WILEY | Hindawi

Retraction

Retracted: Research on Pattern Recognition Method of Transformer Partial Discharge Based on Artificial Neural Network

Security and Communication Networks

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Xi, L. Yu, B. Chen, G. Chen, and Y. Chen, "Research on Pattern Recognition Method of Transformer Partial Discharge Based on Artificial Neural Network," *Security and Communication Networks*, vol. 2022, Article ID 5154649, 12 pages, 2022.

WILEY | Hindawi

Research Article

Research on Pattern Recognition Method of Transformer Partial Discharge Based on Artificial Neural Network

Yu Xi [ID], Li Yu, Bo Chen, Guangqin Chen, and Yimin Chen

Digital Grid Research Institute, China Southern Power Grid, Guangzhou, Guangdong 510000, China

Correspondence should be addressed to Yu Xi; 171847066@masu.edu.cn

Power transformer is pivotal equipment in a power system, which is responsible for energy transmission and transformation, and its operating condition is related to the safe operation of the power system. In the 21st century, computer science has entered a stage of rapid development, advanced network structures and algorithms have been applied to the field of artificial intelligence, and pattern recognition theory and technology have also made great progress. In the past, the identification of partial discharge type mainly relied on the experience of operation and maintenance personnel, and manual analysis and judgment were made based on partial discharge mapping, which was not very accurate. The application of the computer pattern recognition method in the field of partial discharge type identification has changed the status quo of manual identification, and this method has substantially improved the accuracy and efficiency of identification. Pattern recognition using computer technology has been applied to the field of partial discharge analysis. Compared with manual recognition, its recognition results are accurate, recognition speed is fast, and it has great potential for development. This paper proposes an artificial neural network-based model for transformer partial discharge pattern recognition, which combines the advantages of artificial neural networks with accurate extraction of local spatial higher-order features and provides a new solution for transformer partial discharge pattern recognition. Extended experiments show that the method proposed in this paper achieves leading performance and has practical application value.

1. Introduction

China's energy distribution and demand distribution are extremely uneven, with about 80% of energy resources concentrated in the northwest, while nearly 70% of the power load is concentrated in the economically developed areas of the central and eastern regions, showing the characteristics of "source and load cutoff" [1]. The unbalanced geographical distribution of energy and economy requires that in the process of power transmission, transmission line losses should be reduced as much as possible, reducing the area of the transmission corridor, while improving the transmission capacity. Since the 1980s, with the rapid development of power transmission technology, a variety of new methods to improve transmission capacity have emerged [2]. Among them, high-voltage direct current transmission (HVDC) has received widespread attention because of its transmission capacity, large capacity, long

distance, high efficiency, economy, and other characteristics. It can effectively solve the problem of oversupply of electric load, realize cross-regional, long-distance, and large-scale electric power transmission in China's power system, and promote the formation of west-east and north-south electric power transmission patterns [3]. Thus, the high-voltage DC transmission technology can, to a certain extent, provide effective solutions and guarantees for the optimal allocation of energy resources, the development and utilization of new energy sources, the national energy security supply, and the management of ecological and environmental problems in China and can effectively exploit the potential of existing equipment, which is of great significance to China's economic development and environmental management and has considerable social and economic values [4].

As one of the key pieces of equipment of a high-voltage DC transmission system, the insulation condition of the transformer will have a direct impact on the reliability of the

DC transmission system. Figure 1 shows the scale of China's transformer industry, the horizontal coordinate unit is a year and the vertical coordinate unit is a million-kilo ampere, and it can be found that the scale of China's transformer industry has been at a high level in recent years [5]. Unlike ordinary power transformers, the transformer structure is complex and its valve-side windings are subjected to the combined action of alternating current (AC), DC, and pulse voltages. With the increasing level of DC transmission voltage, certain weak links of its internal oil-paper insulation structure will generate or even intensify partial discharge, leading to transformer operation failure. Partial discharge not only characterizes the insulation degradation of transformers but is also a critical factor in causing insulation degradation. Effective identification of the type and severity of partial discharge inside the transformer can promptly detect some latent faults inside the transformer, provide a basis for diagnosis and condition maintenance strategy, and ensure the safe and stable operation of the system. Therefore, the research of transformer partial discharge pattern identification method has extremely important engineering practical value [6]. With the increasing progress and development of wireless communication technology, sensor technology, and computer technology, power big data and power big data analysis technology also came into being, which provides data basis and technical means for the research of transformer partial discharge pattern identification method. These data are characterized by the large volume, many types, fast speed, low-value density, etc. How to make full use of the information implied in these data, extract partial discharge fingerprint features, and combine with intelligent diagnosis methods to accurately and efficiently identify different defect types and discharge severity is one of the hot issues of current research [7].

As one of the most intelligent features and cutting-edge research fields in artificial intelligence, the artificial neural network can effectively mine the implicit information in the data, extract key information, and gain insight into minute features, which is suitable for transformer partial discharge pattern recognition in the context of electric power big data. In recent years, with the proposed and continuous development of deep learning (DL) methods, various improved structural units and optimization algorithms have been proposed one after another. As a branch of machine learning, artificial neural networks are particularly suitable for processing different types of data such as large-scale images, sounds, and texts. It mainly adopts the idea of layer-by-layer training to process large-scale data, from which the deep-level feature expression of the input data is mined, and has outstanding performance in feature expression, data dimensionality reduction, and classification prediction. Artificial neural network-based algorithms have been initially applied in many fields such as medical imaging recognition, speech recognition, face recognition, image recognition, and industrial automation. The application of deep learning in the above fields is based on its powerful feature extraction capability to extract deep features characterizing the nature of input data from shallow features such as pictures, data, and waveforms and to recognize them.

In the same way, transformer partial discharge pattern recognition is also based on the statistical mapping of partial discharge, grayscale images, or electromagnetic wave partial discharge signal data for recognition and classification [8]. Therefore, the continuous development and wide application of deep learning have opened up new ideas for transformer partial discharge pattern recognition. At present, deep learning methods are less applied in the field of transformer partial discharge pattern recognition. Therefore, how to introduce deep learning into transformer partial discharge pattern recognition and study the application effect of the method in partial discharge pattern recognition is significant for transformer maintenance and intelligent diagnosis and early warning.

In the normal operation of a power system, the stability of the power transformer has a great role. Once a power transformer fails, it will have an incalculable impact. One of the main causes of power transformer failure is the internal partial discharge of power transformers. Therefore, the analysis of power transformer partial discharge is an important way to detect power transformer faults and maintain the safety and stability of the power grid [9]. On the basis of analyzing the local discharge signal of power transformer, two aspects need to be studied, firstly, considering that there are more useless noise interference in the discharge signal if the signal analysis is carried out directly, the effect is very poor, and the noise elimination process of the discharge signal is needed first to ensure the stability of the signal; secondly, considering the limitations of traditional methods, this paper uses deep learning to build a suitable local model to solve the limitations. In this paper, we use the deep learning method to build a neural network for localized discharge signal recognition. In traditional recognition methods, feature extraction is a key step in fault diagnosis, and feature extraction mostly relies on personal experience and has some limitations. The main contribution of this paper is to propose a partial discharge pattern recognition method based on a convolutional neural network for the limitations of traditional methods and to input data into the convolutional neural network model. The features are extracted by a convolutional layer to avoid the manual operation of feature extraction and selection and enhance the intelligence of recognition, while the two features of local sensing and weight sharing speed up the training of samples and better adapt to today's big data era. Experimental results show that the proposed artificial neural network-based recognition method has a high accuracy rate, thus demonstrating the feasibility of applying the artificial neural network algorithm to the field of transformer partial discharge type recognition.

The following is a description of the study. The introduction is given in Section 1, and Section 2 goes over the related studies, including transformer partial discharge, partial discharge pattern recognition, and artificial neural networks. Convolutional neural networks, the LSTM deep learning algorithm, model architecture, and algorithm flow are all discussed in Section 3. Experimentation and evaluation are covered in Section 4. Section 5 consists of discussion; finally, in Section 6, the conclusion is drawn.
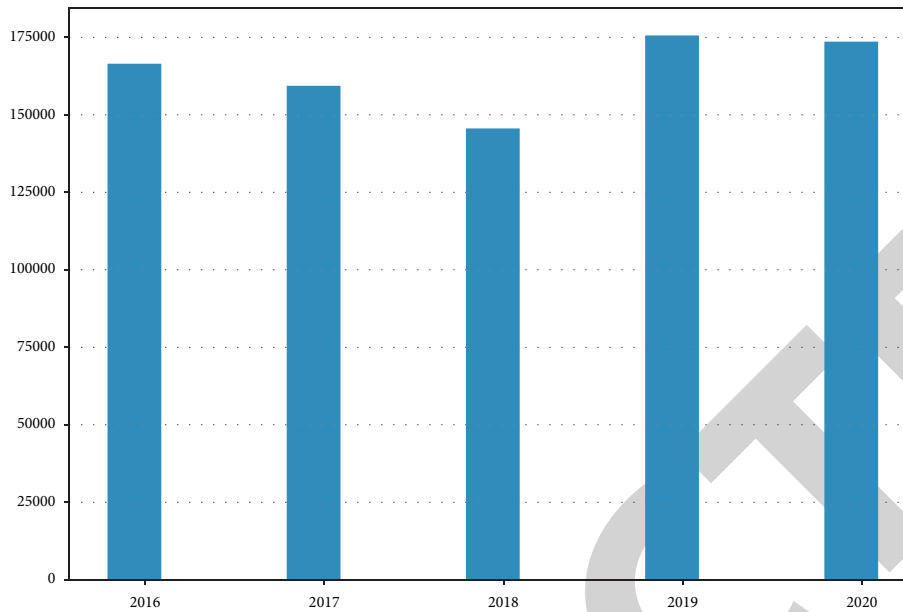
Figure 1: China transformer industry scale.

2. Related Work

In this section, we explain transformer partial discharge, partial discharge pattern recognition, and artificial neural networks in detail.

2.1. Transformer Partial Discharge. With the development of China's power industry, the voltage level of the power system has been increasing, and the capacity and the voltage level of power transformers have also been increasing. Statistics released by the Cliff Electronic Components (CEC) Industry [10] Development and Environmental Resources Department show that by the end of November 2017, the installed capacity of China's power plants of 6,000 kW and above was 1.68 billion kW, an increase of 7.2% year-on-year. From January to the end of August 2017, China completed 388.5 billion kWh of cross-region power transmission, and the total amount of power delivered by China's provinces was 1,030.6 billion kilowatt-hour (kW-h). China's power grid accounts for 220 kilovolts (kV) and above, and transmission line circuit length and public substation equipment capacities are 57.20 million km and 3.027 billion kilovolt-ampere (kVA), respectively, ranking the world's first in total grid size [11]. As pivotal equipment of the power system, the operation condition of the power transformer is related to the safe and stable operation of the power system. According to the statistics of the State Grid Corporation of China in 2008, the number of transformers facing insulation deterioration accounted for 42% of the total number of transformers in accidents. Insulation degradation is the main factor leading to the failure of high-voltage electrical equipment, among which partial discharge is one of the main causes of transformer insulation degradation. As a common insulation material for transformers, oil-impregnated paper insulation bears electrical stress as well as mechanical stress during operation [12]. During the manufacturing and operation of the paper insulation, various defects such as bubbles, cracks, and suspended conductive dots are inevitable. The long-term partial discharge makes these defects deteriorate and eventually lead to insulation breakdown, which endangers the stable operation of transformers and power systems. Figure 2 shows the electrical dendrite carbon marks on the insulation of the enclosing screen and leads in a real power transformer.

The partial discharge can effectively reflect the degree of local insulation deterioration of the transformer, and its signal implies comprehensive information about the internal insulation state of the transformer and the development degree of discharge defects. Therefore, the partial discharge detection and analysis of power transformers can be used to understand the insulation deterioration inside the transformer. The partial discharge can respond to sudden insulation problems and has the advantage of not being impacted by oil changes or other circumstances in the discharge detection process. With the continuous development of digital measurement systems for partial discharge and field anti-interference technology, the field of online detection of partial discharge will be widely used. As the partial discharge of different defects has a large difference in the destructive power of insulation, the partial discharge in different defects also has different discharge characteristics. By extracting the implied features from partial discharge signals of power transformers, the identification of transformer defect types can be realized, so that a more reasonable transformer operation, maintenance, and renewal plan can be formulated to provide an important guarantee for safe transformer operation [13]. Due to the complex mapping relationship between the information carried by partial discharge signals and partial discharges, it is difficult to obtain the differences between different types of partial discharges by human induction alone. Therefore, classification algorithms need to be

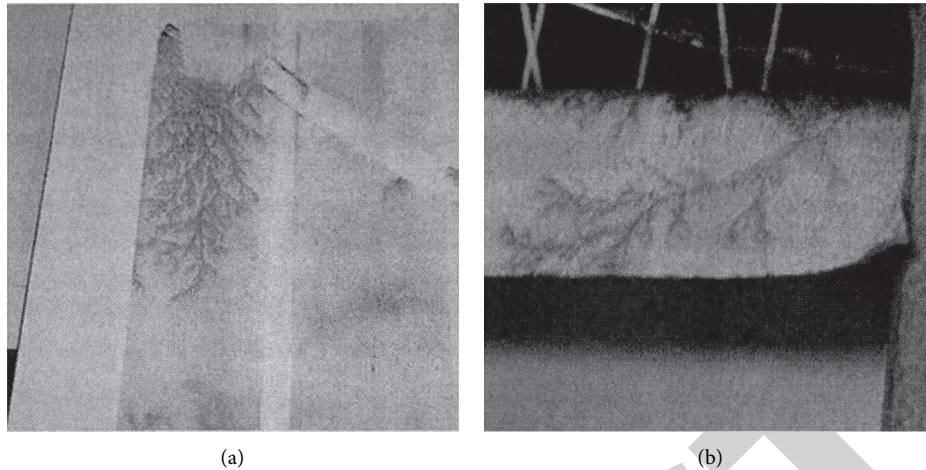(a)   (b)

FIGURE 2: Transformer insulation surface defects.

introduced for learning. The classifier algorithm has a direct impact on the effect of pattern recognition. Also, with the increase of transformer operation years and the complex internal structure, transformers often have partial discharge sources in multiple places at the same time. Therefore, the influence of multiple sources of partial discharges on pattern recognition should be taken into account when identifying.

At this stage, the development of China's power industry has entered a critical period of technological change, and the voltage level requirements have been increasing. At the same time, the voltage increase is also prone to the phenomenon of partial discharge of transformers, and the current generated by partial discharge and the surrounding media will react with each other to produce thermal effects or generate active substances, where the most important problem is that partial discharge will accelerate the aging of the insulator and reduce the insulation performance, which will lead to electrical accidents. Therefore, the optimization of transformer partial discharge detection technology is essential to effectively prevent accidents. The appearance of the partial discharge phenomenon makes the surrounding media form ultrasonic, high-frequency radiation and other effects, which also provides the direction for the upgrade of detection technology. This paper focuses on the partial discharge performance testing of power transformers and analyzes the application type and working principle of the testing technology [14], the development status, and future development trend, in order to provide ideas for the optimization of future testing technology. The power transformer is a key component in the normal operation of the power system, and its operating condition and equipment quality are directly related to the safety and stability of the whole power system. At the same time, the insulation state of power transformer directly affects the overall operation of the transformer, in which partial discharge produces a large number of physical and chemical effects of electricity, light, sound, heat, etc., which is the main cause of insulation aging and deformation of power transformers, which may cause different degrees of power accidents. In order to cope with the transformer operation problems caused by partial discharge, in recent years, relevant experts have developed

various discharge monitoring techniques combined with these effects, such as the electric pulse method, light detection side method, ultrasonic method, ultra-high-frequency method, gas chromatography method, infrared thermal imaging method, and so on, all of which are effectively applied in the partial discharge detection work to help the normal operation of the whole power engineering [15]. In Chinese electrical engineering, the partial discharge phenomenon is roughly divided into three types according to the difference of the types of discharge causes: they are the Townsend discharge, the injection discharge, and the discharge triggered by thermal ionization. In addition, there are various manifestations of discharge, and the small gap partial discharge phenomenon includes pulsed and nonpulsed discharges and subglow discharges. As the partial discharge phenomenon of transformers affects other surrounding substances, which in turn leads to the interaction between the equipment and the surrounding medium, this makes part of the insulator of the transformer react with each other (physicochemical effects, etc.), forming a partial discharge phenomenon. The occurrence of partial discharge may cause the emergence of ultrasonic waves and changes in the composition of the medium, etc., which may cause electrical accidents, resulting in serious consequences. In recent years, with the gradual increase in the number of electrical projects, the Chinese authorities have intensified their research work on partial discharges, aiming at studying more new technologies for discharge detection and strengthening the control of transformers.

2.2. Partial Discharge Pattern Recognition. Currently, the most commonly used classifier algorithms in the field of partial discharge pattern recognition include support vector machines as well as K-neighborhoods. Multikernel support vector machines enhance the kernel tuning capability by combining different kernel functions to form new kernels [16]. However, it also increases the number of tuning parameters. The four kernel functions are linearly combined to form new kernels, and the four discharge types are identified by solving the weights of each kernel function using a

particle swarm optimization algorithm. The results show that the multicore support vector machine can describe the local discharge information more comprehensively. The principle of the distance classifier is to calculate the distance between the sample to be predicted and the reference template, and the class corresponding to the smallest distance in each reference template is used as the prediction result. Among them, distance is an important parameter of this class of classifier, mainly Manhattan distance, Marxian distance, Euclidean distance, and Ming's distance. Different distance metrics are derived from different distance classifiers, including minimum distance classifier, nearest neighbor classifier, K-nearest neighbor (KNN) classifier, confidence interval classifier, and polynomial classifier. Distance classifier has the advantages of simplicity and efficiency; however, because it is a lazy learning method and requires high reliability of the reference template, the application effect is often not satisfactory. Its basic principle is that if a test sample is closest to K similar training samples in the feature space, then the test sample also belongs to this category, and it has achieved good application in the field of partial discharge pattern recognition. However, KNN is a lazy learning method, and it needs to calculate the distance between the new data and all training data when new data are tested, which is computationally intensive [17].

Random forest (RF) is an integrated algorithm based on statistical learning theory, whose model consists of a collection of categorical regression trees learned from a reconstructed training set, and the final prediction result is obtained by voting on each categorical regression tree. RF has high prediction accuracy, is not easy to be overfitted, and has wide applications in medical, bioinformatics, and management fields [18]. There are few studies on RF in the field of partial discharge pattern recognition, but a few papers have proved that RF has a good recognition effect as well as generalization ability in the field of partial discharge pattern recognition. The RF algorithm is used to identify the stages of air gap discharge development for different air gap sizes. The results show that the recognition accuracy of the RF algorithm is significantly better than the classification accuracy of neural networks and SVM. Firstly, 48 discharge patterns were classified according to the AC-DC voltage ratio, defect type, and severity, and feature extraction and selection were performed on the discharge signals. The selected feature quantities were used as input feature quantities for five classification methods, namely, repetitive clip, fuzzy recognition, RF, support vector machine (SVM), and backpropagation neural network (BPNN), and the 48 discharge patterns were identified directly and stepwise, respectively, and the results showed that RF had the highest recognition accuracy in both recognition modes. The EMD-SVD features were used as input feature quantities for the SVM classifier and RF classifier to identify the seven thermal aging stages of oil-paper insulation, and the results showed that the classification accuracy of RF was higher than that of SVM [19].

Support vector machine (SVM) can guarantee good generalization ability according to the statistical theory and the principle of structural risk minimization. It is widely used in the field of partial discharge pattern recognition because it can overcome the problems of small samples, dimensional disasters, local minima, and overfitting. Since the traditional single-kernel function SVM cannot realize the mapping classification of multiple feature space vectors in local discharge and most of the current SVMs use radial basis functions of different scales as kernel functions, their adjustment space capability is limited. With the deepening of SVM research [20], some SVM deformation algorithms have emerged: least-squares support vector machines and multicore support vector machines. The least-squares support vector machine converts the inequality constraint of SVM into equation constraint, which makes the parameter solution faster, but the recognition prediction accuracy is slightly worse than that of SVM. The least-squares support vector machine is used for pattern recognition of four defect types of partial discharges, and its advantages of fast convergence and good generalization ability under the condition of ensuring better recognition accuracy are verified.

2.3. Artificial Neural Networks.

An artificial neural network is a network combining several single neurons simulated by mathematical models, and the input and output weights and bias of each neuron are continuously adjusted by using learning data, and the training is stopped when the model reaches a certain accuracy, so as to obtain the mapping relationship between feature quantity and type. Neural networks can realize the function of any complex nonlinear mapping and have powerful self-organization, self-adaptation, and self-learning capabilities, which have wide applications in the field of partial discharge pattern recognition. BP neural networks (BPNNs) are multilayer feedforward networks with strong nonlinear approximation capability. Factor analysis is used to downscale the feature volume of the discharge data obtained from the accelerated aging test, and the processed features are input to the BPNN for learning to realize the recognition of the aging degree of the samples. The training results verified the effectiveness of the BPNN [21]. However, backpropagation (BP) neural network is sensitive to the selection of initial values, and the algorithm requires long training time and is easy to fall into local minima. In order to overcome the drawback that BPNN may get local optimum, the genetic algorithm is used to perform a global search for the parameters of BPNN, determine the initial value of each parameter, and then use BPNN to perform local optimization search, so as to determine the final network model. The results show that using a combination of the two algorithms can reduce the training time of the BPNN.

Probabilistic neural networks are developed by combining statistical methods with BPNNs, based on Parzen window probability density estimation and Bayesian maximum posterior probability criterion. Probabilistic neural networks have no iterative process and are easy to remodel with newly added samples. Since the partial discharge monitoring process is susceptible to random noise and requires a certain degree of real time, in principle,

probabilistic neural networks meet the requirements of partial discharge pattern recognition [22]. The pattern recognition of two types of partial discharges using probabilistic neural networks is compared with BPNN, and the results demonstrate the effectiveness of probabilistic neural networks. The smoothing factor in the probabilistic neural network affects the degree of correlation between samples and has a large impact on the performance of the network, so a certain search is required to determine the parameter values. The literature uses principal component analysis to reduce the dimensionality of the partial discharge input features, models the reduced samples using the probabilistic neural network, and initially confirms the value of the smoothing factor by taking values at certain intervals, thus improving the accuracy of the partial discharge pattern recognition. The literature uses a genetic algorithm to optimize the smoothing factor and uses the final model to identify the four discharge models [23], which proves that the processing method can effectively determine the appropriate model parameters and improve the identification capability. Since the various methods mentioned above are based on artificially extracted features of partial discharges as recognition information, they cannot fully express the information contained in the partial discharge signals. The network structure of the deep learning network and BPNN is similar, but the difference is that deep learning introduces the "greedy layer-by-layer pretraining" mechanism to determine the initial value of each layer neuron parameter, which can automatically extract features from the original input data.

The convolutional neural network in deep learning is used to learn and recognize the processed partial discharge time-domain signal images. The results show that the recognition ability of the convolutional neural network is higher than that of both the traditional neural network and support vector machine [24]. However, since the scale of the convolutional neural network model is much larger than that of the BPNN model, the combined effectiveness of the convolutional neural network is more powerful, and the network model shows better robustness when new data are added to the test data.

3. Method

The proposed model combines the advantages of CNN, which is good at mining local spatial information of partial discharge pattern, and LSTM, which is good at mining temporal information of partial discharge pattern and can extract both local spatial features and temporal features of partial discharge pattern.

3.1. Convolutional Neural Networks. As a multilayer neural network recognition algorithm widely used in image recognition, a convolutional neural network (CNN) is a feedforward neural network with the main features of local perception, network weight sharing, and multiple convolutional kernels [25]. A typical convolutional neural network usually contains 5 layers, namely, input layer,

convolutional layer, pooling layer, fully connected layer, and output layer (see Figure 3). The convolutional layer consists of multiple eigenfaces, each of which contains multiple neurons. During the feedforward period, the convolutional kernel performs a convolution operation on the input to extract the set of features on the local area. Assuming that the feature vector at the output of the i-th convolutional layer is Hi, there exists

$$H_i = f\left(H_{i-1} \otimes w_i + b_i\right), \tag{1}$$

where $f(\bullet)$ is the activation function and w_i and b_i are the weight vector and offset of the i-th convolutional layer, respectively. As shown in Figure 3, the pooling layer also consists of multiple eigenfaces, and each eigenface corresponds to the eigenface of the convolutional layer. The neurons in the fully connected layer are connected to the eigenfaces of the previous layer, extracting the information with classification features from the eigenfaces of the previous layer and inputting the output to the output layer after linear/nonlinear operations. The role of the output layer is to map the unnormalized output of the fully connected layer to the probability distribution over the predicted output categories.

3.2. LSTM Deep Learning Algorithm. Scholars working on deep learning have proposed an improved recurrent neural network by adding a gating (cell) structure to the recurrent neural network, which is the long short-term memory network (LSTM), to solve the problem that recurrent neural networks cannot obtain the temporal feature information of long interval input. The gating structure mainly includes three gates. The forgetting gate decides how much data are discarded, the input gate determines how much data are input, and the output gate determines how much data are output [26]. Because there might be lags of undetermined duration between critical occurrences in a time series, LSTM networks are well suited to categorizing, processing, and making predictions based on time series data. LSTM networks were created to solve the problem of vanishing gradients that can occur when training traditional RNNs. The structure of the gating unit of the long and short-term memory network is shown in Figure 4. In Figure 4, the symbol $\otimes$ indicates the multiplication of vector elements; the symbol $\oplus$ indicates the splicing of vectors; x_t is the input of the gating unit at the current moment; h_{t-1} indicates the output of the gating unit at the adjacent previous moment; c_{t-1} is the state of the gating unit at the adjacent previous moment; c_t is the state of the gating unit at the current moment; h_t is the state of the gating unit at the current output of the gating unit at the current moment; σ denotes the function sigmoid($\bullet$); and tanh denotes the function tanh($\bullet$).

3.3. Model Architecture. The proposed model is a combination of convolutional neural network (CNN) and long and short-term memory network (LSTM), combining the advantages of both: extracting the spatial features of the partial
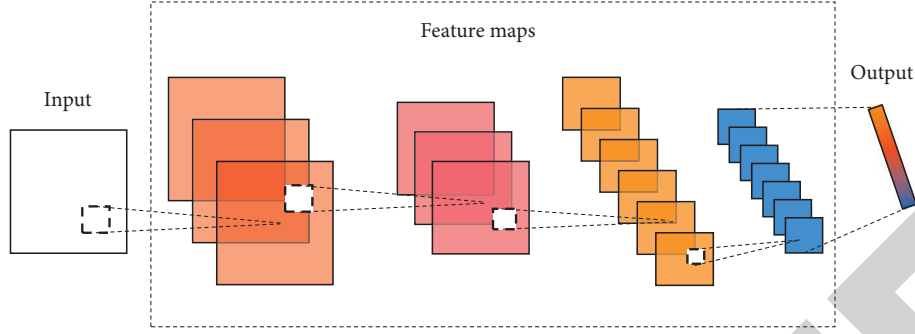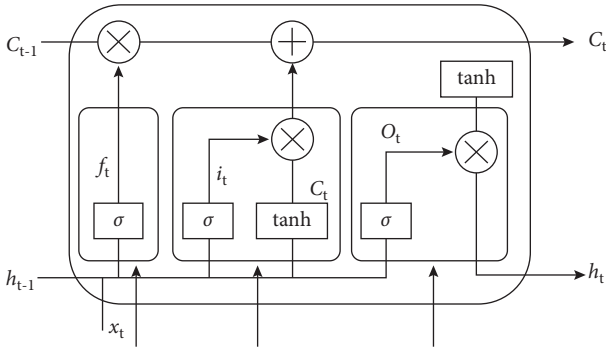
Figure 3: Typical convolutional neural network.

Figure 4: LSTM cell unit structure.

discharge map by taking advantage of the advantages of convolutional and neural networks in extracting the local features of the network input and extracting the temporal features on the partial discharge map at different moments by taking advantage of the advantages of long and short-term memory networks in extracting the temporal features between the inputs at different moments. Finally, pattern recognition is performed by the softmax classifier. The network architecture of the proposed model deep learning algorithm is shown in Figure 5.

In the CNN processing layer, what needs to be extracted is the spatial information of the input atlas, and it is needed to generate the corresponding feature vectors [27]. Let N be the number of the input partial discharge atlas, and for the input single atlas, assume that the pixel matrix of the atlas is $m \times n$ and the size of the convolution kernel is $k \times k$. Convolution is performed in the same way, and the periphery of the input image is filled with zero elements, and the fill length is k_2, and the image matrix is expanded from size $m \times n$ to $(m + k) \times (n + k)$. Let v_{ij} be the pixel value of the i-th row and j-th column of the expanded input image; then, the elements of the intra-window matrix with v_{ij} as the top corner element of the window of the convolution kernel field are

$$Xij = \begin{cases} V_{ij} & \cdots & V_{i(j-k-1)}, \\ \vdots & \vdots, \\ V_{(i-k-1)(j)} & \cdots & V_{(i+k-1)(j+k-1)}. \end{cases} \quad (2)$$

Convolution operation based on (1) for the window matrix X_{ij} is

$$Y_{ij} = f\left(X_{ij} \otimes w_i + b_i\right). \quad (3)$$

The activation function in (3) uses the ReLu function:

$$f(x) = \max(0, x). \quad (4)$$

The maximum pooling method is used to pool the output Y_{ij} of the convolution layer after the convolution operation of the input mapping is completed, and the maximum feature value within the pooling window is obtained to achieve the secondary optimization of the feature set.

$$R_n = \max\left(Y_{ij}\right). \quad (5)$$

After the feature extraction of all the input local discharge profiles, the feature matrix extracted by the convolutional neural network layer is

$$R = (R_1, R_2, \ldots, R_N). \quad (6)$$

Since the input set of local discharge profiles is a profile that fluctuates to some extent with time, assuming that the n-th input profile corresponds to the acquisition moment t, R_n can be used as the input to the long and short-term memory network at time t. Denoting R_n by R_t, the cell update process of the corresponding long and short-term memory network is

$$f_t = \sigma\left(W_f \cdot [R_t \cdot x_t] + b_f\right), \quad (7)$$

$$i_t = \sigma\left(W_i \cdot [R_t \cdot x_t] + b_i\right), \quad (8)$$

$$o_t = \sigma\left(W_o \cdot [R_t \cdot x_t] + b_o\right), \quad (9)$$

$$C_t = \sigma\left(W_c \cdot [R_t \cdot x_t] + b_c\right), \quad (10)$$

$$c_t = f_t \otimes c_{t-1} + i_t \otimes C_t, \quad (11)$$

$$h_t = o_t \otimes \tanh(c_t), \quad (12)$$

where W_f, W_i, and W_o are the weight matrices of the forgetting gates, input gates, and output gates, respectively; b_f, b_i, and b_0 are the bias terms of the forgetting gates, input gates, and output gates, respectively; C_t is the state of the input unit; and W_c and b_c are the weight matrices and bias terms of the state of the input unit, respectively.
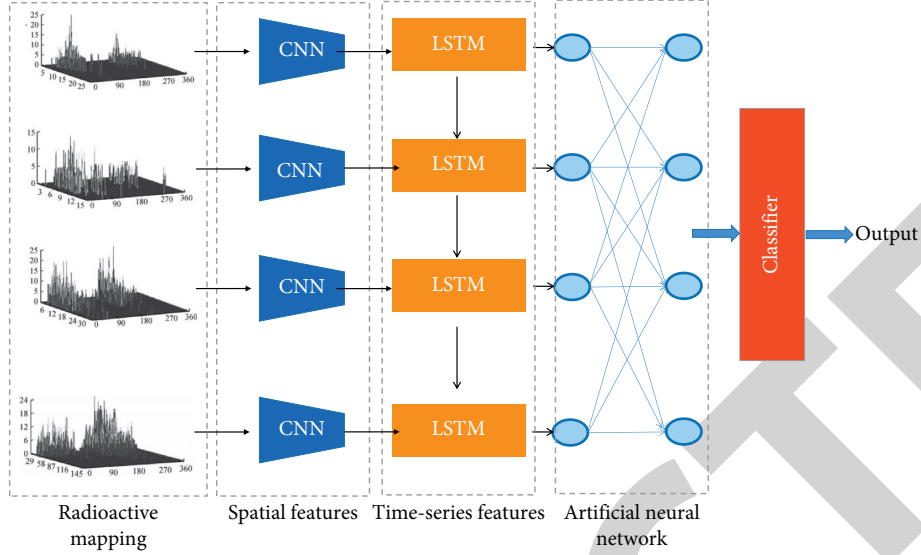
Figure 5: Network structure diagram.

3.4. Algorithm Flow. The partial discharge pattern recognition process based on an artificial neural network for transformer gap leakage electromagnetic wave is shown in Figure 6. Firstly, the partial discharge mapping set coupled to the ultra-high-frequency (UHF) antenna sensor is divided into two parts: a training set and test set, and then the partial discharge mapping is preprocessed to ensure the normalization of the pattern recognition network input. Finally, the grayscale matrix of the preprocessed training set is input to the proposed pattern recognition network to be trained, and the training process is backpropagated using the gradient descent method, and the training ends when the error value between the network output and the expected value is less than or equal to the expected value. The training results are validated.

4. Experimentation and Evaluation

4.1. Experimental Setup. Table 1 shows the major parameters for the proposed model hybrid network. The experimental environment of this article is as follows: the hardware environment is Linux system, NVIDIA GTX 2080Ti; the software environment is Python3.5, sklearn0.20.3, and other toolkits. The recognition object is the partial discharge mapping of four types of typical insulation defects in power transformers: metal protrusion defects, oil-paper air gap defects, oil-paper discharge defects along the surface, and suspended potential defects. For 240 sets of discharge data, 200 sets were selected for training and 40 sets were selected for testing. Firstly, 10 sets of 200 training data are selected as the initial training sample set (the ratio of 4 types of discharge samples is $2:3:3:2$ to ensure that each type of discharge sample is included in the initial training sample set), and an artificial neural network-based local discharge pattern recognition classifier is constructed.

4.2. Performance Comparison Results. The parameters in Table 1 are used to create the suggested model network, the number of convolutional layers is 5, the number of pooling

layers is 5, the number of layers of LSTM is 1, and the rest of the parameters are the system default. The grayscale processed partial discharge mapping's single-channel matrix data are fed into the network for training and recognition, and the overall prediction accuracy P_r is used as the evaluation parameter of the pattern recognition ability of the proposed model network, and P_r is calculated as

$$P_r = \frac{N_p}{N_{\text{sum}}}, \tag{13}$$

where N_p is the number of samples where the predicted local discharge type and the actual discharge type match and N_{sum} is the total number of samples. Table 2 shows the proposed network's identification results for various types of typical power transformer insulation defects. As can be seen, the proposed network has the highest accurate identification rate of 100 percent for suspension potential insulation defects and the lowest identification rate of 89 percent for local discharge insulation defects along the surface.

To compare the pattern recognition performance of the proposed network, CNN, and LSTM network for typical insulation defects of partial discharge of power transformers, the same partial discharge mapping data are input to the convolutional neural network (CNN) and the long short-term memory network (LSTM network), where the parameter settings of CNN and LSTM network are the same as those of the proposed network, and the remaining parameters are the system default. Table 3 shows that the proposed network outperforms the CNN and LSTM network in terms of partial discharge pattern recognition for metal protrusion defects, oil-paper air gap defects, and along-surface discharge defects, but for suspension potential defects, the proposed network has the same recognition effect as CNN, which is 100%, and this effect corresponds to the fluctuation of partial discharge patterns of different types of insulation. The stability of suspended potential defect discharge profiles is significantly better than that of other types of discharge profiles, and the advantage of the LSTM network is the
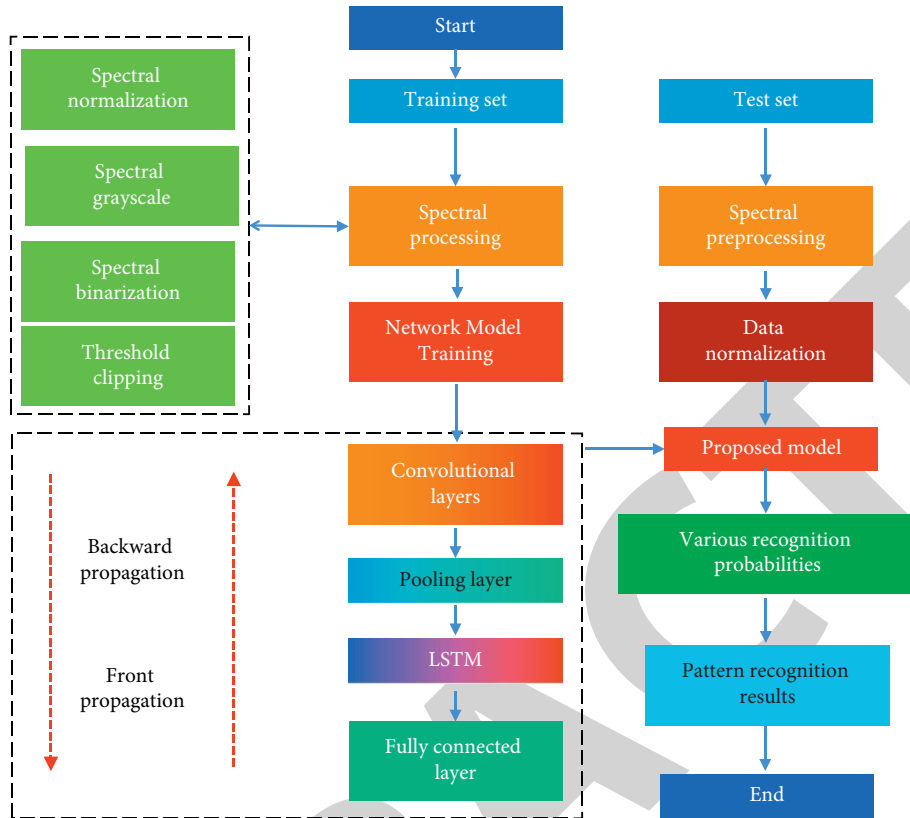
FIGURE 6: Algorithm flowchart.

TABLE 1: Model parameter setting.

Test items	Parameter setting
Convolution kernel size and step size	(5 * 5, 1)
Pooling method	Max-pooling
Pooling window size and step size	(2 * 2, 1)
Fill mode	Same mode
Activation function	ReLu
Weights	Random normal distribution
Number of LSTM nodes	100
Optimization function	SGD
Learning rate	0.001
Batch size	64
Epoch	100

extraction of temporal feature information, so the LSTM network does not provide much help in the proposed network for pattern recognition of suspended potential defect types, but the discharge profiles of metal protrusion defects, oil-paper air gap defects, and along-surface discharge defects are significantly better than those of other types of discharge profiles, and the advantage of LSTM network in extracting the temporal feature information of the profiles gains manifestation and improves the recognition accuracy of the CNN for these defect types.

Figure 7 shows the variation curves of recognition rate obtained by using the model proposed in this paper and by learning with CNN and LSTM models. Figure 7 shows that

the method described in this paper requires 70 epochs of training to get a high sample recognition rate, whereas CNN and LSTM require 85 epochs of training to reach the same recognition result. As demonstrated in Figure 7, the proposed model requires significantly less learning time than the CNN and LSTM models to achieve the same recognition accuracy. The above results show that the learning cost of the proposed model is much smaller than that of the CNN and LSTM model learning methods. Further study shows that the suggested method outperforms CNN and LSTM in terms of ultimate convergence performance, demonstrating the method's innovation and effectiveness.

5. Discussion

In this section, we discuss all the equations and all the figures in detail. In (1), $f(\bullet)$ is the activation function and w_i and b_i are the weight vector and offset of the i-th convolutional layer, respectively. As shown in Figure 3, the pooling layer also consists of multiple eigenfaces, and each eigenface corresponds to the eigenface of the convolutional layer. The structure of the gating unit of the long and short-term memory network is shown in Figure 4. The network architecture of the proposed model deep learning algorithm is shown in Figure 5. In the CNN processing layer, we represent the pixel matrix of the atlas as $m \times n$ and the size of the convolution kernel as $k \times k$; also, v_{ij} is the pixel value of the i-th row and j-th column of the expanded input image. Since the input set of local discharge profiles is a profile that

TABLE 2: Detection results of different types of defects.

P_r (%)	Metal protrusion defects	Oil-stained air gap defects	Surface discharge defects	Suspension potential defects
Metal protrusion defects	94	4	2	0
Oil-stained air gap defects	6	90	3	1
Surface discharge defects	5	6	89	0
Suspension potential defects	0	0	0	100

TABLE 3: Performance comparison.

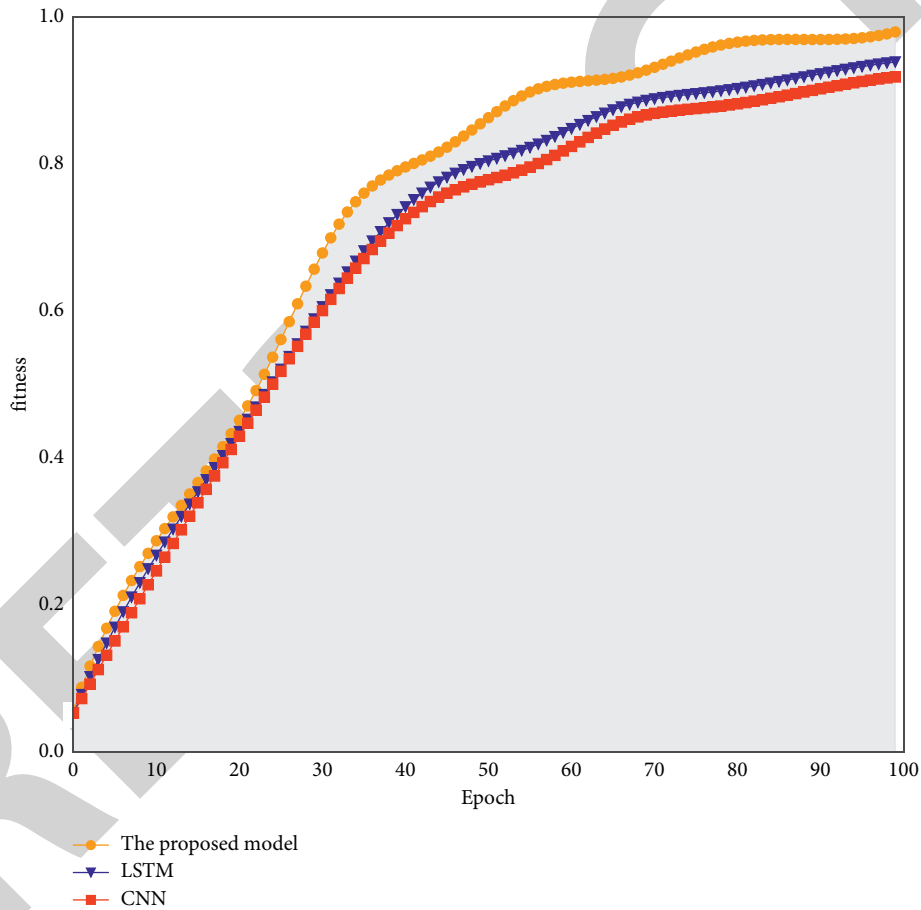| P_r (%)	Metal protrusion defects	Oil-stained air gap defects	Surface discharge defects	Suspension potential defects
CNN	83	80	85	100
LSTM	80	78	82	98
The proposed model	94	80	89	100

FIGURE 7: Model training process.

fluctuates to some extent with time, assuming that the $n-$th input profile corresponds to the acquisition moment t, R_n can be used as the input to the long and short-term memory network at time t. Denote R_n by R_t. In equations (7)–(12), W_f, W_i, and W_o are the weight matrices of the forgetting gates, input gates, and output gates, respectively; b_f, b_i, and b_0 are the bias terms of the forgetting gates, input gates, and output gates, respectively; C_t is the state of the input unit; and W_c and b_c are the weight matrices and bias terms of the state of

the input unit, respectively. In (13), N_p is the number of samples where the predicted local discharge type and the actual discharge type match and N_{sum} is the total number of samples. Table 2 shows the proposed network's identification results for various types of typical power transformer insulation defects. As can be seen, the proposed network has the highest accurate identification rate of 100 percent for suspension potential insulation defects and the lowest identification rate of 89 percent for local discharge insulation

defects along the surface. Table 3 shows that the proposed network outperforms the CNN and LSTM network in terms of partial discharge pattern recognition for metal protrusion defects.

6. Conclusion

The power transformer is one of the important pieces of equipment in the power system, and its health condition is directly related to the safe operation of the whole power system. With the development of the power system and the increase of voltage level, the partial discharge has become one of the main causes of transformer insulation deterioration, and effective detection of partial discharge is one of the most effective means to prevent sudden transformer insulation accidents. Different types of partial discharges have different degrees of harm to power transformers, and the maintenance measures taken are different, so the identification of the type of partial discharge signal is particularly important. In this paper, according to the problems faced by transformer partial discharge signal denoising and discharge pattern recognition in power system, deep learning with better recognition effect is combined with pattern recognition, and the denoised signal is input into an artificial neural network for pattern recognition to determine the transformer partial discharge type. The efficient and rapid identification of the type and severity of partial discharge defects in transformers is beneficial to the timely detection of internal latent faults and provides a basis for the formulation of maintenance strategies to ensure the safety and stability of their operation.

The algorithm integrates the advantages of CNN and LSTM, which is good at mining local spatial information of partial discharge pattern, which is good at mining temporal feature information of partial discharge pattern. Then, the extracted feature information is input to the LSTM network to extract all the partial discharge pattern recognition features including the temporal feature information, and finally, the pattern recognition of partial discharge insulation defects is realized by the fully connected layer and softmax classifier. The findings show that the suggested algorithm is capable of recognizing partial discharge insulation flaws in a pattern; the pattern recognition ability of CNN, LSTM, and the proposed network is compared for different partial discharge insulation defects. The proposed method is better than CNN and LSTM in terms of recognition ability. It is demonstrated that the proposed method has leading performance and can be used for real power system detection. In the future, we plan to carry out the use of dynamic neural networks and attention mechanism networks to further improve the performance of transformer partial discharge pattern recognition.

Data Availability

The datasets used during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] G.-M. Ma, H.-Y. Zhou, and C. Shi, "Distributed partial discharge detection in a power transformer based on phase-shifted FBG," *IEEE Sensors Journal*, vol. 18, no. 7, pp. 2788–2795, 2018.

[2] M. Kunicki, A. Cichoń, and S. Borucki, "Measurements on partial discharge in on-site operating power transformer: a case study," *IET Generation, Transmission & Distribution*, vol. 12, no. 10, pp. 2487–2495, 2018.

[3] C. Gao, L. Yu, Y. Xu, W. Wang, S. Wang, and P. Wang, "Partial discharge localization inside transformer windings via fiber-optic acoustic sensor array," *IEEE Transactions on Power Delivery*, vol. 34, no. 4, pp. 1251–1260, 2018.

[4] D. Dukanac, "Application of UHF method for partial discharge source location in power transformers," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 25, no. 6, pp. 2266–2278, 2018.

[5] J. Jia, C. Hu, Q. Yang, Y. Lu, B. Wang, and H. Zhao, "Localization of partial discharge in electrical transformer considering multimedia refraction and diffraction," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 8, pp. 5260–5269, 2020.

[6] S. Qian, H. Chen, Y. Xu, and L. Su, "High sensitivity detection of partial discharge acoustic emission within power transformer by sagnac fiber optic sensor," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 25, no. 6, pp. 2313–2320, 2018.

[7] M. Siegel, S. Coenen, M. Beltle et al., "Calibration proposal for UHF partial discharge measurements at power transformers," *Energies*, vol. 12, no. 16, Article ID 3058, 2019.

[8] M. Mondal and G. B. Kumbhar, "Detection, measurement, and classification of partial discharge in a power transformer: methods, trends, and future research," *IETE Technical Review*, vol. 35, no. 5, pp. 483–493, 2018.

[9] Y. B. Wang, D. G. Chang, Y. H. Fan et al., "Acoustic localization of partial discharge sources in power transformers using a particle-swarm-optimization-route-searching algorithm," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 24, no. 6, pp. 3647–3656, 2017.

[10] H. Y. Zhou, G. M. Ma, M. Zhang, H. C. Zhang, and C. R. Li, "A high sensitivity optical fiber interferometer sensor for acoustic emission detection of partial discharge in power transformer," *IEEE Sensors Journal*, vol. 21, no. 1, pp. 24–32, 2019.

[11] X. Zhang, G. Zhang, Y. Li, J. Zhang, and R. Huang, "On the feasibility of gap detection of power transformer partial discharge UHF signals: gap propagation characteristics of electromagnetic waves," *Energies*, vol. 10, no. 10, Article ID 1531, 2017.

[12] X. Hu, W. H. Siew, M. D. Judd, A. J. Reid, and B. Sheng, "Modeling of high-frequency current transformer based partial discharge detection in high-voltage cables," *IEEE Transactions on Power Delivery*, vol. 34, no. 4, pp. 1549–1556, 2019.

[13] J. Li, X. Zhang, X. Han, and X. Yao, "A partial discharge detection method for SF 6 insulated inverted current transformers adopting inner shield case as UHF sensor," *IEEE Transactions on Power Delivery*, vol. 33, no. 6, pp. 3237–3239, 2018.

[14] S. A. Ward, A. Faraskoury, M. Badawi et al., "Towards precise interpretation of oil transformers via novel combined

techniques based on DGA and partial discharge sensors," *Sensors*, vol. 21, no. 6, Article ID 2223, 2021.

[15] H. Chai, B. T. Phung, and D. Zhang, "Development of UHF sensors for partial discharge detection in power transformer," in *Proceedings of the 2018 Condition Monitoring and Diagnosis (CMD)*, pp. 1–5, IEEE, Perth, WA, Australia, September 2018.

[16] W. Sikorski, "Development of acoustic emission sensor optimized for partial discharge monitoring in power transformers," *Sensors*, vol. 19, no. 8, 2019.

[17] S. G. Ha, J. Cho, J. Lee, B. W. Min, J. Choi, and K. Y. Jung, "Numerical study of estimating the arrival time of UHF signals for partial discharge localization in a power transformer," *Journal of electromagnetic engineering and science*, vol. 18, no. 2, pp. 94–100, 2018.

[18] L. Zhou, J. Cai, J. Hu, G. Lang, L. Guo, and L. Wei, "A correction-iteration method for partial discharge localization in transformer based on acoustic measurement," *IEEE Transactions on Power Delivery*, vol. 36, no. 3, pp. 1571–1581, 2020.

[19] W. Sikorski, K. Walczak, W. Gil, and C. Szymczak, "On-Line partial discharge monitoring system for power transformers based on the simultaneous detection of high frequency, ultra-high frequency, and acoustic emission signals," *Energies*, vol. 13, no. 12, Article ID 3271, 2020.

[20] S. Kanakambaran, R. Sarathi, and B. Srinivasan, "Identification and localization of partial discharge in transformer insulation adopting cross recurrence plot analysis of acoustic signals detected using fiber Bragg gratings," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 24, no. 3, pp. 1773–1780, 2017.

[21] L. A. M. M. Nobrega, G. V. R. Xavier, M. V. D. Aquino, A. Serres, C. C. R. Albuquerque, and E. Costa, "Design and development of a bio-inspired UHF sensor for partial discharge detection in power transformers," *Sensors*, vol. 19, no. 3, p. 653, 2019.

[22] C. B. A. De, M. B. D. De, F. G. Baptista, A. L. Andreoli, and J. A. C. Ulson, "Assessment of macro fiber composite sensors for measurement of acoustic partial discharge signals in power transformers," *IEEE Sensors Journal*, vol. 17, no. 18, pp. 6090–6099, 2017.

[23] C. Pan, J. Tang, X. Tao, Y. Zhang, and S. Ma, "Partial discharge and breakdown characteristics of moving transformer oil contaminated by metallic particles," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 25, no. 5, pp. 1774–1784, 2018.

[24] N. Xue, J. Yang, D. Shen et al., "The location of partial discharge sources inside power transformers based on TDOA database with UHF sensors," *IEEE Access*, vol. 7, Article ID 146732, 2019.

[25] G. Zhang, X. Zhang, J. Tang, and H. Cheng, "Study on localization of transformer partial discharge source with planar arrangement UHF sensors based on singular value elimination," *AIP Advances*, vol. 8, no. 10, Article ID 105232, 2018.

[26] X. Zhou, X. Wu, P. Ding et al., "Research on transformer partial discharge UHF pattern recognition based on CNN-lSTM," *Energies*, vol. 13, no. 1, p. 61, 2020.

[27] H. Karami, M. Azadifar, A. Mostajabi et al., "Partial discharge localization using time reversal: application to power transformers," *Sensors*, vol. 20, no. 5, Article ID 1419, 2020.