

# PRPD data analysis with Auto-Encoder Network

Songyuan Li<sup>1</sup>, Yuyan Man<sup>1</sup>, Chi Zhang<sup>1</sup>, Qiong Fang<sup>1</sup>, Suya Li<sup>1</sup>, Min Deng<sup>2,\*</sup>

<sup>1</sup>Electric Power Research Institute, State Grid Tianjin Electric Power Company, 300384 Tianjin, China

<sup>2</sup>Red Phase INC., 361005 Xiamen, Fujian

**Abstract:** Gas Insulated Switchgear (GIS) is related to the stable operation of power equipment. The traditional partial discharge pattern recognition method relies on expert experience to carry out feature engineering design artificial features, which has strong subjectivity and large blindness. To address the problem, we introduce an encoding-decoding network to reconstruct the input data and then treat the encoded network output as a partial discharge signal feature. The adaptive feature mining ability of the Auto-Encoder Network is effectively utilized, and the traditional classifier is connected to realize the effective combination of the deep learning method and the traditional machine learning method. The results show that the features extracted based on this method have better recognition than artificial features, which can effectively improve the recognition accuracy of partial discharge.

## 1 Introduction

Due to its small size, stable operation, and low electromagnetic pollution, Gas Insulated Switchgear(GIS) is more and more widely used [1-4]. However, during the design, transportation, manufacturing, and long-term operation of GIS equipment, various insulation defects often appear inside the equipment, which in turn induce equipment insulation failure. Partial Discharge(PD) is an important feature of various latent insulation faults in the early stage [6-7], and different partial discharge types have different degrees of damage to equipment. Therefore, it is important to monitor and identify the partial discharge phenomenon timely and effectively [8-9].

GIS equipment radiates Ultra-High Frequency(UHF) signals to the outside when a partial discharge occurs. Effective use of UHF signals to extract available identification features is of great significance in real-time monitoring of GIS equipment insulation conditions. At present, the methods for UHF data mainly include Time Resolved Partial Discharge (TRPD) and Phase Resolved Partial Discharge (PRPD)[10-11]. Compared with the TRPD, the PRPD has mature technology and good stability. We mainly studies the feature extraction method based on PRPD data.

Currently, the partial discharge pattern recognition method for PRPD data mainly constructs artificial features, which has the disadvantages of strong subjectivity and great uncertainty. In view of the above problems, we proposes to introduce Convolutional Neural

---

\* Corresponding author: maoheng@redphase.com.cn

Network (CNN) and its extended network, Auto-Encoder Network, which can reconstruct the input data and then extract more essential features about the input data through the coding network. The research shows that CNN [12], with the powerful feature adaptive extraction ability, can effectively replace the traditional artificial feature engineering to explore deeper feature of the original partial discharge signal data.

## 2 Artificial feature extraction method for PRPD data

PRPD is a commonly used mode in the field of partial discharge pattern recognition. The traditional PRPD data statistical feature extraction method mainly introduces five kinds of statistical operators related to standard normal distribution to use to quantitatively analyze the feature distribution, which are skewness  $S_k$ , steepness  $K_u$ , phase asymmetry  $\Psi$ , discharge factor  $Q$ , and phase correlation coefficient  $C_c$  [13]. Skewness is used to measure the degree of asymmetry of the data distribution relative to the normal distribution. The steepness is used to describe the degree of protrusion of a shape compared to the shape of a normal distribution. The phase asymmetry is used to describe the phase distribution difference of the initial voltage on the positive and negative half cycles of the spectrum. The discharge amount factor  $Q$  and the phase correlation coefficient are used to describe the contour difference between the positive and negative half cycles of the spectrum.

## 3 Gray-scale maps of partial discharge

For the characteristics that convolutional neural networks are good at processing image data, we convert PRPD data into a gray-scale maps of partial discharge. The pulse signal amplitude  $q$  is plotted on the vertical axis and divided into  $M$  intervals. The power frequency phase  $p$  is the horizontal axis and is divided into  $N$  intervals. Then, We can divide the  $q$ - $\phi$  plane of the PRPD data into  $M \times N$  intervals, count the pulse number  $n_{u,v}$  ( $u \in M, v \in N$ ) of each interval, construct the  $q$ - $\phi$ - $n$  map, and then normalize the  $n_{u,v}$ :

$$n'_{u,v} = n_{u,v} / n_{\max} \quad (1)$$

where  $n'_{u,v}$  is the number of normalized pulses,  $n_{u,v}$  is the actual number of pulses, and  $n_{\max}$  is the maximum number of pulses of the  $q$ - $\phi$ - $n$  map [16]. Using the PRPD data to construct the gray-scale maps of partial discharge can effectively extract the feature with the gray-scale maps of partial discharge can obtain more global feature information.

## 4 Partial discharge pattern recognition based on Auto-Encoder Network

Auto-Encoder Network is a neural network used to efficiently encode input data based on the original hierarchical structure of the neural network. The purpose of the Auto-Encoder Network is to reproduce the input signal as much as possible. In order to realize this recurrence, the Auto-Encoder Network must capture the components that most represent the input data. Therefore, it realizes the extraction of the most essential features of the input data [15].

### 4.1 Data processing

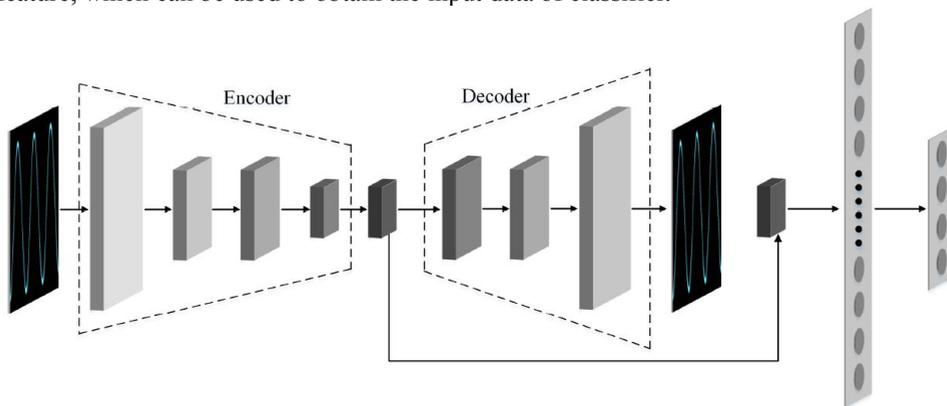
We relies on the self-developed data acquisition platform to simulate the complex environment in the actual operation of GIS equipment, and collects four types of partial discharge UHF data. The four types of partial discharge UHF PRPD data finally obtained are divided into training sets and test sets, and the specific division is as follows(Table 1)

**Table. 1** Partial discharge UHF PRPD data partition table

	Corona discharge	Free metal particle discharge	Insulation void discharge	Floating electrode discharge
Train set	6900	6900	6900	6900
Test set	1725	1725	1725	1725

### 4.2 Network structure

Aiming at the practical problems of four partial discharge pattern recognition, we proposes a partial discharge type identification network structure based on Auto-Encoder Network (Fig. 1). The coding network part of the Auto-Encoder Network is equivalent to the feature extractor of the input data. The Auto-Encoder Network is trained to reconstruct the input data. At this time, the output of the coding network can be regarded as the identification feature, which can be used to obtain the input data of classifier.



**Fig. 1** Identification of partial discharge pattern recognition network based on auto-encoder network

At this time, the classifier is no longer limited to the deep learning method, and the traditional classic machine learning classifiers such as SVM, random forest, etc. can be used. Through the above method, the feature extraction capability of the Auto-Encoder Network can be fully utilized, combining the traditional machine learning method and the deep learning method.

The network is composed of a coding network, a decoding network, and a classifier network. The coding network extracts the features of the input signal, and the decoding network reconstructs the original signals through the features extracted by the coding network, while the classifier network uses the features extracted by the coding network to perform the classification task.

## 5 Identification results and analysis of partial discharge types

Through a large number of experiments, Table 2 shows the Accuracy comparison between partial discharge type identification of five artificial features based on PRPD data extraction and features extracted based on deep learning methods in traditional machine learning classifiers (SVM, random forest, BP neural network).

The bolded parts in Table 2 indicate the higher classification accuracy of the same classifier in each discharge type. It can be seen from Table 2 that for the same classifier, the features extracted based on the Auto-Encoder Network have higher average accuracy in SVM, random forest, and BP neural network than the traditional artificial features, and the average accuracy rate can better evaluate the performance of the model. This also reflects from the side that the features extracted through the Auto-Encoder Network have stronger recognition than the artificial features. At the same time, we finds that when the input features are derived from the Auto-Encoder Network, the recognition performance of the three types of classifiers for different partial discharge types is also quite different. The SVM and BP neural networks show better classification performance for corona discharge and floating electrode discharge, while random forests have higher recognition level for free metal particle discharge. Moreover, the recognition accuracy of each type of classifier for the insulation void discharge is at a low level regardless of which feature is used. This is because the data type of PRPD itself has the disadvantage of not being able to reflect the main characteristics of insulation void discharge [14]. In summary, the features extracted based on the Auto-Encoder Network have a better effect in the identification task of partial discharge than the artificial features.

**Table. 2** The result of classifiers. The s.f. is the statistical feature and the n.f. is the feature extracted from network.

	Corona discharge	Free metal particle discharge	Insulation void discharge	Floating electrode discharge	Average accuracy
SVM (s.f.)	86.21%	42.74%	43.68%	46.23%	54.72%
SVM (n.f.)	<b>99.63%</b>	<b>94.55%</b>	<b>89.36%</b>	<b>96.21%</b>	<b>94.94%</b>
BP (s.f.)	91.80%	48.08%	67.25%	66.55%	68.42%
BP(n.f.)	<b>98.74%</b>	<b>93.87%</b>	<b>87.45%</b>	<b>97.56%</b>	<b>94.41%</b>
Random forest(s.f.)	86.38%	56.81%	66.59%	39.38%	62.29%
Random forest(n.f.)	<b>89.16%</b>	<b>97.65%</b>	<b>86.65%</b>	<b>94.77%</b>	<b>92.01%</b>

## 6 Conclusion

For the pattern recognition problem of GIS partial discharge, we first converts PRPD data into gray-scale maps of partial discharge and then constructs a convolutional neural network based on self-encoding technology to extract the characteristics of gray-scale maps of partial discharge and learn with traditional machine learning. The method is organically combined, which effectively improves the recognition ability of the traditional classifier for partial discharge signals. At the same time, the experiment proves that when using the same classifier, the features extracted based on the Auto-Encoder Network have stronger recognition than the artificial features extracted based on the traditional method, which can effectively improve the recognition accuracy of the partial discharge type.

## References

1. D. Kopejtkova, T. Molony, S. Kobayashi, et al. *CIGRE* **23**,101(1992).
2. H. Duan, Z. Li, H. zhan, et al. *Fuel & Energy Abstracts* **37**,58(2015)
3. H. Zhan, L. Yao, Z. Li, et al. *Power System Technology* **38**.9:2570-2574(2014).
4. W. Tang, S. Zheng, C. Li, et al. *Power System Technology* **38**.1,241-247(2014)(in Chinese).
5. X. Zhang, J. Tang, C. Pan, et al. *Power System Technology* **40**.10,3272-3278(2016).
6. Y. Qian, C. Huang, X.Jiang, et al. *Power System Technology* **29**.1,40-43(2005)
7. J. Tang. *High Voltage Engineering* **38**.6,1281-1291(2012)
8. G. Wang, Y. Hao, Y. Li., *Advanced Technology of Electrical Engineering and Energy* **20**.2,52-57(2001).
9. L. Zhong, S. Ji, Y. Cui, et al. *High Voltage Apparatus*, **51**.3,15-21(2015).
10. C. Chang, C. Chang, J. Jin, et al. *IEEE Transactions on Dielectrics and Electrical Insulation* **12**.2,374-386(2005).
11. T. Jiang, J. Li, Y. Zheng, et al. *Energies* **4**.7,1087-1101(2011).
12. Y. Lecun, Y. Bengio, H. Hinton, et al. *Nature* **521**.7533,436-444(2015).
13. E. Gulski, F. Kreuger. *IEEE Trans on Electrical Insulation* **27**.1,82-92(1992) .
14. Z. Tang, M. Tang, J. Li, et al. *High Voltage Technology* ,**43**.7,2263-2277(2017).
15. J. Zhang, H, Wang, Y. Guo, et al. *Journal of Computer Applications* (2018).
16. Q. Ding. *High Voltage Technology* (2011).