Research on Partial Discharge Diagnosis Based on Data Augmentation and Convolutional Neural

Yi Yang¹, Chao Gu¹, Jiafeng Qin¹, Zhenyu Shao², Fang Song², Yao Li²,*

¹Equipment Status Evaluation Center, State Grid Shandong Electric Power Research Institute, Jinan 250000, China
²Star Institute, PDStars Electric Co., Ltd., Nanjing 210000, China

*Corresponding author e-mail: liyao@pdstars.com

Abstract. Partial discharge is a common phenomenon in the operation of electrical equipment. The detection of partial discharge is an important method to evaluate the insulation state of electrical equipment. Accurate and effective identification and evaluation of partial discharge is of great significance for the operation and maintenance of power equipment. This paper first introduces the basic characteristics of partial discharge, commonly used detection methods and defects. In view of the low efficiency, poor generalization and low accuracy of shallow neural network diagnosis, a new method based on data augmentation and convolution neural networks is proposed. The partial discharge diagnosis method of Convolutional Neural Networks (CNNs) is used to establish a partial discharge diagnosis system. Through the verification of the real PD diagnosis case, the partial discharge diagnosis system based on this method has higher recognition rate and stability.

1. Introduction

Partial discharge refers to the discharge phenomenon that occurs in a localized range of electrical equipment insulation under a sufficiently strong electric field. It can occur in the vicinity of the conductor or in other places. Partial discharge has a certain influence on the dielectric strength of the medium. The slight partial discharge has less influence on the insulation of the power equipment, and the insulation strength decreases slowly. The strong partial discharge will cause the insulation strength to drop rapidly [1]. Partial discharge is one of the important causes of insulation damage of high voltage power equipment [2-3]. Therefore, timely detection of faults in the PD of the running equipment and correct classification and identification of the faults are essential for the normal operation of the power equipment.

Currently PD diagnosis recognition technology, typically extracted from the raw data of a specified set of feature values, in order to achieve compression of the data amount and the dimension reduction [4]. Typical methods such as wavelet analysis [5-6], fractal feature parameter method [7-8], waveform feature method [9], etc.; researchers also use traditional machine learning methods such as BP (Back Propagation) neural network [10], to achieve partial discharge Identification of features.

The validity of feature extraction diagnosis depends largely on the selection of feature parameters, while the selection of parameters by manual method depends on a large number of theoretical and practical experience, and the number of samples that can be processed by manual method is relatively...
limited, so the selected data features often do not have generalization. The diagnosis based on neural network usually uses BP shallow network model, which is limited by the number of computing units in shallow network. It has limited performance and low recognition accuracy for complex network. Adding network layers directly to shallow model is easy to cause gradient dispersion and does not have good expansibility [11].

This paper introduces a diagnostic method based on data augmentation and convolutional neural network. Firstly, the data of the field interference coupling and Gaussian noise are processed by the data augmentation method to construct a balanced data sample set. Then the convolutional neural network is built based on the data sample set. The feature extraction and classification of the original data can be automatically realized, and the complicated artificial feature extraction process is avoided. From the results, the method has high diagnostic efficiency and diagnostic accuracy.

2. Another section of your paper

2.1. Data Feature Description

The training and verification of neural networks relies on a large number of samples, and the quality of the sample set largely determines the performance of the neural network model [12-13]. The training and test data used in this paper are derived from measured partial discharge data from multiple substation sites. The field measurement data is characterized by distinctive features and true data. However, due to different data sources and different durations of defects, the number of samples collected for different defects is quite different, which easily leads to sample imbalance [14]. If the neural network training is directly performed based on the unbalanced sample, the trained model has poor generalization ability and is prone to over-fitting. Data augmentation is one of the important means to solve the sample imbalance. For a small number of samples, by processing the existing samples and generating new samples, the sample size of each type of label can be balanced to avoid the influence of sample imbalance on the training results [15].

Data augmentation methods commonly used in the field of image recognition include image processing, rotation processing, scaling processing, random clipping, etc., which are suitable for identifying scenes in which objects have different angles in different samples. In the partial discharge pattern recognition, the angle of the recognition object is usually relatively simple. In addition, PD images are often coupled with various kinds of noise, such as mobile phone noise, light noise, etc. At the same time, the image pixels and shapes from different sources are different. The model trained with a single sample is only applicable to the image with specified color and specifications, and its generalization is poor. Therefore, the traditional data augmentation method is not applicable.

2.2. Data Augmentation Method Based on Environmental Noise Coupling and Gauss Data Augmentation Method Based on Environmental Noise Coupling and Gauss

In order to solve the problems encountered in partial discharge image detection described above, this paper proposes a data augmentation method combining environmental noise coupling and Gaussian noise. The process flow is shown in the following figure:
Firstly, common jamming on the input data overlay, including but not limited to radar jamming, mobile phone jamming, energy-saving lamp jamming and so on. The superposition method is: converting the partial discharge raw data collected by the acquisition front end into three-dimensional data whose phase is the x-axis, the period is the y-axis, and the amplitude is the z-axis; Then according to the data characteristics of the on-site interference, the three-dimensional data corresponding to each interference is generated separately; then the partial discharge raw data and the interference data are subjected to amplitude accumulation, and the interference superimposed data is obtained.

Gaussian noise processing is performed on the superimposed data: the interference superimposed data is converted into a picture file, and the RGB (Red, Green, Blue) value of the picture file is extracted, and the values of the three values of R, G, and B are all [0, 255]; For each R, G, B pixel value of each point of the picture file, the noise is calculated by a two-dimensional Gaussian distribution function. The two-dimensional Gaussian distribution function is:

\[ G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \]  

In the formula, \( x \) and \( y \) are the horizontal and vertical coordinates of the points in the graph, \( \pi \) is the circumference, \( e \) is the natural constant, and \( \sigma \) is the standard deviation of the normal distribution [16].

Gauss noise is additive noise. After calculating the Gauss distribution, the superimposed pixel values can be obtained by directly superimposing the pixel values. A new pixel value is used to save the image, and the image processed by Gauss noise is obtained. The image is merged into the sample library with the original data as training and testing samples for deep learning.

A practical example based on interference superposition and Gaussian noise is as follows:

\[ G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \]
Figure 2. Data augmentation example

The data augmentation method solves the problem of data coupling noise in real application scenarios by superimposing common interference in the field, and solves the problem of different image data specifications and pixels from different sources by processing Gauss noise. The two methods work together to make the enhanced samples more representative and practical.

After applying the data augmentation method, the number of samples used in this paper is as follows:

| PD type       | Original sample number | Augmentation sample size |
|---------------|------------------------|--------------------------|
| Corona        | 15384                  | 20000                    |
| FloatingElectrode | 16930               | 20000                    |
| Void          | 8642                   | 20000                    |
| Surface       | 11200                  | 20000                    |
| Particle      | 5921                   | 20000                    |
2.3. Mixed partial discharge signal
In addition to the typical partial discharge samples listed in the table above, some of the samples used in this paper also contain mixed partial discharge samples with two or more partial discharge characteristics. For example, the signal shown in Fig. 3 contains both corona discharge and floating electrode discharge. Experts confirm that the designated tag is 55% corona discharge and 45% floating electrode discharge.

![Figure 3. Mixed partial discharge sample](image)

Partial discharge is a widespread defect in electrical equipment. There are often multiple mixed signals in power equipment that have long running time and lack of maintenance. The introduction of mixed samples further enhances the ability of the sample set to cover real PD signals.

3. Partial Discharge Detection System Based On Convolutional Neural Network

3.1. Design of Convolutional Neural Network
Convolutional neural networks are a kind of deep learning network that uses local perception and simulation of real horizons [17]. It is one of the representative algorithms of deep learning. Different from the traditional method of artificial extraction, the core idea of convolutional neural network is to extract the features of the target automatically by convolution operation and establish a structural model. In 2012, with the success of AlexNet, convolutional neural networks have been widely used in the fields of image processing, natural language processing and other local and global information with obvious correlation [18-19].

This paper designs a supervised deep convolutional neural network. The main structure is shown in Fig. 4:

![Figure 4. Deep convolutional neural network](image)
In Fig. 4, the Conv layer is the convolutional layer, its main function is to perform feature extraction on the input data using a convolution operation [20]. Each convolution layer contains multiple convolution kernels, each convolution kernel has a corresponding weight coefficient, and the local information is spliced to the whole information through the translation of different convolution kernels on the image [21]. The Pool layer is the pooling layer, also known as a down sampling layer, its main function is feature selection and information filtering. Through the pooling operation, the feature map volume is reduced, the network computation complexity is reduced, and the convergence speed is accelerated. On the other hand, the pooling operation can extract the main features again based on the features extracted by the convolution layer, while filtering out the features with low correlation and weak importance. The FC layer is a fully connected layer that connects all features and delivers the output values to the classifier for classification [22-23].

The network used in this paper is an 8-layer network consisting of five convolution layers and three full connection layers. Among them, the first, second and fifth convolution layers are supplemented by the maximum pooling layer. The first convolution layer input size is 256 x 256 x 3, of which 256 x 256 is the pixel size of the image, and 3 represents the number of channels of the image. The training samples used in this paper are RGB images processed by interference coupling and Gauss noise, which contain red, green and blue pigments, so the number of channels is 3. The first convolution layer uses 96 convolution cores with size of 11 *11 *3 and step size of 4. The feature extracted by the first layer convolution kernel is used as the input of the second layer convolution after maximum pooling. The second layer convolution uses 256 convolution cores with size of 5 *5 *48 to convolute, and also carries out maximum pooling operation in convolution. The third convolution layer is connected to the second convolution layer, and there are 384 convolution cores with a size of 3 x 3 x 256. The fourth convolution layer has 384 convolution cores of 3 *3 *192 in size, and the fifth convolution layer has 256 convolution cores of 3 *3 *192 in size. The fifth convolution layer is connected with the largest pooling layer, and then output to three full-junction layers in sequence, each of which has 4096 neurons.

In the model, the excitation function of all excitation layers selects ReLU (Rectified Linear Unit), because ReLU function has better performance in the gradient descent process than other activation functions, and the convergence speed of ReLU function is also better than Sigmoid. Function and Tanh function [24-25].

After the convolutional neural network is constructed, a 10-fold cross-validation is performed on 100,000 typical samples and 50,000 atypical sample sets containing data-enhanced samples, that is, the sample set was divided into 10 samples, 9 of which were taken as training set and the other one as test set. Repeat the training process ten times, and take the model with the highest accuracy as the final model. The accuracy of the five groups of models is as follows:

| Numbering | Number of samples | Accuracy Rate |
|-----------|-------------------|---------------|
| 1         | 150000            | 97.16%        |
| 2         | 150000            | 96.72%        |
| 3         | 150000            | 96.30%        |
| 4         | 150000            | 97.58%        |
| 5         | 150000            | 95.34%        |

Finally, this paper selects Model 4 as a model for field applications. The classification accuracy of this model in multi-category tasks is 97.58%, which is better than the traditional method of 85%~91%.

3.2. Diagnostic Scheme Design and Deployment

Based on the deep learning diagnostic model with high diagnostic accuracy and practicability, a complete partial discharge diagnostic system was built. The system framework is shown in the following figure:
Figure 5. Partial discharge diagnosis system

The sample library provides PD samples, establishes an intelligent diagnosis model, and establishes a partial discharge intelligent diagnosis system, with the verified and tested diagnostic model as the core, real-time diagnosis of the data collected by the monitoring device, judging the current running state of the device, and giving type diagnosis and alarm prompts for the PD phenomenon, providing guidance for manual troubleshooting. At the same time, the data that detected the fault is verified and filtered and added to the sample library to form a complete closed loop.

The system can be used in the scenes of partial discharge state detection, data analysis and device state assisted diagnosis of various types of power equipment. It not only reduces the requirement of field operation and maintenance personnel and the dependence of experts, but also improves the level of equipment condition control, the working efficiency of live detection and the reliability of power network operation. At the same time, the real data entered on the site continuously expands the size of the sample library, further promoting the continuous iterative optimization of the deep learning convolution model.

3.3. Application Analysis

For example, in the transformer field monitored by the PD diagnostic system, an abnormal UHF (Ultra High Frequency) signal is continuously collected at the C-phase observation window position of the transformer. The position of the sensor is shown in the red circle mark position as shown below:

Figure 6. Sensor installation position

The typical signal amplitude collected by the sensor is 51dB, and there are stable large and small two-cluster discharge pulses, and the signal has power frequency correlation. The typical map is shown below:
The atlas was diagnosed as floating electrode discharge by deep learning model. A large number of diagnoses were floating electrode discharge in continuous time, and triggered system alarm. Based on the diagnostic results of in-depth learning, the technicians synthetically analyzed the abnormal data and determined that there was a floating electrode discharge source near the sensor, and the discharge situation was serious. After on-site confirmation, the final location confirms that there exists discharge phenomenon in the upper iron yoke of phase C of transformer. The discharge type is floating electrode discharge, which is consistent with the diagnosis results.

The dismantling inspection of the equipment found that the red circle marking place had a continuous discharge phenomenon due to the foreign body left behind, and the marking area had obvious ablation marks, and the discharge was more serious. As shown in the following figure.

This case proves that the convolutional neural network model used in this paper has an accurate ability to judge the PD characteristics.

4. Conclusion
This paper designs a PD data diagnosis system based on data enhancement and deep convolution neural network. The system first solves the problem of sample imbalance and insufficient sample quantity by
data enhancement of partial discharge map. Based on this, the deep discharge convolutional neural network algorithm is used to diagnose the partial discharge map. Through the example verification, the system can accurately identify the fault type, has good recognition ability and diagnostic ability, and solves the problem of relying on artificial extraction features and low recognition accuracy in the traditional PD diagnostic method.

Acknowledgments
This work was financially supported by Science and Technology Project fund of State Grid Shandong Electric Power Company (NO: 520626180017).

References
[1] Liu Hong-bin, Liu Lian-rui, Liu Shao-yu, Deng Chun. Application of On-Line Detection on Transmission and Distribution Equipment in North China Grid [J]. North China Electric Power, 2009(08): 35-37.
[2] Zhang Jie, Fu Quan-yong, Yuan Ye. Application Research on Partial Discharge Charge Detection Technology of Transformer [J]. Transformer. 2018, 55(08): 66-71.
[3] Zhong Li-peng, Ji Sheng-chang, Cui Yan-jie, Wang Yuan yuan, Meng Yan, Liang Nai-feng, Sun Da-lu. Study on Typical Defective Discharge Characteristics of Transformer and Its Charge Detection Technology [J] High Voltage Apparatus. 2015, 51(03): 15-21.
[4] Zhang Sha, Song Jian-cheng, Wen Min-min, Tian Mu-qin. Current Status and Prospects of Typical Partial Discharge in Power Transformer Insulation System [J]. Transformer. 2018, 55(04): 32-40.
[5] Cai-xin, S.U.N., Xin, L. and Jian, L., 2001. Research on complementarity between wavelet and fractal theory and relevant application in PD pattern recognition. Proceedings-Chinese Society of Electrical Engineering, 21(12), pp.73-76.
[6] Dan Wen-gang, Chen Xiang-xu, Zheng Jian-chao. Classification Of Partial Discharge Distribution Patterns Using Wavelet Transform And Neural Network [J]. Proceedings of the Csee. 2002(09): 2-6+19.
[7] Satish, L. and Zaengl, W.S., 1995. Can fractal features be used for recognizing 3-d partial discharge patterns. IEEE Transactions on Dielectrics and Electrical Insulation, 2(3), pp.352-359.
[8] Li Hua; Cheng Chang-kui, Chen Jiao, Gan Jian-wei, Hu Lin, Li Xin-bo. Fractal Dimension Research of the Partial Discharge UHF Signal in GIS Based on EMD [J]. High Voltage Apparatus. 2014, 50(06): 104-110.
[9] Gao Kai, Ni Hao, Si Wen-rong. GIS partial discharge detection and analysis of its waveform characteristics [J]. East China Electric Power. 2010, 38(10): 1512-1517.
[10] Jiang Guo-qing, Li Lu, Li Dao, Xu Dong-sheng. Partial Discharge Pattern Recognition of High Voltage Reactor Based on BP Network [J]. Electrotechnical Application. 2015, 34 (04): 86-89.
[11] Liu Bing, Zheng Jian. High Voltage Apparatus. Partial Discharge Pattern Recognition of Transformer Based on Convolutional Neural Network [J]. 2017, 53(05): 70-74+81.
[12] Tanner M A, Wong W H. The calculation of posterior distributions by data augmentation [J]. Journal of the American statistical Association, 1987, 82(398): 528-540.
[13] Frühwirth-Schnatter S. Data augmentation and dynamic linear models [J]. Journal of time series analysis, 1994, 15(2): 183-202.
[14] Chen Wen-bing, Guan Zheng-xiong, Chen Yun-jie. Data augmentation method based on conditional generation confrontation network [J]. Journal of Computer Applications. 2018, 38(11): 3305-3311.
[15] Salamon J, Bello J P. Deep convolutional neural networks and data augmentation for environmental sound classification [J]. IEEE Signal Processing Letters, 2017, 24(3): 279-283.
[16] Mandelbrot B B. A fast fractional Gaussian noise generator [J]. Water Resources Research, 1971,
7(3): 543-553.

[17] LeCun, Y., Bengio, Y. and Hinton, G., 2015. Deep learning. nature, 521(7553), p.436.
[18] Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets [J]. Neural computation, 2006, 18(7): 1527-1554.
[19] Salakhutdinov R, Hinton G E. Deep Boltzmann Machines[C]//AISTATS. 2009, 1: 3.
[20] Wang Zhen, Gao Mao-ting. Design and Implementation of Image Recognition Algorithm Based on Convolutional Neural Network [J]. Modern Computer. 2015(20): 61-66.
[21] Fialka, O. and Cadik, M., 2006, July. FFT and convolution performance in image filtering on GPU. In Information Visualization, 2006. IV 2006. Tenth International Conference on (pp. 609-614). IEEE.
[22] Yin Bao-cai, Wang Wen-tong, Wang Li-chun. Review of Deep Learning [J]. Journal of Beijing University of Technology. 2015, 41(01): 48-59.
[23] Chang Liang, Deng Xiao-Ming, Zhou Ming-Quan, Wu Zhong-Ke, Yuan Ye, Yang Shuo, Wang Hong-An. Convolutional Neural Networks in Image Understanding [J]. Acta Automatica Sinica. 2016, 42(09): 1300-1312.
[24] Lu Hong-tao, Zhang Qin-chuan. Applications of Deep Convolutional Neural Network in Computer Vision [J]. Journal of Data Acquisition and Processing. 2016, 31(01): 1-17.
[25] Sun Y, Wang X, Tang X. Deep learning face representation from predicting 10,000 classes[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2014: 1891-1898.