Hilbert ID Considering Multi-Window Feature Extraction for Transformer Deep Vision Fault Positioning

XIAOXIN WU, YIGANG HE, (Member, IEEE), CHENYUAN WANG, WENJIE WU, CHUANKUN WANG, AND JIAJUN DUAN
School of Electrical Engineering and Automation, Wuhan University, Wuhan 430072, China
Corresponding author: Yigang He (yghe1221@whu.edu.cn)

ABSTRACT Most of the power transformer fault diagnostic researches so far focuses on its fault type diagnosis, but there are less related researches on fault positioning, and the diagnostic methods of which are still less intelligent. This paper proposes a two-dimensional Hilbert ID considering multi-window feature extraction for deep vision fault positioning of the transformer winding. Firstly, sweep frequency response data containing complex fault characteristics is obtained based on pspice simulation. Next, a multi-window feature extraction method with logarithmic constraints is introduced to process the original data to obtain feature sequences. Then the proposed Hilbert visualization is used to further highlight the graphic feature of the feature sequences, and obtain Hilbert ID (MAPE) dataset. Finally, it is used to conduct transfer learning on the convolutional neural network. Different intelligent positioning methods are compared, and the proposed deep vision fault positioning method is 6.51% higher than other methods on average. What’s more, the positioning effects based on different data processing methods are also compared. The accuracy of the proposed Hibert ID (MAPE) dataset is 10.35% higher than the other data processing methods on average. Finally, the positioning accuracy of Hilbert ID (MAPE+CC) combining two feature sequences can reach 96.09%, having an increase of 2.50%.

INDEX TERMS Convolutional neural network (CNN), deep transfer learning (DTL), fault positioning, Hilbert visualization, multi-window feature extraction, power transformer, sweep frequency response analysis (SFRA).

I. INTRODUCTION Fault diagnosis of key power equipment such as power transformer is an important part of power systems for its safe and economic operation [1], [2]. With the further construction of smart grid, power equipment will continue to develop in the direction of intelligence and high integration. However, the current power transformer fault diagnostic methods still have relatively weak ability to extract transformer fault features, so the diagnostic effect is not satisfactory, and can’t meet the needs of smart grid construction. Most studies focus on power transformer fault type diagnosis, while there are fewer related studies on fault positioning methods [3]–[5].

Transformer fault diagnostic methods include Dissolved Gas Analysis (DGA) [6], Frequency Response Analysis (FRA) [7], Infrared Thermography (IRT) [8] and so on. Sweep Frequency Response Analysis (SFRA), as one of FRA method, has higher sensitivity for winding fault diagnosis [9], [10]. It uses transfer function to obtain the Frequency Response (FR) curve to provide information for diagnosis. Currently, in the industrial field, the interpretation of the FR curve is mainly performed by skilled engineers who have extensive experience in transformer detection and maintenance [11]. The FR curve of the standard, and healthy transformer can be regarded as the original characteristics of
the transformer, which is often called fingerprint [12]. And faults such as winding axial deformation, radial deformation, and inter-turn short circuits may cause changes of parameters in the equivalent circuit of the transformer windings, thus cause deformation of the FR fingerprint. Comparing the FR curve of a transformer in a fault state with its fingerprint can diagnose a specific fault type [13], [14]. In this case, the experience and professional ability of the diagnostic expert is critical. Because it requires strong prior knowledge of the impact of each parameter on the FR curve. This can’t meet the developing trend of the smart substation. In order to overcome this limitation, many studies have focused on improving the interpretability of the FR curve through calculation techniques. Many statistical indexes(SI) [15], such as Correlation Coefficient (CC), Mean Squared Error (MSE), Sum Squared Max–Min Ratio Error (SSMMRE), etc. are introduced in FRA-based diagnosis. However, most SI methods compress the characteristics of the FR curve to a single number, and has limited reflection on the characteristics of the FR curve which results in the unsatisfactory diagnostic accuracies. There is also reference that piecewise calculate the SI to extract more local features of the curve and finally achieve fault type determination and approximate positioning, but the fault location is still relatively broad and less accurate [16]. In order to further extract the characteristics of the FR curve and improve the diagnostic accuracy, some machine learning models [17]–[19] have also been introduced into FRA-based transformer fault diagnosis. For example, reference [19] introduces Support Vector Machine (SVM) to transformer FRA-based diagnosis. It diagnoses three types of transformer fault states, and achieved good results. However, SVM is essentially a two-classifier. With the number of fault types increasing, it faces a lot of diagnostic challenges.

Some data visualization methods have also been introduced into the transformer FRA diagnostic field [20]–[23]. Performing wavelet transform on Pulse Frequency Response Analysis (PFRA) data, reference [23] compares the wavelet time-frequency spectrum image of fault states with that of the normal state. And the rules of image difference of different transformer state are summarized. However, the diagnosis of this kind of visualization method is intrinsically limited to visual comparison, and has not been combined with the intelligent methods so as to further improve the intelligence of diagnosis.

At present, only a few studies involve in the fault positioning of transformer windings. If the fault location can be more accurately located, it will greatly reduce the workload of transformer maintenance and improve economic efficiency. In addition, the existing researches on the fault positioning methods often ignore the complex and diverse types of faults when setting faults, and only select a few specific faults without fully considering the severity of the transformer fault [24]–[26].

In short, the current researches on FRA-based transformer fault diagnosis have the following problems: 1) there are less researches on fault positioning, and the diagnostic methods are not intelligent enough; 2) the severity and diversity of faults at the same fault location have not been fully considered; 3) the feature extraction of FR curve is still limited to traditional SI methods, while the data visualization techniques used by a few studies lack an in-depth feature extraction process.

Deep Learning (DL) which is an important artificial intelligence method [27], [28], can realize automatic identification and fast processing of abnormal states [29], [30]. As one of DL techniques, Convolutional Neural Network (CNN) which has shown significant advantages in the field of image processing [31], [32], can extract the deep-level features of the images. However, CNN usually has relatively large parameters which requires a large amount of data to train. As a kind of machine learning method, Transfer Learning (TL) can transfer the knowledge learned in domain B to knowledge of the target domain A which has not been known, thereby effectively reducing the dataset of the domain A required for training [33] and improving accuracy. Combining TL with CNN, we get Deep Transfer Learning (DTL), which can still effectively extract complex data features and get nice results when the amount of training data is not large enough [34], [35]. Using the deep networks to diagnose the FR curve feature data can finally improve the final diagnostic result and intelligence level.

The data visualization method intuitively condenses a large amount of data onto a two-dimensional image, which effectivly saves information storage space and highlights the graphic features, conducive to identification. However, in the current researches, the purpose of the visualization of the FR data is often to visually see the change level of the FR curve, rather than as an input to an intelligent diagnostic deep network. Therefore, it is common in the existing references to directly draw the FR curve [36], [37]. But such kind of images often have the majorities of space remain unused. Inputting such kind of images into CNN for training, the subtle features of the FR curve in the images can’t be well recognized. Therefore, it is necessary to creatively obtain an image different from the conventional FR fingerprint by using a two-dimensional filling curve to highlight the graphic characteristics of the data. It’s a new idea which is an inspiration from the known knowledge to illustrate the FRA data of the old question [38].

This paper proposes a transformer winding fault positioning method based on multi-window feature extraction, two-dimensional Hilbert ID and DTL model. The flowchart of the positioning method is shown as Fig. 1. First, the equivalent ladder network model of the transformer winding is constructed using pspice simulation. Next, the model’s own errors and different fault locations, fault types and fault severities are comprehensively considered in the transformer winding model to obtain the original FR dataset containing complex fault characteristics. Further, a multi-window feature extraction method is introduced to process the original FR dataset and obtain the FR feature sequence dataset. And then the feature sequence dataset is converted into a Hilbert
ID dataset using the Hilbert visualization theory. Finally, the dataset is input into the deep diagnostic network to conduct the transfer training and testing. And the positioning effect is compared with various methods, and the proposed positioning method is further improved.

The rest of the article is organized as follows. Section II deliberates the principle and innovation of the proposed positioning method; Section III illustrates the specific implementations of the proposed method; Section IV compares the proposed methods with various positioning methods and data processing methods; Section V further improves the proposed method; Section VI finally draws the conclusion.

II. PROPOSED METHOD

A. THE PRINCIPE OF DATASET ACQUISITION

The transformer winding can be regarded as a linear circuit model which consists of components such as resistors, inductors and capacitors, in the high frequency range (>1kHz). When a frequency-varying excitation power signal is applied at the beginning of the winding, the response signal measured at the end of the winding will change with the frequency. The FR data of the winding can be obtained by measuring the input signal and the response signal, and calculating according to the transfer function as shown in formula (1). Winding faults caused due to electrical, thermal, mechanical and other reasons will cause changes of the components in the equivalent circuit such as inductors and capacitors and finally result in changes of the FR curve. For example, inductive faults in the winding are usually caused by short turns. Because the inductance of the coil is directly proportional to the number of turns in the coil, a short circuit in the coil reduces the inductance of the equivalent circuit. Therefore, by comparing the measured FR curve with the FR fingerprint, the winding deformation fault can be diagnosed and identified.

$$H_f = 20 \lg \left| \frac{U_2(f)}{U_1(f)} \right|$$  \hspace{1cm} (1)

where $U_1(f)$ denotes the input signal; $U_2(f)$ denotes the response signal.

This paper validates the proposed positioning method based on a transformer ladder network simulation model that ignores high-order mutual inductance. The ladder network naturally establishes a mapping between the physical winding of the transformer and the simulated part, which is conducive to study the positioning of the deformed windings. Changing parameters of different components in the ladder network units can simulate different severity and types of faults at different winding locations [39] and obtain FR data for exploring the feasibility of the proposed deep vision fault positioning method. Pspice software is used to perform simulations of various fault and normal conditions of the transformer winding. The construction and related parameters of the transformer winding circuit model refers to reference [40]. The model is as shown in Figure 2. The equivalent circuit model of the transformer winding is composed of $i$ units (in this article $i = 1, 2, \cdots, N_{\text{area}}, N_{\text{area}} = 7$). Each unit is composed of coil inductance $L_i$, mutual inductance $M_i$, capacitance to ground $C_{gi}$, capacitance $C_{si}$ between coils, and resistance $R_i$. Because the influence of $R_i$ is small, this article ignores the changing of $R_i$. Then, the various faults located on the unit $i$ of the transformer winding can be simplified to the changing of the parameter $Pa_i = [L_i M_i C_{gi} C_{si}]$ in the equivalent circuit.
network’s ability to detect minor latent faults. Train a deep diagnostic network can effectively improve the accuracy of FR dataset containing complex fault characteristics can be used to group the data with the same label (i.e., the same fault location), due to the randomness of the type and severities of the fault occurring at the same location. Therefore, the fault severity is set to vary randomly from 10% to 50%. The formulas to monitor and locate minor latent faults, the fault severity due to the ever, to prove the powerful ability of the proposed method about feature extraction. Second, it is generally assumed that when the feature extraction is considered, which means ±5% random errors are superimposed on the component parameters of the normal ladder network to obtain the corresponding FR data.

B. MULTI-WINDOW FEATURE EXTRACTION UNDER LOGARITHMIC CONSTRAINTS

In order to fully extract the fault characteristics of the FR data, this article introduces a method for calculating the FR characteristic index: multi-window feature calculation. This method is different from most of the SI methods which directly calculated the entire FR curve into a single characteristic index. This method maps the FR data into a new feature sequence rather than an index. In this method, a window with a specific length is defined. It scans in steps from the beginning of the frequency band to the end. Each time the window moves, the feature value of the data framed in the window. To scan the entire frequency band, the feature value of the data framed in the window is calculated. The calculation formula for the length of the obtained feature sequence is, as shown in equation (4).

\[ N = \text{floor} \left( \frac{(x_n - x_0) - W_w}{s} \right) + 1 \]  

where \( x_0 \) denotes the \( x \) coordinate of the start point and \( x_n \) denotes the \( x \) coordinate of the end point. \( s \) and \( W_w \) are positive integer multiples of the FR sweep frequency step \( \Delta f \). \( s \) denotes the step length of the window. And \( W_w \) denotes the width of the window. To scan the entire frequency band, \( W_w \geq 2s \). \text{floor} represents the lower bound function and returns the largest previous integer.

\( x_k \) denotes the \( x \) coordinate of the symmetry of the \( k \)-th \( (k = 1, 2, \ldots, N) \) window; \( x_{k, \text{left}} \) denotes the \( x \) coordinate of the left border of the \( k \)-th window; while \( x_{k, \text{right}} \) denotes the \( x \) coordinate of the right border of it. Each window contains \( n_w \) points. The calculation of the four variables mentioned are as shown in equations (5)-(8).

\[ x_k = x_0 + \frac{W_w}{2} + (k - 1) \cdot s \]  

\[ x_{k, \text{left}} = x_k - \frac{W_w}{2} \]  

\[ x_{k, \text{right}} = x_k + \frac{W_w}{2} \]  

\[ n_w = \frac{W_w}{\Delta f} + 1 \]  

where \( \Delta f \) is the step length of the sweep frequency.

The multi-window feature calculation method can be used to piecewise calculate different SI of the FR data proposed.
in different researches, such as Correlation Coefficient (CC), Root Mean Square Error (RMSE), Mean of Absolute Percentage Error (MAPE), etc. In this way, different feature sequences can be obtained and served as pre-processing data of the visualization process. In addition, to weaken the dominant influence of extreme values on images and further highlight the characteristics of two-dimensional images obtained through later visualization, and to give full play to the recognition ability of CNN, this article proposes to limit the excessive data difference by introducing logarithmic constraints to the multi-window feature calculation before the feature sequences are put into the visualization process. Taken the MAPE calculation for example, when it has the logarithmic constraints, it can be calculated as equation (9).

\[
\text{MAPE}^* = \log \left( \frac{\frac{1}{n_w} \sum_{i=1}^{n_w} |y_{m,k}^e - y_{m,k}^p|}{n_w} \right) \times 100
\]  

where MAPE* denotes the MAPE that introduces the logarithmic constraints; \(y_{m,k}^e\) denotes the \(y\) coordinate of the \(l\)-th point of the standard FR fingerprint in the \(k\)-th window; and \(y_{m,k}^p\) denotes the \(y\) coordinate of the \(l\)-th point of the of the FR curve which is being calculated.

### C. TWO-DIMENSIONAL HILBERT ID

By using the data visualization methods, the FR feature sequences of the transformer winding obtained from multi-window feature extraction can be effectively input into the deep learning model for training, to finally obtain the deep vision fault positioning model. However, because different visualization methods have different forms and capabilities of feature enhancement, which will directly affect the final training and testing effect of the deep vision fault positioning model, a suitable visualization method for further process is of great importance in this FRA diagnostic scenario.

The Hilbert curve is a special curve that can fill the entire space continuously. This type of curve exists in Euclidean space whose size is greater than 1. And the order of the Hilbert curve can be adjusted to make it traverse every specified points in space [41]. Fig. 4 shows examples of Hilbert curves of different orders (\(n\) denoted the order of the curve). So far, Hilbert curves have been successfully used in the field of data visualization processing [42]. Introducing the Hilbert curve into the diagnostic scenario, the transformer FR feature sequences obtained by the multi-window feature extraction method described in II-B can be arranged in order of the extending direction of the Hilbert curve. And the numerical values of the feature sequences can be represented by different color. Then the FR feature sequences will be finally converted into images, effectively realizing the highlighting of the graphic feature of the sequences. This process can be named as Hilbert visualization, the images obtained from which containing the FR information can be named as Two-dimensional Hilbert ID. And it is abbreviated as Hilbert ID hereinafter. The diagram of the Hilbert visualization process is shown as Fig. 5. To avoid remaining useless features, the red track of Hilbert curve is eliminated in the Hilbert ID. To completely display the feature sequence on the Hilbert ID, the number of points of the selected Hilbert curve must be much more than or equal to that of the feature sequence. The rest points of the selected Hilbert curve left over for the reason will be used to draw the feature sequence again or set to zero, as long as the operation on the entire feature sequence dataset is the same. In addition, due to the infinite extension characteristic of the Hilbert curve, introducing it into the field of fault diagnosis can facilitate the continuous expansion of the feature sequence types. As the number of feature sequence types increasing, the length of the required Hilbert curve will also increase, which means we need a higher-order Hilbert curve for visualization.
Due to the infinite extension characteristic of the Hilbert curve, the deep vision fault positioning method based on this visualization method will have a high scalability, fully exerting the learning ability of the deep diagnostic method and increasing the visual perception and judgment.

At last, the Hilbert ID (MAPE) dataset (the MAPE in parentheses indicates the selected feature index) would be input into a pre-trained CNN network for transfer learning. And the final FRA deep vision fault positioning framework can be obtained. Due to the limitation of article length, since the complex principle of DTL is not the innovation point of this article, and it is difficult to be described briefly, this article will not introduce it in detail. Only the specific implementation of the DTL process will be described later.
III. METHOD IMPLEMENTATION

A. DATASET CONSTRUCTION

When the transformer winding is deformed, short-circuited, or the pitch is changed, the circuit parameters will change accordingly. This kind of change can be detected by FRA. This paper selects the sweep frequency range from 1kHz to 1 MHz.

Obtain the FR dataset of different fault locations containing complex fault characteristics according to the method described in II-A. The transformer winding is divided into 7 parts from top to bottom as shown in Fig. 3. Then the dataset label is $j = \{0, 1, 2, 3, 4, 5, 6, 7\}$ in which ‘0’ represents the normal state. 80% of the obtained dataset is used for transfer learning and 20% is used for testing. The specific fault locations, types of the dataset and their corresponding quantities are shown as Table 1.

According to the multi-window feature calculation method described in II-B, the feature index MAPE is selected to perform the calculation on the original FR dataset, and then a 6-order Hilbert curve is used to visualize the FR dataset into the Hilbert ID (MAPE) dataset.

B. TRANSFER LEARNING AND POSITIONING RESULTS OF THE DEEP NETWORK

At present, there are dozens of CNN image recognition models pre-trained for large image databases. To select a suitable CNN pre-trained network with better performance, this section explores the transfer learning effects of various CNNs in the FRA fault diagnostic scenario of this paper. Firstly, replace the last layer of the network: for most networks, the last layer is a fully connected layer. Replace it with a new fully connected layer so that the number of outputs is equal to the number of diagnostic labels. While for certain networks (such as SqueezeNet), the last learnable layer is a $1 \times 1$ convolutional layer. In this case, replace the old convolutional layer with a new one, whose number of filters is equal to the number of diagnostic labels. Then freeze the parameters of the first 10 layers of the pre-trained network. And input the Hilbert ID (MAPE) dataset into different pre-trained CNN networks for transfer learning. The common hyperparameters settings of different networks are shown in Table 2. The final positioning effects of the CNN diagnostic models are shown in Table 3. As can be seen from Table 3, ResNet50, NASNet, ResNet18 are the top 3 CNN of good diagnostic performance. Fully considering the network size and the number of parameters, this paper selects NASNet as the pre-trained network for further study in the FRA diagnostic scenario.

IV. COMPARISON OF DIFFERENT METHODS

In order to fully explain the advantages of the proposed deep vision fault positioning method which introduces the Hilbert ID and multi-window feature extraction, the positioning effect of the proposed method is first compared with that of certain intelligent diagnostic methods; then compared with that of multiple data processing methods.

A. COMPARISON OF POSITIONING EFFECTS OF DIFFERENT INTELLIGENT DIAGNOSTIC METHODS

In order to fully represent the advantages of the proposed deep vision fault diagnostic method for transformer winding fault
positioning, this article also uses SVM, K-nearest Neighbor (KNN), Decision Tree (DT) to compare their positioning effects. These different intelligent methods are trained and tested by the feature sequence dataset obtained by the multi-window feature calculation of MAPE. Table 4 lists the important parameters and the final accuracy of these three models. Comparing with the NASNet diagnostic result in Table 3, it could be seen that the proposed positioning method based on Hilbert ID (MAPE) has the relatively better effect, with an accuracy of 93.75%, and is 6.51% higher than other intelligent methods (SVM, KNN, DT) on average. Fig. 6 shows the confusion matrix of the final diagnostic results of each method. Comparing Fig. 6 (d) with Fig. 6 (a), (b), and (c), it can be clearly noticed that the proposed method is more accurate. Even if the positioning is wrong, its wrong judgement is still near the right one.

B. COMPARISON OF POSITIONING EFFECTS BASED ON DIFFERENT DATA PROCESSING METHODS

In order to further validate the effectiveness of the proposed Hilbert ID, this paper also compares the Hilbert visualization process with other visualization methods, such as the direct drawing method (the feature sequence is directly drawn into a curve, and the obtained image is named as Curve-ID), histogram method (the feature sequence is drawn into a histogram, and the obtained image is named as Bar-ID). Based on the same kind of intelligent model (NASNet), we test different kinds of visualization process. The final positioning accuracy is shown in the line I of the Table 5. It can be seen from the table that the positioning effect of the Hilbert ID is better than that of the other two visualization methods in the line I, with an average of 5.73% higher.

Similarly, to further illustrate the importance of the logarithmic constraints during the feature extraction process, the FR dataset conducted feature extraction without logarithmic constraints is drawn to Hilbert ID for diagnosis. And its positioning accuracy 73.44% is shown in line III in Table 5. Comparing it with the result in the line I, we can see that the lack of multi-window feature extraction reduces the accuracy by 21.66%. And it can also be clearly seen that the graphical feature of Hilbert ID without logarithmic constraints is visually hard to recognized.

Therefore, both the multi-window feature extraction and the Hilbert visualization method have significant effects on the final positioning results. Compared with the other four kinds of data processing methods Curve ID, Bar ID, Hilbert ID (without feature extraction), Hilbert ID (without logarithmic constraints) in Table 5, the positioning accuracy of the proposed Hilbert ID(MAPE) is 10.35% higher on average.
V. FURTHER IMPROVEMENT

After proving the advantages of the proposed deep vision fault positioning method which introduces Hilbert ID and multi-window feature extraction, in order to find a way to further improve the accuracy of the proposed method, we first use the multi-window feature calculation method to extract different feature sequences based on different SI, and compares their diagnostic results. Then with the infinite extension characteristic of the Hilbert curve, we fuse two feature sequences into one image to increase its number of features and ultimately improve the positioning effect.

Because the multi-window feature calculation method can be used to obtain different feature sequences based on different SI as described in II-B. This paper also compares the deep vision fault positioning results of the Hilbert ID datasets obtained based on different kinds of feature sequences to further explore the positioning methods.

Firstly, respectively obtain the feature sequence datasets by multi-window feature extraction based on different SI such as CC, RMSE, and MAPE. Then implement Hilbert visualization on these three feature sequence datasets to obtain three Hilbert ID datasets: Hilbert ID (CC), Hilbert ID (RMSE), Hilbert ID (MAPE). Finally, input the three datasets respectively into NASNet with hyperparameter settings as described in Table 2 for transfer learning. The final positioning accuracies based on different SI are also compared. And Hilbert ID (CC) performs relatively better than the other two, with an average increase of 1.68%.

The one-dimensional frequency response data is converted into a two-dimensional Hilbert ID, which provides a new expression for FR fingerprint. The Hilbert ID not only highlights the graphic feature of the feature sequences, but also facilitates the visualization of multi-dimensional feature sequences because of the infinite extension characteristic of the Hilbert curve. When considering not only MAPE but also CC of the FRA data during the Hilbert visualization, the accuracy of the proposed method is 10.35% higher on average. The final positioning results of Hilbert ID datasets of different SI (MAPE, CC, RMSE) are also compared. And Hilbert ID (CC) performs better than the other two, with an average increase of 1.68%.

VI. CONCLUSION

This paper proposes a Hilbert ID considering multi-window feature extraction and combines it with a deep transfer learning model for transformer FRA fault positioning. And the final positioning accuracy of Hilbert ID (MAPE)-based NASNet can reach 93.75%. Compared with other intelligent positioning methods, the accuracy of the proposed method is 6.51% higher on average. In addition, compared with the other four data processing methods, the accuracy of the proposed method is 10.35% higher on average. The final positioning results of Hilbert ID datasets of different SI (MAPE, CC, RMSE) are also compared. And Hilbert ID (CC) performs relatively better than the other two, with an average increase of 1.68%.

The one-dimensional frequency response data is converted into a two-dimensional Hilbert ID, which provides a new expression for FR fingerprint. The Hilbert ID not only highlights the graphic feature of the feature sequences, but also facilitates the visualization of multi-dimensional feature sequences because of the infinite extension characteristic of the Hilbert curve. When considering not only MAPE but also CC of the FRA data during the Hilbert visualization, the accuracy of the proposed deep vision fault positioning method can reach 96.09%.

REFERENCES

[1] M. D. Judd, L. Yang, and I. B. B. Hunter, “Partial discharge monitoring of power transformers using UHF sensors. Part I: Sensors and signal interpretation,” IEEE Elect. Insul. Mag., vol. 21, no. 2, pp. 5–14, Mar. 2005.
[2] M. Khanali, A. Hayati-Soloot, H. K. Hoidalen, and S. Jayaram, “Study on locating transformer internal faults using sweep frequency response analysis,” Electr. Power Syst. Res., vol. 145, pp. 55–62, Apr. 2017, doi: 10.1016/j.epsr.2016.11.016.
[3] N. Hashemnia, A. Abu-Siada, and S. Islam, “Improved power transformer winding fault detection using FRA diagnostics—Part 2: Radial deformation simulation,” IEEE Trans. Dielectr. Electr. Insul., vol. 22, no. 1, pp. 564–570, Feb. 2015, doi: 10.1109/TDEI.2014.004592.
[4] V. Nurmanova, M. Bagheri, A. Zollanvari, K. Alikahmet, Y. Akhmetov, and G. B. Gharehpetian, “A new transformer FRA measurement technique to reach smart interpretation for inter-disk faults,” IEEE Trans. Power Del., vol. 34, no. 4, pp. 1508–1519, Aug. 2019, doi: 10.1109/TPWRD.2019.2909144.
G. Qian, Z. Zhao, D. Zou, H. Liu, C. Li, and C. Yao, “Detection of power transformer winding deformation based on continuous wavelet transform,” Gaodianya Jishu/High Voltage Eng., vol. 43, pp. 2016–2023, Jun. 2017, doi: 10.11336/j.1003-6520.hve.201705270537.

A. J. Ghanizadeh and G. B. Gharehpetian, “ANN and cross-correlation based features for discrimination between electrical and mechanical defects and their localization in transformer winding,” IEEE Trans. Dielectr. Electr. Insul., vol. 21, no. 5, pp. 2374–2382, Oct. 2014, doi: 10.1109/TDEI.2014.004364.

K. C. Karimirad and G. B. Gharehpetian, “A new algorithm for localization of radial deformation and determination of deformation extent in transformer windings,” Electric Power Syst. Res., vol. 78, no. 10, pp. 1701–1711, Oct. 2008, doi: 10.1016/j.epsr.2008.02.017.

F. Ren, S. Ji, Y. Liu, Y. Shi, and L. Zhu, “Application of Gauss–Newton iteration algorithm on winding radial deformation diagnosis,” IEEE Trans. Power Del., vol. 34, no. 4, pp. 1736–1746, Aug. 2019, doi: 10.1109/TPWRED.2019.2919927.

A. Khormali and M. Khalooei, Intelligent Bearing Fault Diagnosis with Convolutional Long-Short-Term-Memory Recurrent Neural Network. 2019.

T. Zhang, G. Su, C. Qing, X. Xu, B. Cai, and X. Xing, “Hierarchical lifelong learning by sharing representations and integrating hypothesis,” IEEE Trans. Syst., Man, Cybern. Syst., to be published, doi:10.1109/TSMC.2018.2884996.

J. Wang, S. Li, Z. An, X. Jiang, W. Qian, and S. Ji, “Batch-normalized deep neural networks for achieving fast intelligent fault diagnosis of machines,” Neurocomputing, vol. 329, pp. 53–65, Feb. 2019.

T. Zhang, X. Wang, X. Xu, and C. L. P. Chen, “GCB-net: Graph convolutional broad network and its application in emotion recognition,” IEEE Access, vol. 8, pp. 154–255, 2014, doi:10.1109/ACCESS.2014.2325029.

D. Yey, B. Chatterjee, S. Dalmi, S. Munshi, and S. Chakravorti, “A deep learning framework using convolution neural network for classification of impulse fault patterns in transformers with increased accuracy,” IEEE Trans. Dielectr. Electr. Insul., vol. 24, no. 6, pp. 3894–3897, Dec. 2017, doi:10.1109/TDEI.2017.007693.

X. Yue and P. Beausery, “Transfer learning for one class SVM adaptation to limited data distribution change,” Pattern Recognit. Lett., vol. 100, pp. 117–123, Dec. 2017, doi: 10.1016/j.patrec.2017.10.030.

J. Wang, H. Zheng, Y. Huang, and X. Ding, “Vehicle type recognition in surveillance images from labeled Web-nature data using deep transfer learning,” IEEE Trans. Intell. Transp. Syst., vol. 19, no. 9, pp. 2913–2922, Sep. 2018, doi: 10.1109/TITS.2017.2756767.

I. A. S. Jayasinghe, Z. D. Wang, P. N. Jarman, and A. D. Darwin, “Winding movement in power transformers: A comparison of FRA measurement connection methods,” IEEE Trans. Dielectr. Electr. Insul., vol. 13, no. 6, pp. 1342–1349, Dec. 2006, doi:10.1109/TDEI.2006.258206.

H. Lin, Z. Zhang, W. Tang, Q. Wu, and J. Yan, “Equivalent gradient area based fault interpretation for transformer winding using binary morphology,” IEEE Trans. Dielectr. Electr. Insul., vol. 24, no. 3, pp. 1947–1956, Jun. 2017, doi: 10.1109/TDEI.2017.006237.

T. Zhang, C. L. P. Chen, L. Chen, X. Xu, and B. Hu, “Design of highly nonlinear substitution boxes based on I-ching operators,” IEEE Trans. Cybern., vol. 48, no. 12, pp. 3349–3358, Dec. 2018, doi: 10.1109/TCYB.2018.2846186.

S. Maulik and L. Satish, “Localization and estimation of severity of a discrete and localized mechanical damage in transformer windings: Analytical approach,” IEEE Trans. Dielectr. Electr. Insul., vol. 23, no. 3, pp. 1266–1274, Jun. 2016, doi: 10.1109/TDEI.2015.005408.

M. S. Jahan, R. Keypour, H. R. Izadfar, and M. T. Keshavarzi, “Detecting the location and severity of transformer winding deformation by a novel adaptive particle swarm optimization algorithm,” Int. Trans. Electr. Energy Syst., Jan. 29, no. 1, pp. 1–21, Jan. 2019, doi:10.1002/etep.2666.

T. He, J. Tai, Y. Shan, X. Wang, and X. Liu, “A fast acoustic emission beamforming localization method based on Hilbert curve,” Mech. Syst. Signal Process., vol. 133, Nov. 2019, Art. no. 106291.

E. Estevez-Rams, C. Perez-Demymdenko, B. Aragon Fernandez, and R. Lora-Serrano, “Visualizing long vectors of measurements by use of the Hilbert curve,” Comput. Phys. Commun., vol. 197, pp. 118–127, Dec. 2015, doi:10.1016/j.cpc.2015.08.019.
XIAOXIN WU received the B.Sc. degree from Wuhan University, China, in 2019, where she is currently pursuing the M.Sc. degree in electrical engineering. Her research interests include deep learning and transformer fault diagnosis.

YIGANG HE (Member, IEEE) received the M.Sc. degree in electrical engineering from Hunan University, Changsha, China, in 1992, and the Ph.D. degree in electrical engineering from Xi’an Jiaotong University, Xi’an, China, in 1996.

In 1990, he joined the College of Electrical and Information Engineering, Hunan University, where he was promoted to an Associate Professor and a Professor, in 1996 and 1999, respectively. From 2006 to 2011, he was the Director of the Institute of Testing Technology for Circuits and Systems, Hunan University. He was a Senior Visiting Scholar with the University of Hertfordshire, Hatfield, U.K., in 2002. In 2011, he joined the Hefei University of Technology, Hefei, China. In 2017, he joined Wuhan University, Wuhan, China, where he is currently the Vice Dean of the School of Electrical Engineering and Automation. He has authored over 200 journals and conference papers in his research areas and several chapters in edited books. His current research interests include circuit theory and its applications, testing and fault diagnosis of analog- and mixed-signal circuits, electrical signal detection, smart grids, radio frequency identification technology, and intelligent signal processing. He has been serving on the technical program committee of a number of international conferences. He was a recipient of a number of national and international awards, prizes, and honors.

CHENYUAN WANG received the B.Sc. degree from Wuhan University, Wuhan, China, in 2018, where she is currently pursuing the master’s degree in electrical engineering. Her research interests include the reliability assessment, condition monitoring, and fault diagnosis of high-power modules.

WENJIE WU received the B.Sc. degree from Wuhan University, Wuhan, China, in 2017, where she is currently pursuing the master’s degree in electrical engineering. Her research interests include the transformer fault diagnosis and prognosis.

CHUANKUN WANG received the B.Sc. degree from the Hefei University of Technology, China, in 2019. He is currently pursuing the M.Sc. degree in electrical engineering with Wuhan University, China. His research interests include IGBT reliability and power electronics reliability.

JIAJUN DUAN received the B.Sc. degree from Chongqing University, Chongqing, China, in 2016. He is currently pursuing the Ph.D. degree in electrical engineering with Wuhan University, Wuhan, China. His current research interests include deep learning and intelligent monitoring systems for transformers.

* * *