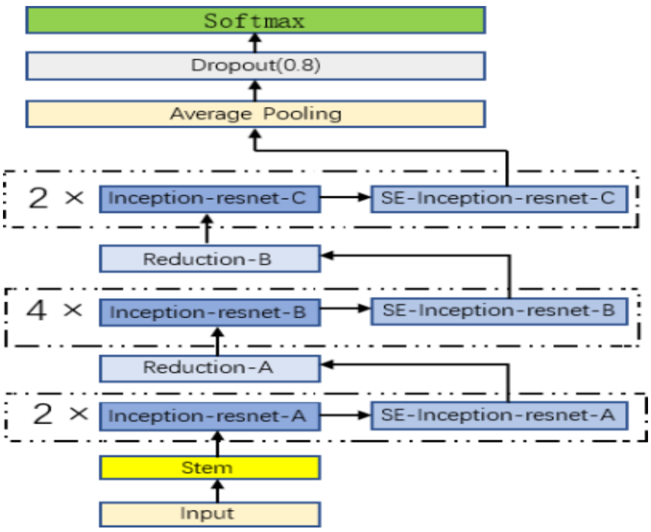


Figure 2.3.1-1: Improved Inception-Resnet v2 network

As shown in Figure 2.3.1-2, the left picture is the Inception-ResNet module, and the right picture is the SE-Inception-ResNet module. SE-Block mainly includes two operations: Sequence and Excitation. The Sequence operationuses global average pooling to encode the entire feature space on the channel as a global feature. The calculation is shown in equation (2.3.1-1):

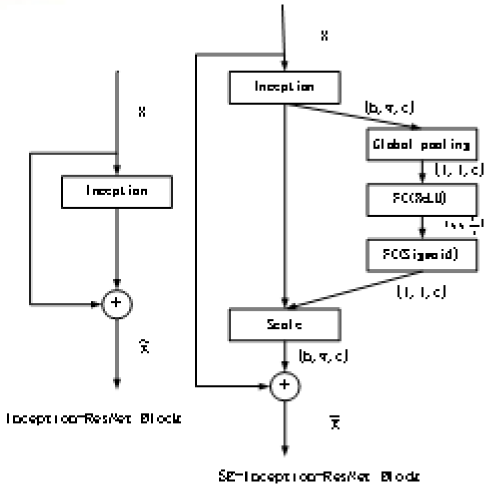


Figure 2.3.1-2: SE-Inception-Resnet network structure

Where u_c represents the spatial feature of the Inception module on the channel, H represents the height of the feature space, W represents the width of the feature space, and C represents the number of channels in the feature space.

Then the extracted global features are operated to extract the relationships between the channels. In order to reduce the complexity of the model and to extract the generalization capability, two full connection layers are used here. The first full connection layer is to reduce the dimension. And then activated by the ReLU function. Finally, the second full connection layer is used to restore the initial dimension and the weight coefficient of each channel is obtained by the sigmoid activation function. The calculation is shown in Formula (2.3.1-2):

Where z is the feature space after global average pooling, $W_1 \in R^{\frac{C}{r}\times C}$, $W_2 = R^{C\times\frac{C}{r}}$, r is the dimensionality reduction coefficient, a is the Sigmoid activation function, and β is the ReLU activation function.

Finally, the learned weight coefficients of each channel are fused with the feature map of the Inception module, and the calculation is shown in Formula (2.3.1-3):

Among them, s_c is the weight coefficient of SE module, and u_c is the spatial feature of Inception module on the channel.

In order to achieve a good balance between model performance, complexity and computational efficiency, this paper combines Inception-ResNet and SE-Inception-ResNet, and correspondingly reduces the depth of the model to achieve the purpose of reducing the computational amount and increasing the computational efficiency.

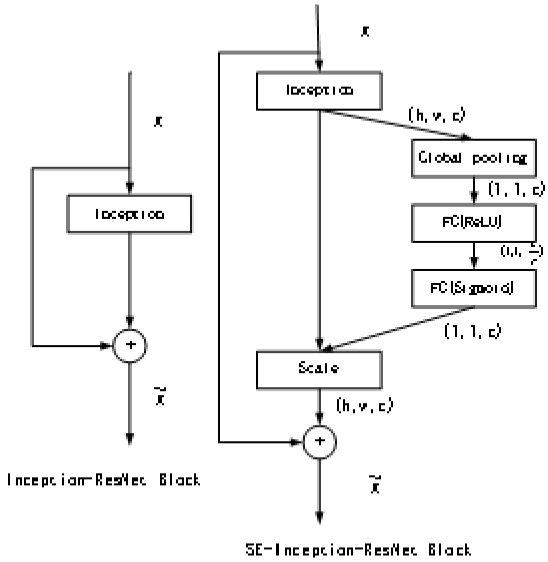


Figure 2.3.1-3: SE-Inception-Resnet network structure

2.3.2 Algorithm flow

As shown in Figure 2.3.2-1, this paper proposes a power grid fault diagnosis method based on improved Inception-Resnet graphical semantic extraction. Firstly, the model converts the PMU samples of the faulty device from discrete data features to continuous image features. Secondly, the converted radar map samples are trained with improved Inception-ResNet to obtain a fault classification model for identifying simple and complex faults.

$$s = F_{ex}(z,W) = \sigma\left(g(z,W)\right) = \sigma\left(W_2 ReLU(W_1 z)\right) \tag{2.3.1-2}$$

$$\tilde{x}c = Fscale(u_c, s_c) = s_c \cdot u_c \tag{2.3.1-3}$$

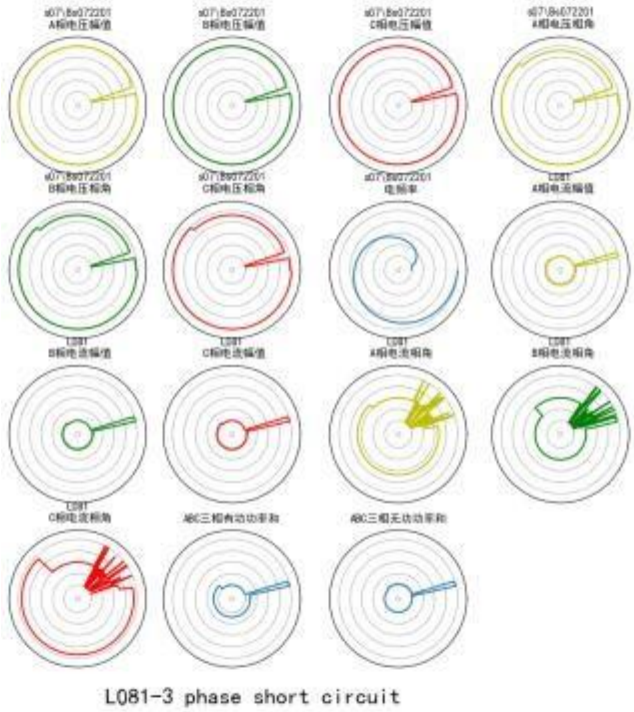
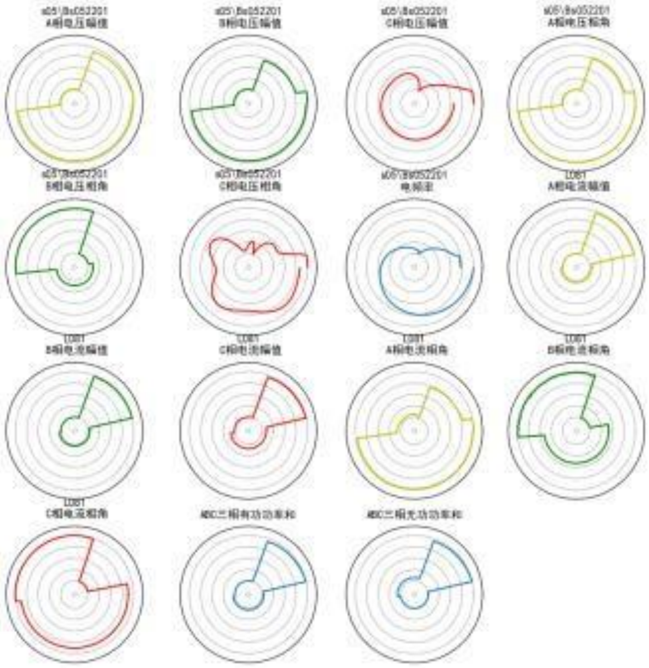


Figure 3.1-1(a): Simple fault

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 51, No. 2, JULY - DEC 2022



L081-Single phase earth fault

Figure 3.1-1(b): Complex fault

Figure 3.1-1: Example of simple and complex faults

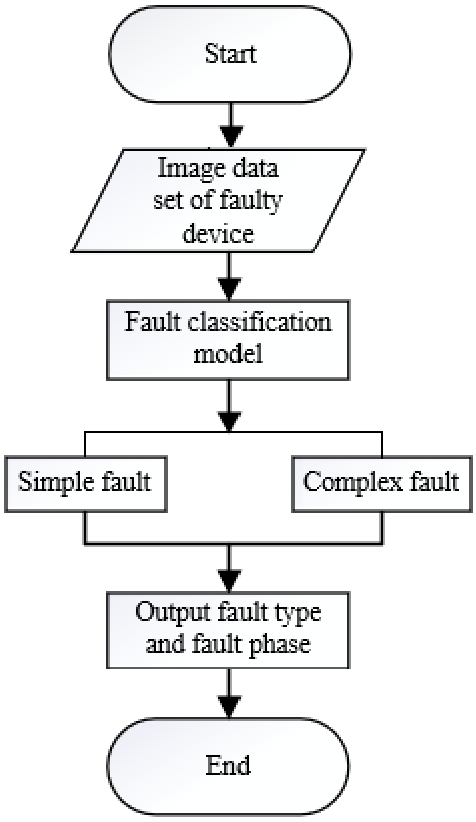


Figure 2.3.2-1: Overall flowchart of power grid faultdiagnosis

3. Experimental Research

3.1 Data processing

According to the judgment requirements of fault type, the collected PMU data shall be classified into training sets on the basic of fault complexity. In the light of the fault complexity, the PMU data samples are divided into simple fault, complex fault and misoperation fault. The sample composition is shown in table 3.1-1.

Table 3.1-1: Sample composition of fault complexity judgment

Fault Complexity	Simple Fault	Complex Fault
Number of Samples	5000	5000

Select the A, B, C three-phase voltage, the amplitude and phase angle of three-phase current, electrical frequency, active power and reactive power of the fault equipment, as well as the PMU data radar diagram of each device. Sample examples of simple and complex faults are shown in Figure 3.1-1

3.2 Experiment and result analysis

3.2.1 Experimental parameter setting and analysis

The paper collects the fault data of instantaneous fault, permanent fault, single-phase short-circuit grounding fault, two-phase short-circuit grounding fault, two-phase short-circuit fault, three-phase short-circuit fault and refusal fault from IEEE39 model, and the power grid fault diagnosis experiment based on improved Inception-Resnet image semantic extraction.

Among them, 6000 single-phase short-circuit grounding fault samples, two-phase short-circuit grounding fault samples, two-phase short-circuit fault samples and three-phase short-circuit fault samples were collected for fault complexity judgment.

This article uses the above-mentioned data set, and each fault diagnosis experiment is divided into training set, verification set and test set according to 8:1:1.

In order to verify the influence of the number of layers of SE-Inception-ResNet and the compression rate of SE Block on the accuracy of fault diagnosis model, the experiment were conducted with models with layers of 4, 6, 8 and 10 and compression rates of 1, 8, 16 and 32 respectively under other conditions unchanged. The evaluation results of test set are shown in Figure 3.2.1-1, which are analyzed from the number of layers of SE-Inception-ResNet. In general, the accuracy rate increases gradually with the increase of the number of layers, among which the upward trend is obvious from layer 4 to layer 8, and tends to be flat from layer 8 to layer 10, which shows that the change range of accuracy rate is not very large when the number of layers increases to a certain number of layers. From the analysis of compression ratio, generally speaking, the smaller the compression ratio, the smaller the impact on the accuracy. When the compression ratio is 1, 8 and 16, the downward trend is not obvious, and when the compression ratio reaches 32, the accuracy will decline

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 51, No. 2, JULY - DEC 2022

significantly.

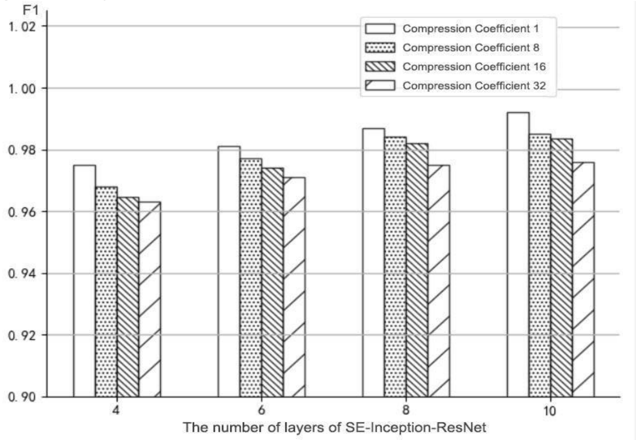


Figure 3.2.1-1: The influence of training parameters on theaccuracy of the fault diagnosis model

In order to verify the impact of SE-Inception-ResNet layers and SE Block compression ratio on the running time of fault diagnosis model, the experiment were carried out with 4, 6, 8,10 layers and 1, 8, 16, 32 layers, respectively, under other conditions unchanged. The evaluation results of test set are shown in Figure 3.2.1-2.From the analysis of the number of SE-Inception-ResNet layers, the training time gradually increases with the increase of the number of layers. When the number of layers is 4 and 6, the training time will not increase significantly. When the number of layers is greater than 8, the training time will be greatly increased. From the analysis of the compression rate, the smaller the compression coefficient, the longer the training time required. When the compression coefficient is set to 1, the training duration for uncompressed use is also the maximum; when the compression ratio is set to
8 and 16, the training duration decreases; when the compression ratio reaches 32, the training duration drops sharply, indicating that the larger the compression coefficientis, the greater the influence will be on the training duration ofthe model.

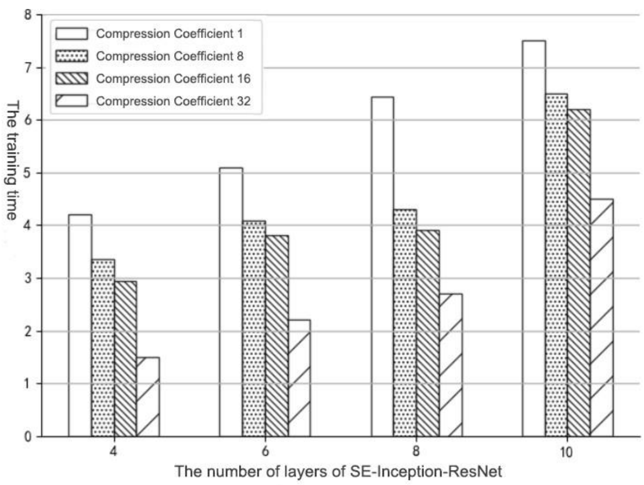


Figure 3.2.1-2: The influence of training parameters on thetraining time of the fault diagnosis model

On the whole, when the number of training layers reaches 8 layers, the growth of F1 tends to be flat. At the same time, the accuracy rate is at a higher level when the compression rate is 16.When the number of training layers is 8 layers and the compression rate is 32, the cost is the training duration of is the smallest, so when the number of training layers is 8 and the compression rate is 16, the accuracy and training duration are in a balanced state. Through the evaluation of the test set, the balance between training time and accuracy is comprehensively considered. This paper uses a model with a training layer of 8 and a compression rate of 16 for fault diagnosis.

3.2.2 Analysis of experimental results

Table 3.2.2-1: Two categories (simple, complex)

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 value (%)	Training duration (s)
LSTM	92.5	92.5	92.5	92.5	622.08
CNN	94.3	92	92.1	92	39.60
ResNet	96.8	96.8	96.8	96.8	376.56
Inception-ResNet	97.4	97.4	97.4	97.4	313.20
Improved Inception-ResNet	99.3	99.3	99.3	99.3	234.00

In this paper, binary discrimination experiments of simple and complex faults are carried out for LSTM, CNN, ResNet, Inception-ResNet and improved Inception-ResNet respectively. The number of training rounds is 100. The evaluation indexes such as accuracy, precision rate, recall rate and F1 value were used for comprehensive evaluation. The evaluation results of the test set are shown in Table 3.2.2-1. Among them, the convolution neural network represented by CNN shows strong advantages in F1 value and training time within 100 rounds, while the deep features extracted by CNN and ResNet in the training process are limited and the features are not rich enough, and the F1 value only reaches 92% and 94%. The deeper and wider Inception-ResNet model and the improved Inception-ResNet model have the same level of accuracy. However, the improved Inception-ResNet model can integrate the deep features of each channel and compress the feature space to improve the calculation efficiency. The model can perform better under the same number of rounds, so that training can be guaranteed.

4. Conclusion

This article mainly introduces the improved Inception-ResNet model and the related principles of SE-Inception-ResNet. Based on this method, a binary discrimination model of simple faults and complex faults is constructed. Through simulation experiments, the influence of experimental parameters on the diagnosis model is analyzed. Finally, the experimental performance of different networks is comprehensively analyzed through evaluation indicators such as accuracy rate, precision rate, recall rate, and F1 value.

References

[1] Xie L, Mo Y, Sinopoli B. “False data injection attacks in electricity markets” [C]. 2010 First IEEE International Conference on Smart Grid Communications. IEEE, 2010: 226-231.

[2] Zhao Binyan. “Research on power grid fault diagnosis method based on multi-information fusion” [D]. Nanjing: Nanjing University of Science and Technology, 2017.

[3] Ju Ping, Zheng Shiyu, Xu Qun, etc. “Survey of wide-area measurement system research” [A]. Electric Power Automation Equipment, 2004, 24(7)

[4] Han Yingduo, Wang Zhonghong, Lin Kongxing, etc. “Three frontier topics in the power system-flexible power transmission technology, intelligent control, GPS-based dynamic safety analysis and monitoring system” [J]. Journal of Tsinghua University (Natural Science Edition), 1997, 37(7)1-6

[5] DUAN Gang，YAN Yaqin，XIE Xiaodong，et al． “Development status quo and tendency of wide area phasor measuring technology” [J]． Automation of Electric Power Systems，2015，39(1)：73-80．

[6] Zhang Xu. “Practical research on fault diagnosis system for regional power grid” [D]. Beijing: North China Electric Power University, 2005.

[7] Dong Shizhen. “Research method of power grid fault diagnosis”[J]. Guangdong Electric Power, 2009, 22(5): 32-36.

[8] Ge Mengying, Yu Zhongzhong, Zhou Lan, Ma Yuxi. “Deep learning image classification algorithm based on collaborative semi-supervised”[J]. Computer Simulation, 2019, 36(02): 206-210.

[9] Liu Cuiling, Lin Long, Yu Chongzhong, et al. “Research on Peanut Hyperspectral Image Classification Method Based on Deep Learning” [J]. Computer Simulation, 2020, 037(003):189-192,283.

[10] Han Yaochang, Wang Jie, Shi Tong, et al. “Remote sensing image classification method based on improved GoogLeNet”[J]. Journal of Projectiles, Rockets, Rockets and Guidance, 2019, v.39; No.194(05):145-148+ 159.

[11] He K, Zhang X, Ren S, et al. “Deep Residual Learning for Image Recognition”[C]//Conference Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 27-30 June 2016. 770-778.

[12] Tang Hao, He Chu. “SAR image scene classification based on full convolution network and improved conditional random field-cyclic neural network” [J]. Computer Applications, 2016, 36(12): 3436-3441.

[13] Ge Mengying, Yu Chongzhong, Zhou Lan, Ma Yuxi. “Deep learning image classification algorithm based on collaborative semi-supervision” [J]. Computer Simulation, 2019, 36(02): 206-210.

[14] Thukaram D, Khincha H P, Vijaynarasimha H P. “Artificial neural network and support vector machine approach for locating faults in radial distribution systems” [J]. IEEE Transactions on Power Delivery, 2005, 20(2): 710-721.

[15] ZHOU S, WANG X, QIAN Q.” Bayesian networks based distributed fault diagnosis approach for power grids” [J]. Power System Technology, 2010, 34(9): 76-81.