Improved Winding Mechanical Fault Type Classification Methods Based on Polar Plots and Multiple Support Vector Machines

JIANGNAN LIU\(^1\), ZHONGYONG ZHAO\(^1\), (Member, IEEE), KAI PANG\(^2\), DONG WANG\(^2\), CHAO TANG\(^1\), (Member, IEEE), AND CHENGUO YAO\(^3\), (Member, IEEE)

\(^1\)College of Engineering and Technology, Southwest University, Chongqing 400716, China
\(^2\)Electric Power Research Institute, State Grid Henan Electric Power Company, Henan 450000, China
\(^3\)State Key Laboratory of Power Transmission Equipment and System Security and New Technology, Chongqing University, Chongqing 400044, China

Corresponding author: Zhongyong Zhao (zhaozy1988@swu.edu.cn)

This work was supported in part by the Venture and Innovation Support Program for Chongqing Overseas Returnees under Grant cx2019123, and in part by the Natural Science Foundation of Chongqing under Grant cstc2019jcyj-msxmX0236.

ABSTRACT

The accurate and fast diagnosis of transformer winding deformation faults is of significance to power suppliers and utilities. An improved winding mechanical deformation fault classification method is proposed. In this study, the transformer frequency response data is used to draw polar plots, and then its texture features are extracted for fault classification. The classification model constructed by multiple support vector machines is successfully obtained and shows good classification effect. Besides, this article uses an improved genetic algorithm based on the Emperor-Selective mating scheme and catastrophic operation, to optimize the parameters of support vector machine. The feasibility and accuracy of the proposed method are verified with experimental data obtained from a model transformer, and the proposed method is demonstrated to exhibit better performance compared with the traditional method.

INDEX TERMS

Transformers, fault diagnosis, fault detection, frequency response, support vector machines, genetic algorithms, classification algorithms, digital images, image processing, image texture.

I. INTRODUCTION

Power transformers are among the most important and expensive elements of transmission and distribution networks, which are directly related to the performance and reliability of the network [1]. Winding deformation is one of the main causes of transformer accidents [2]. Generally, mechanical faults can occur on transformer windings, which can subsequently lead to transformer failures [3]. Significant failures can result in long-term interruptions of transformer, and the repairs are costly and time-consuming [4]. The early detection of winding mechanical faults can prevent the transformer from consecutive and catastrophic failure [5]. Thus, it is necessary to diagnose transformer winding mechanical fault when the fault is at an early stage, to provide guidance on transformer maintenance [6].

At present, many transformer fault diagnosis methods have been proposed and practically implemented [7]. For instance, the temperature measurement [8], dissolved gas-in-oil analysis (DGA) [9], dielectric response analysis [10], vibration analysis [11], partial discharge analysis (electrical and acoustic) [12], [13], transfer function measurement [14], ultrasonic method and online short circuit impedance (SCI) [15]. All of these methods are significant. Some of them are already commercialized; however, some other power frequency techniques such as online SCI methods are still in the development stage due to the lack of sufficient accuracy and robustness.

Among the commercialized techniques, frequency response analysis (FRA) is a simple and popular technique [16]. However, despite the wide application of FRA in the power industry, the diagnostic accuracy of FRA can still be improved, according to the feedback of numerous field tests. In addition, although the online FRA has been studied, it still remains some challenges for its actual application, for example, the difficulties in online measurement, impact of electromagnetic interference on measurement, and lack of standard interpretation should be considered.
Thus, many improved versions of the FRA method have been widely studied [17]–[19]. Nevertheless, most of the existing literature only considered the amplitude information of frequency response, while the uses of phase-frequency characteristics were often neglected, except for a few studies [20]–[22]. In [22], transformer polar plots were drawn by using both the amplitude and phase information of the frequency response, which was used as a new signature for fault diagnosis. In [23], polar plot and digital image processing (DIP) techniques were used to detect the radial deformation (RD) fault of transformer. There is no doubt that the new polar plot is superior to the traditional FRA fingerprints that only provide amplitude information in fault diagnosis. Thus, the polar plot method is also adopted and studied in this contribution.

By using the DIP technique, the FRA polar plot image features can be extracted. There have been efforts to apply DIP to transformer fault diagnosis [24]–[26]. Ref. [27] studied the polar plot texture features of several common faults, which is of significance in fault diagnosis.

Another problem is the diagnosis of winding faults. Despite the diagnosis of winding mechanical deformations can be realized by the traditional statistical indicator method in both the FRA single signature and polar plot signature. However, up to now, the accurate diagnosis and recognition of winding deformations from FRA traces still remains some challenges. Some advanced artificial intelligence methods have been used and studied. Support Vector Machine (SVM) is one of these intelligent methods, which has proven to be suitable for processing the FRA amplitude-frequency signatures with highly dimensional and small samples [28], [29]. In addition, the parameters of SVM model are significant for the performance of classifiers. As one of the most widely accepted intelligent algorithms, the genetic algorithm (GA) has often been used in SVM parameter optimization [30], [31].

However, in the above studies, the method in [22], [23] was not verified by relevant physical experiments. In addition, the feature extraction of signatures can be improved. In [27], only the gray-level co-occurrence matrix (GLCM) features were extracted and focused, which has limitations to some extent. However, most of the existing reported literature only used a single SVM classifier for transformer fault classification [29], [32], which can be modified by a combination strategy of multiple SVMs, to improve the accuracy, calculation speed and robustness of fault classification. In addition, there remains much room for improvement in the standard genetic algorithm (SGA).

To improve the performance of winding deformation fault diagnosis, this study proposes an improved fault classification method based on FRA polar plot images and a combination of multiple SVM classifiers. The main steps of the method are summarized as follows.

- First, the polar plot with both amplitude and phase information of FRA is used to replace the traditional FRA signature as the foundation of fault diagnosis.

- Second, three texture features of the polar plot image are extracted by using DIP technology, including histogram of oriented gradient (HOG) features, local binary pattern (LBP) features, and GLCM features, which are extracted from the perspective of global description, local description and mathematical statistics, respectively.

- Third, three SVM models are trained using different texture features as input. An improved genetic algorithm (IGA) that integrates several recognized methods is used to optimize the SVM parameters. Three SVM models with independent decision-making ability are obtained.

- Finally, three independent SVM models are combined into a strong classifier for fault diagnosis.

With the above background, this study provides some new ideas for winding fault type classification considering the state-of-the-art method. The four main contributions of this study are as follows: a) polar plot, b) texture features, c) multiple SVMs classifier, d) IGA. In addition, the experimental data obtained from a model transformer is adopted to verify the feasibility and accuracy of the proposed method.

The rest of this article is organized as follows. In Section II, the FRA polar plot is introduced. Section III and IV present explanations for the feature extraction, IGA-SVM and classification strategy, respectively. In Section V, the experimental test setup is introduced and the classification results are discussed. Finally, Section VI concludes this article.

### II. FREQUENCY RESPONSE ANALYSIS POLAR PLOT

FRA, firstly introduced by Dick and Erven [33], adopts a sweep frequency sinusoidal signal to excite transformer windings and measure the response signal in the frequency domain to construct a frequency response signature [34]. The frequency response of transformer winding is expressed as:

\[
H(f) = 20 \log_{10} \left( \frac{U_{out}(f)}{U_{in}(f)} \right)
\]

where \( U_{in} \) is the excitation voltage, \( U_{out} \) is the response voltage, \( f \) is the frequency of the applied sinusoidal excitation, and \( H \) is the frequency response. Both \( U \) and \( H \) are phasors with amplitude and phase.

At present, only the amplitude information of the FRA signature is analyzed in the most cases of FRA application, and the phase information is often neglected. However, Ref. [35] noted that using a combination of amplitude and phase information is more sensitive to the transformer winding faults than using only amplitude information. The FRA polar plot is an image that combines the amplitude and phase information of the frequency response. The polar plot is plotted as follows: when a sweep sinusoidal signal is exerted to the terminal of the transformer winding, the corresponding amplitude and phase of FRA signature are recorded at each frequency sampling point. In the polar coordinates, the specific point, which corresponds to the frequency response at the exact frequency, is plotted by taking the measured phase as the polar angle and the measured amplitude as the polar diameter. By this way, the FRA polar plot can be drawn in polar coordinates.
III. FEATURE EXTRACTION

As mentioned before, this study takes the polar plot as a fingerprint of a transformer. In the field of biometric technology, many fingerprint and signature recognition are depending on texture feature comparison [36]. The biometric technology can be migrated and applied to the transformer fault diagnosis. Thus, as the signature of transformers, it is also feasible to realize the fault diagnosis by identifying the texture features of polar plot.

In the image processing, the texture can be defined as a function of spatial variation of the brightness intensity of the pixels [36]. This study selects three widely used texture features: HOG features, LBP features and GLCM features as the basis of fault recognition. These features are described in detail below.

A. HOG FEATURES

The HOG feature is a feature descriptor that is used to detect objects in computer vision and image processing. The feature is constructed by calculating the gradient direction histogram of the local region of an image. In an image, the appearance and shape of the local target can be described by the directional density distribution of the gradient or edge. HOG features are essentially statistical information of the gradients.

The gradients mainly exist at the edge, thus, the HOG features can be used to describe the texture features. Here is the procedure: First, the image is divided into small- connected areas, which are known as cell units. Then, the direction histogram of the gradient or edge of each pixel in the cell unit is collected. Finally, these histograms can be combined to form feature descriptors. Fig. 2 shows a visualization of the HOG feature.

B. LBP FEATURES

The LBP feature, firstly proposed by Ojala et al. [37], is an operator that is used to describe the local texture feature of an image. The core concept of LBP is to treat the gray value of the central pixel as the threshold value, and obtain the corresponding binary code to represent the local texture features compared with its neighborhood. From the perspective of texture analysis, the texture features of a pixel in an image refer to the relationship between the certain point and the neighboring pixel points. Thus, the LBP feature is a descriptor that measures the relationship between a pixel and its surrounding pixels. Fig. 3 shows a visualization of the LBP feature.

In this study, the LBP value is calculated following the definition that an ordered set of binary comparisons of pixel intensities between the center pixel and its eight surrounding pixels [38]. Therefore, the radius of the LBP operator is 1, and the LBP value of the center pixel is calculated by the binary coding of 8 neighbors.
TABLE 1. GLCM eigenvalues of typical samples in several winding conditions.

| Index        | Sample  | Normal | DSV fault | SC fault | AD fault |
|--------------|---------|--------|-----------|----------|----------|
| ASM of 0° GLCM | 0.95353 | 0.948762 | 0.951456 | 0.950745 |
| ASM of 45° GLCM | 0.950146 | 0.945135 | 0.94689 | 0.946356 |
| ASM of 90° GLCM | 0.955881 | 0.951579 | 0.954128 | 0.953832 |
| ASM of 135° GLCM | 0.950027 | 0.944458 | 0.949192 | 0.948155 |
| ENT of 0° GLCM | 0.138135 | 0.149454 | 0.142829 | 0.144577 |
| ENT of 45° GLCM | 0.145632 | 0.157747 | 0.153049 | 0.154374 |
| ENT of 90° GLCM | 0.131567 | 0.142259 | 0.135994 | 0.136677 |
| ENT of 135° GLCM | 0.14589 | 0.159169 | 0.148138 | 0.150573 |
| IDM of 0° GLCM | 0.99205 | 0.991238 | 0.991621 | 0.991452 |
| IDM of 45° GLCM | 0.99045 | 0.989419 | 0.989312 | 0.989235 |
| IDM of 90° GLCM | 0.993352 | 0.992693 | 0.992999 | 0.993045 |
| IDM of 135° GLCM | 0.990388 | 0.989066 | 0.990511 | 0.990173 |
| COR of 0° GLCM | 31.42402 | 28.50541 | 29.77443 | 29.1633 |
| COR of 45° GLCM | 24.82175 | 22.27595 | 20.87198 | 20.81833 |
| COR of 90° GLCM | 36.84403 | 33.53478 | 35.13749 | 35.21437 |
| COR of 135° GLCM | 24.56766 | 21.0629 | 25.50242 | 24.35411 |

C. GLCM FEATURES

The GLCM statistical method, proposed by Haralick et al. [39], is a widely used texture analysis method. It is based on the assumption that the spatial distribution relationship among the pixels contains the texture information of the image. The image texture is formed by the repetitive appearance of pixel gray in a spatial position. GLCM is a joint distribution of two pixel gray levels with a certain spatial position relationship. Briefly, GLCM is the joint histogram of two pixel gray levels, which is a second-order statistic.

However, the direct input of classifiers is not the GLCM but the mathematical features extracted from the GLCM. In this study, the following widely used mathematical features are adopted [40] (in the following equations, \( p(i,j) \) is the normalized GLCM):

1) **ANGULAR SECOND MOMENT (ASM)**

\[
ASM = \sum_{i=1}^{I} \sum_{j=1}^{J} p(i,j)^2
\]

The ASM is the sum of squares of the elements of the GLCM. The ASM reflects the uniformity of the gray distribution and texture thickness.

2) **INVERSE DIFFERENCE MOMENT (IDM)**

\[
IDM = \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{p(i,j)}{1 + (i-j)^2}
\]

The IDM is a measure of the local variation of the image texture that reflects the degree of the regularity of the texture.

3) **ENTROPY (ENT)**

\[
ENT = - \sum_{i=1}^{I} \sum_{j=1}^{J} p(i,j) \times \log_{10} p(i,j)
\]

The ENT represents the amount of information in an image, which is a measure of the randomness of image content, and represents the complexity of texture.

4) **CORRELATION (COR)**

The COR is a measure of the similarities of GLCM elements in row or column directions. This parameter can be used to determine the main direction of the texture.

\[
COR = \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{i \times j \times p(i,j) - \mu_1 \times \mu_2}{\delta_1 \times \delta_2}
\]

\[
\mu_1 = \sum_{i=1}^{I} \sum_{j=1}^{J} p(i,j)
\]

\[
\mu_2 = \sum_{i=1}^{I} \sum_{j=1}^{J} p(i,j)
\]

\[
\delta_1^2 = \sum_{i=1}^{I} (i - \mu_1)^2 \sum_{j=1}^{J} p(i,j)
\]

\[
\delta_2^2 = \sum_{j=1}^{J} (j - \mu_1)^2 \sum_{i=1}^{I} p(i,j)
\]

Because GLCM features are statistical features, they are not easily visualized. Therefore, in Table 1, the eigenvalue data of some representative samples are attached.

IV. IGA-SVM AND CLASSIFICATION STRATEGY

It seems unreasonable to use only SVM when other intelligent methods such as LSTM and GRU are in wide use. However, in fact, it is very difficult to obtain much FRA data from the actual transformers. Such a small quantity of data is insufficient for training a neural network with a complex structure. Nevertheless, SVM is suitable for solving the problem of small sample pattern classification. This study proposes a strategy that combines multiple SVMs to obtain a more robust classifier. Another reason for using multiple SVMs lies in the
fact that the dimensions of the texture features differ from each other, if only a single SVM model is used, the simple fusion strategy of features may lead to consequence that high-dimensional features cover up low-dimensional features. This problem can be avoided by using multiple SVMs.

A. IGA-SVM

Fig. 4 shows the flowchart for training a SVM model.

![Flowchart for training a SVM model](image)

1) EMPEROR-SELECTIVE (EMS) MATING SCHEME

The mating scheme plays an important role in the convergence and robustness of GA. The EMS scheme has been demonstrated to be more effective than other schemes. It allows the fittest individual to procreate freely with the rest of the population, resulting in a greater diversity and faster convergence, and ensures that the high-quality gene is preserved during the random process. Fig. 5 shows a schematic of two different mating schemes.

As seen in Fig. 5, the mating strategy of SGA is to treat the individuals in the population equally and mate randomly among them. Such a mating scheme cannot produce good offspring stability. The EMS scheme regards the optimal individual as the emperor of the population, and the emperor has the right to mate with any individual in the population. This mating scheme enables the genes of the optimal individual to be smoothly passed to the next generation, and ensures that the high-quality gene is preserved during the random process. Fig. 5 shows a schematic of two different mating schemes.

![Comparison of SGA and EMS-GA mating schemes](image)

2) SELECTION OPERATION

The selection operation of the SGA usually adopts roulette gambling. However, there are still some defects in this selection operation. Individuals with high fitness will occupy the population rapidly in the early iteration stage, and the population will stop evolving in the later iteration stage. Therefore, another selection operation is adopted to ensure the transmission of high-quality genes.

In the proposed selection operation, the population is first sorted according to the fitness degree, and divided into three parts. All superior individuals are copied to replace the inferior individuals, which constructs a new generation of the population, as shown in Fig. 6. This selection operation can protect the diversity of the population and avoid local optimization.

![Diagram of the IGA selection operation](image)

3) CROSSOVER AND MUTATION OPERATION

The crossover and mutation operation of the SGA were completed with fixed crossover rate $P_C$ and mutation rate $P_M$, which may lead to the individual genes with high fitness being destroyed and the convergence speed of the algorithm being reduced. Therefore, in this study, the improved adaptive function is introduced to make the $P_C$ and $P_M$ adjust nonlinearly with the individual fitness to improve the convergence speed and stability of the algorithm. When the individual fitness is superior to the average fitness, the stability of excellent genome can be maintained by reducing the individual crossover rate and mutation rate; when the individual fitness is inferior to the average fitness, the diversity of population gene can be enhanced by increasing the individual crossover rate and mutation rate. The adaptive adjustment function of
A and \( P \) is shown as (6) and (7) [41].

\[
P_C = \begin{cases} 
\frac{P_{C\text{max}} - P_{C\text{min}}}{1 + e^{\frac{2(M_{f'} - f_{avg})}{P_{C\text{max}} - P_{C\text{min}}}}} & (f' < f_{avg}) \\
1 + e^{\frac{2(M_{f'} - f_{avg})}{P_{C\text{max}} - P_{C\text{min}}}} & (f' \geq f_{avg}) 
\end{cases}
\]

\[
P_M = \begin{cases} 
\frac{P_{M\text{max}} - P_{M\text{min}}}{1 + e^{\frac{2(M_f - f_{avg})}{P_{M\text{max}} - P_{M\text{min}}}}} & (f < f_{avg}) \\
1 + e^{\frac{2(M_f - f_{avg})}{P_{M\text{max}} - P_{M\text{min}}}} & (f \geq f_{avg}) 
\end{cases}
\]

where: \( A \) is the transformation constant value equal to 9.903438, \( P_{C\text{max}} \) and \( P_{C\text{min}} \) are the maximum and the minimum of the crossover rate, respectively, \( P_{M\text{max}} \) and \( P_{M\text{min}} \) are the maximum and the minimum of the mutation rate, respectively, \( f_{avg} \) is the average fitness of the population, \( f' \) is the higher fitness value of two mating individuals, and \( f \) is the individual fitness value.

4) CATASTROPHIC OPERATION
The local optimum problem of the SGA can be solved by the laws of natural evolution. According to the laws of natural evolution, if we want to escape the trap of local optimal value, we must destroy all the best individuals at present and provide the possibility for other better solutions.

When the best individual has not evolved in successive generations, or each individual has been too close to each other and the population lacks diversity, a catastrophic operation can be considered and applied. The specific step of catastrophic operation is to randomly select the individuals in the population, and carry out random mutation operations on them, to increase the genetic diversity of the population.

By using the catastrophic operation, the monopoly advantage of the original gene can be broken, the gene diversity for the population can be created, and the algorithm is more prone to converge to the global optimal value.

B. CLASSIFICATION STRATEGY
In order to distinguish the three IGA optimized SVM models obtained by the abovementioned texture features, the SVM models are named SVM_HOG, SVM_LBP and SVM_GLCM. To test the performance of each SVM model, 1,000 random experiments are conducted to test the three SVM models separately. The random tests are carried out according to the following steps. The data set is divided into a test set and a training set. The test set is made up of 10 samples that are randomly picked from the data set, and these samples contain all fault types. The remaining samples in the data set are used as the training set. The SVM model is then trained and obtained with the training set; after that, each SVM model is tested with the test set. The performance of the three SVM models in the random tests is shown in Fig. 7.

After obtaining the three independent SVM models, how to combine them into a strong classifier determines the performance of the final classifier. The most appropriate solution is to assign weights to the outputs of each SVM model and combine them into one strong classifier. The problem is how to assign weights to these SVM models. The first approach is equal weight assignment, namely, the principle of the minority being subordinate to the majority. This weight distribution method is simple and easy to implement, and has strong practical significance. However, from the performance comparison diagrams of the three SVMs in Fig. 7, the classification accuracy of SVM_HOG and SVM_LBP is significantly larger than that of SVM_GLCM. Thus, the strategy of equal weight distribution still has some shortcomings.

This strategy ignores the differences between several SVM models. Moreover, it fails to process the worst-case scenario in which the outputs of three SVM models are inconsistent with each other.

Therefore, in this study, the following strategy is adopted. First, following the principle of the minority being subordinate to the majority in general, the three SVM models are voted to determine the final classification results of the samples. If the outputs of the three models are inconsistent, then the principle of the minority being subordinate to the majority will not be applicable. Because judging the correctness of model is a random process, the Markov chain model is used to decide which SVM model and its output are adopted as the final classification result of the sample.

We conducted numerous random experiments to collect a sufficient number of cases in which the outputs of the three models are inconsistent. The classification accuracy of SVM_HOG and SVM_LBP is much higher than that of SVM_GLCM; therefore, the output of SVM GLCM is directly excluded, and one of the outputs of SVM_HOG and SVM_LBP is selected as the final classification result. When the output of SVM_HOG is the correct judgment, the record state of the Markov chain is 1. When the output of SVM_LBP is the correct judgment, the record state of the Markov chain is 2. The Markov chain model with a chain length of 100 is established and 99 state transitions are recorded in a Markov chain. The state transition probability can be calculated by using the frequency of transitions between states. The initial transition probability is obtained by the ratio of test accuracy of SVM_HOG and SVM_LBP.
Fig. 8 shows the state transition visualization of the Markov chain model. The specific SVM model is determined by the output of the Markov chain model.

V. EXPERIMENTAL RESULT
A. RESULTS AND ANALYSIS
In this study, the experimental transformer FRA data are obtained from the previous work reported in [7]. A special model transformer was manufactured to perform the winding simulated fault experiments. All experiments were carried out using FRA end-to-end open circuit connection, and the amplitude and phase of the frequency response were recorded in experimental data. Fig. 9 shows images of the experimental transformer model. More information about the model transformer and the detailed experimental setup can be found in [42], [43].

Winding RD, inter-disk Short Circuit (SC) and Disk Space Variation (DSV) faults were simulated on the model transformer. In the case of the SC fault, 15 groups of different frequency response data were measured. In the case of the DSV fault, 21 groups of different FRA data were obtained. RD fault experiments were carried out with 17 different conditions. Fig. 11 shows some typical FRA polar plots in different situations under different fault types. To be clear, the proposed method is based on the texture of polar plot to achieve fault classification. The existence of scale will seriously affect the accurate extraction of polar plot texture. Thus, the scales were removed in the polar plots shown in Fig. 11. The scale and data distribution range are found in Fig. 1.

Fig. 10 shows a typical fitness curve when the IGA is used to optimize the SVM parameters. Two obvious phenomena are observed. In the early iteration stage of the algorithm, the average fitness of the population rapidly converges to the optimal fitness, which is owing to the EMS adopted by the algorithm. Each individual of the population mates with the best individual, which results in improved average fitness and rapid convergence. In the later iteration stage, it can be observed that the average fitness of the population changes significantly between several generations, which is caused by the catastrophic operation. The fitness of the population is high, and the catastrophic operation is very likely to cause a decline of the average fitness. Meanwhile, the gene diversity improves, which is helpful for searching the global optimum.

After combining several SVM models into a strong classifier according to the above methods, the performance of the final classifier is also evaluated by the aforementioned random experiments. Fig. 12 shows the confusion matrix of the final classifier. Labels 1, 2 and 3 correspond to the DSV, SC and RD fault, respectively. The labels at the bottom of Fig. 12 are the actual labels of the winding fault, which are the targets of the classifier. The left labels in Fig. 12 are the output labels obtained by the classifier. Taking the output and target of the classifier as the index, the classified number of corresponding samples and their percentage to the total number of samples is also displayed in the box. If the classification result
FIGURE 11. Comparison of FRA polar plots in different situations under three fault types: (a) polar plots in DSV fault condition (b) polar plots in SC fault condition (c) polar plots in RD fault condition.

is correct, the box is displayed in green; if incorrect, the box is displayed in red. The correct classification rate and error classification rate of corresponding classes are calculated in the gray box. The final statistical results of the correct classification rate and the error classification rate of all samples are shown in the blue box.
It can be seen in Fig. 12 that the accuracy of the final classifier reaches 97.8%, and the performance is excellent. It should be noted that the FRA data are not divided into several sub-frequency bands in this contribution; instead, the entire frequency band of 1-1,000 kHz is analyzed. From the results, the proposed method has reached a satisfactory accuracy, thus, it is not necessary to divide the frequency band. In addition, the most likely mistake that the classifier makes is to divide SC fault into DSV fault. Compared with other fault types, one might suspect that the classifier does not fully extract the features of SC fault. In fact, the measured FRA data of the SC fault are the rarest of the three fault types. This phenomenon is closely related to the lack of SC fault measured data. If the proposed method is industrialized in the future, the problem of insufficient winding fault training data will be solved, and the classification accuracy of the classifier will be further improved.

Table 2 shows the comparison of IGA-based and SGA-based parameter optimization algorithm for the SVM model. The prediction accuracy and time consumption of the three SVM models are compared. It can be concluded that the time consumption of IGA is longer than that of SGA, but the accuracy of the IGA optimized model is much larger than that of the SGA optimized model. The searching ability of IGA is stronger, and the quality of the obtained solution is higher.

**TABLE 2. Comparison of IGA-based and SGA-based parameter optimization algorithm for SVM model.**

| Algorithm | Time consumption | Model       | Training accuracy | Testing accuracy |
|-----------|------------------|-------------|-------------------|------------------|
| IGA       | 38.21s           | SVM_HOG     | 100% (43/43)      | 100% (10/10)     |
|           |                  | SVM_LBP     | 100% (43/43)      | 100% (10/10)     |
|           |                  | SVM_GLCM    | 90.70% (39/43)    | 70% (7/10)       |
| SGA       | 35.23s           | SVM_HOG     | 100% (43/43)      | 90% (9/10)       |
|           |                  | SVM_LBP     | 97.67% (42/43)    | 90% (9/10)       |
|           |                  | SVM_GLCM    | 93.02% (40/43)    | 70% (7/10)       |

**B. COMPARISON WITH THE TRADITIONAL METHOD**

To compare the classification result of the proposed method with the traditional method, the classification results of previous contribution [7] using traditional mathematical indexes and single SVM method are also established. Fig. 13 (a) compares the recognition accuracy of the proposed method with that of the traditional method under different fault types. Fig. 13 (b) illustrates the classification accuracy.

![Confusion Matrix](image)

**FIGURE 12.** Confusion matrix of the final classifier. The confusion matrix records the number of samples that the predicted result (output label) of the classifier match or mismatch with the real label (target label), and calculates the corresponding accuracy.

![Comparison of classification accuracy between proposed method and traditional method](image)

**FIGURE 13.** Comparison of classification accuracy between proposed method and traditional method: (a) comparison of the recognition accuracy of the proposed method and the traditional method under different fault types; (b) comparison of comprehensive classification accuracy.

Form Fig. 13, the following conclusions can be drawn. Although the SC fault classification accuracy of the proposed method is slightly lower than that of the traditional method, the RD fault classification accuracy is much higher. In addition, the shaded area surrounded by the red line in Fig. 13 (a) is much larger than that surrounded by the blue line, which shows that the proposed method has a more balanced performance for the classification of various faults. This phenomenon benefits from the combination of multiple SVMs, which once again indicates that the proposed multi-SVM model is more intelligent and powerful than the single SVM model.

In addition, several typical contributions of the existing literature are also compared with the method proposed in this
study. The classification accuracy of the fault types achieved by Ref. [44] is 96.74%, and that of Ref. [45] is 97.2%. Both of them are smaller than the 97.8% of the proposed method. Although the samples for training and testing SVM models are quite different in various literature, the comparison of accuracy can indirectly demonstrate the effectiveness of the proposed method.

VI. CONCLUSION

This article presents an improved method for classifying and recognizing transformer winding mechanical faults. The HOG, LBP and GLCM texture features of the transformer polar plot are extracted and input to the multi-SVM classifiers for fault type identification. The IGA is used to optimize the SVM parameters to achieve a better performance. The following conclusions can be drawn:

1) According to the proposed method, the obtained classifier model is capable of classifying transformer winding deformation fault types;
2) The texture feature extracted from the polar plot can represent the state information of transformer winding;
3) The classification accuracy of the multi-SVMs classifier is higher than that of the traditional single SVM classifier.

However, there are still some limitations to this study. The data set of this study is not large enough to be widely representative. Moreover, this study only discusses the classification of fault types, and does not pay attention to the degree and location of fault.

The experiment results show that this method is a highly accurate fault classification method. It can provide guidance in fault diagnosis work if massive FRA fault data are obtained in the future. In addition, the identification of transformer fault location and degree can be taken into account in the next work. Data mining and neural networks with more complex structures can be used for fault diagnosis when large-scale data of transformer winding deformation faults are available.

REFERENCES

[1] A. R. Abbasi, M. R. Mahmoudi, and Z. Avazzadeh, “Diagnosis and clustering of power transformer winding fault types by cross-correlation and clustering analysis of FRA results,” *IET Gener., Transmiss. Distrib.*, vol. 12, no. 19, pp. 4301–4309, Oct. 2018.

[2] M. Mostafaei and F. Haghjoo, “Flux-based turn-to-turn fault protection for power transformers,” *IET Gener., Transmiss. Distrib.*, vol. 10, no. 5, pp. 1154–1163, Apr. 2016.

[3] A. J. Ghazizadeh 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.

[4] S. Bagheri, Z. Moravej, and G. B. Gharehpetian, “Classification and discrimination among winding mechanical defects, internal and external electrical faults, and inrush current of transformer,” *IEEE Trans. Ind. Informat.*, vol. 14, no. 2, pp. 484–493, Feb. 2018.

[5] M. H. Samimi, S. Tenbohnen, A. A. S. Akmal, and H. Mohseni, “Effect of different connection schemes, terminating resistors and measurement impedances on the sensitivity of the FRA method,” *IEEE Trans. Power Del.*, vol. 32, no. 4, pp. 1713–1720, Aug. 2017.

[6] H. Tarimoradi and G. B. Gharehpetian, “Novel calculation method of indices to improve classification of transformer winding fault type, location, and extent,” *IEEE Trans. Ind. Informat.*, vol. 13, no. 4, pp. 1531–1540, Aug. 2017.

[7] J. Liu, Z. Zhao, C. Tang, C. Yao, C. Li, and S. Islam, “Classifying transformer winding deformation fault types and degrees using FRA based on support vector machine,” *IEEE Access*, vol. 7, pp. 112494–112504, 2019.

[8] Y. P. Liu, J. Y. Yin, X. Z. Fan, and B. W. Wang, “Distributed temperature detection of transformer windings with externally applied distributed optical fiber,” *Appl. Opt.*, vol. 58, no. 29, pp. 7962–7969, Oct. 2019.

[9] D. Kweon and Y. Kim, “Interpretation of turn-to-turn insulation fault by dissolved gas analysis,” *IEEE Trans. Dielectr. Electr. Insul.*, vol. 25, no. 4, pp. 1560–1566, Aug. 2018.

[10] J. Gao, L. Yang, Y. Wang, X. Liu, Y. Lv, and H. Zheng, “Condition diagnosis of transformer oil-paper insulation using dielectric response fingerprint characteristics,” *IEEE Trans. Dielectr. Electr. Insul.*, vol. 23, no. 2, pp. 1207–1218, Apr. 2016.

[11] K. X. Hong and G. X. Lin, “State classification of transformers using nonlinear dynamic analysis and Hidden Markov models,” *Measurement*, vol. 147, p. 10, Dec. 2019.

[12] H. Suryavanshi, J. Velandy, and M. Sakhiviel, “Wavelet power ratio signature spectrum analysis for prediction of winding insulation defects in transformer and shunt reactor,” *IEEE Trans. Dielectr. Electr. Insul.*, vol. 24, no. 4, pp. 2649–2659, Aug. 2017.

[13] H. Tabarsa, M. A. Hejazi, and G. B. Gharehpetian, “Detection of HV winding radial deformation and PD in power transformer using stepped-frequency hyperboloid method,” *IEEE Trans. Instrum. Meas.*, vol. 68, no. 8, pp. 2934–2942, Aug. 2019.

[14] Q. Yang, P. Y. Su, and Y. Chen, “Comparison of impulse wave and sweep frequency response analysis methods for diagnosis of transformer winding faults,” *Energies*, vol. 10, no. 4, p. 16, Apr. 2017.

[15] M. M. Islam, G. Lee, and S. N. Hettipatne, “A review of condition monitoring techniques and diagnostic tests for lifetime estimation of power transformers,” *Electr. Eng.*, vol. 100, no. 2, pp. 581–605, Jun. 2018.

[16] V. Nurmanova, M. Bagheri, A. Zollanvari, K. Aliakhamet, 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.

[17] X. Zhao, C. Yao, A. Abu-Siada, and R. Liao, “High frequency electric circuit modeling for transformer frequency response analysis studies,” *Int. J. Electr. Power Energy Syst.*, vol. 111, pp. 351–368, Oct. 2019.

[18] M. H. Samimi, P. Hillenbrand, S. Tenbohnen, A. A. S. Akmal, H. Mohseni, and J. Faiz, “Investigating the applicability of the finite integration technique for studying the frequency response of the transformer winding,” *Int. J. Electr. Power Energy Syst.*, vol. 110, pp. 411–418, Sep. 2019.

[19] A. A. Devadiga, N. Harid, H. Griffiths, N. Al Sayari, B. Barkat, S. Jayaram, H. Jafar, T. Koshizaki, and Y. Taniguchi, “Winding turn-to-turn short-circuit diagnosis using FRA method: Sensitivity of measurement configuration,” *IET Sci., Meas. Technol.*, vol. 13, no. 1, pp. 17–24, Jan. 2019.

[20] O. Aljohani and A. Abu-Siada, “Application of DIP to detect power transformers axial displacement and disk space variation using FRA polar plot signature,” *IEEE Trans. Ind. Informat.*, vol. 13, no. 4, pp. 1794–1805, Aug. 2017.

[21] O. Aljohani and A. Abu-Siada, “Application of digital image processing to detect transformer bushing faults and oil degradation using FRA polar plot signature,” *IEEE Trans. Dielectr. Electr. Insul.*, vol. 24, no. 1, pp. 428–436, Feb. 2017.

[22] O. Aljohani and A. Abu-Siada, “Application of digital image processing to detect short-circuit turns in power transformers using frequency response analysis,” *IEEE Trans. Ind. Informat.*, vol. 12, no. 6, pp. 2062–2073, Dec. 2016.

[23] A. Abu-Siada and O. Aljohani, “Detecting incipient radial deformations of power transformer windings using polar plot and digital image processing,” *IET Sci., Meas. Technol.*, vol. 12, no. 4, pp. 492–499, Jul. 2018.

[24] S. Mortazavian, M. M. Shabestary, Y. A.-R.-I. Mohamed, and G. B. Gharehpetian, “Experimental studies on monitoring and metering of power transformer windings using polar plot and digital image processing and UWB transceivers,” *IEEE Trans. Ind. Informat.*, vol. 11, no. 6, pp. 1334–1345, Dec. 2015.

[25] A. Abu-Siada, I. Radwan, and A. F. Abdou, “3D approach for fault identification within power transformers using frequency response analysis,” *IET Sci., Meas. Technol.*, vol. 13, no. 6, pp. 903–911, Aug. 2019.
Z. Zhao, C. Yao, C. Tang, C. Li, F. Yan, and S. Islam, “Diagnosis of transformer winding deformation faults based on the analysis of binary image obtained from FRA signature,” IEEE Access, vol. 7, pp. 40463–40474, 2019.

X. Zhao, C. Yao, C. Li, C. Zhang, S. Dong, A. Abu-Siada, and R. Liao, “Experimental evaluation of detecting power transformer internal faults using FRA polar plot and texture analysis,” Int. J. Electr. Power Energy Syst., vol. 108, pp. 1–8, Jun. 2019.

C. Zhang, Y. He, S. Jiang, T. Wang, L. Yuan, and B. Li, “Transformer fault diagnosis method based on self-powered RFID sensor tag, DBN, and MKSVM,” IEEE Sensors J., vol. 19, no. 18, pp. 8202–8214, Sep. 2019.

J. Li, Q. Zhang, K. Wang, J. Wang, T. Zhou, and Y. Zhang, “Optimal dissolved gas ratios selected by genetic algorithm for forecasting dissolved gas content in power transformers,” Energies, vol. 11, no. 9, p. 19, Sep. 2018.

S. Koroglu and A. Demircali, “Diagnosis of power transformer faults based on multi-layer support vector machine hybridized with optimization methods,” Electr. Power Compon. Syst., vol. 44, no. 19, pp. 2172–2184, Nov. 2016.

Y. Bennmahamed, M. Teguar, and A. Boubakeur, “Diagnosis of power transformer oil using PSO-SVM and KNN classifiers,” in Proc. Int. Conf. Electr. Sci. Technol. Maghreb (CISTEM), Oct. 2018, pp. 1–4.

E. P. Dick and C. C. Erven, “Transformer diagnostic testing by frequency response analysis,” IEEE Trans. Power App. Syst., vol. AP-97, no. 6, pp. 2144–2153, Nov. 1978.

S. Alshahiabi, Y. Khan, A. Beroual, and N. Malik, “A review of frequency response analysis methods for power transformer diagnostics,” Energies, vol. 9, no. 11, p. 879, Oct. 2016.

M. H. Samimi, S. Tenbohlen, A. A. S. Akmal, and H. Mohseni, “Improving the numerical indices proposed for the FRA interpretation by including the phase response,” Int. J. Electr. Power Energy Syst., vol. 83, pp. 585–593, Dec. 2016.

L. Armi and S. Fekri-Ershad, “Texture image analysis and texture classification methods—A review,” Int. J. Image Process. Pattern Recognit., vol. 2, no. 1, pp. 1–29, 2019. [Online]. Available: https://www.researchgate.net/publication/332438968_Texture_image_analysis_and_texture_classification_methods_A_review

T. Ojala, M. Pietikäinen, and D. Harwood, “A comparative study of texture measures with classification based on featured distributions,” Pattern Recognit., vol. 29, no. 1, pp. 51–59, Jan. 1996.

F. Tajeri and S. F. Ershad, “Porosity detection by using improved local binary patterns,” in Proc. 11th WSEAS Int. Conf. Electron., Hardware, Wireless Opt. Commun., Proc. 11th WSEAS Int. Conf. Signal Process., Robot. Autom., Proc. 4th WSEAS Int. Conf. Nanotechnol., 2012, pp. 116–121.

R. M. Haralick, K. Shanmugam, and I. Dinstein, “Textural features for image classification,” IEEE Trans. Syst., Man, Cybern., vol. SMC-3, no. 6, pp. 610–621, Nov. 1973.

R. M. Haralick, “Statistical and structural approaches to texture,” Proc. IEEE, vol. 67, no. 5, pp. 786–804, May 1979.

J. Jing and S. Yong, “An improved adaptive genetic algorithm,” (in Chinese), Comput. Eng. Appl., vol. 41, no. 18, pp. 64–69, 2005.

Z. Zhao, C. Tang, C. Yao, Q. Zhou, L. Xu, Y. Gui, and S. Islam, “Improved method to obtain the online impulse frequency response signature of a power transformer by multi scale complex CWT,” IEEE Access, vol. 6, pp. 48934–48945, 2018.

Z. Zhao, C. Yao, C. Li, and S. Islam, “Detection of power transformer winding deformation using improved FRA based on binary morphology and extreme point variation,” IEEE Trans. Ind. Electron., vol. 65, no. 4, pp. 3509–3519, Apr. 2018.

H. Gu, Z. Zhao, Y. Li, C. Li, S. Zhai, and X. Wang, “The recognition of power transformer winding deformation type based on SVM,” in Proc. IEEE Int. Conf. Mechatronics Autom., Aug. 2016, pp. 1881–1885.

L. Zhou, J. Jiang, X. Zhou, Z. Wu, T. Lin, and D. Wang, “Detection of transformer winding faults using FRA and image features,” IET Electr. Power Appl., vol. 14, no. 6, pp. 972–980, Jun. 2020.
CHAO TANG (Member, IEEE) was born in Sichuan, China, in 1981. He received the M.S. and Ph.D. degrees in electrical engineering from Chongqing University, China, in 2007 and 2010, respectively. He was with the Tony Davies High Voltage Laboratory, University of Southampton, U.K., where he was involved in some researches on the dielectric response characteristics and space charge behaviors of oil-paper insulation as a Ph.D. Student from 2008 to 2009, and a Visiting Scholar from 2013 to 2013 and from 2015 to 2016. He is currently an Associate Professor with the College of Engineering Technology, Southwest University, China. His research activities are mainly in the field of online monitoring of insulation conditions and fault diagnosis for high-voltage equipment.

CHENGUO YAO (Member, IEEE) was born in Nanchong, Sichuan, China. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from Chongqing University, Chongqing, China, in 1997, 2000, and 2003, respectively. He became a Professor with the School of Electrical Engineering, Chongqing University, in 2007. He has been a Visiting Scholar with Old Dominion University, USA, from 2017 to 2018. His current works include online monitoring of insulation condition and insulation fault diagnosis for HV apparatus, pulsed power technology, and its application in biomedical engineering.