IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received December 3, 2021, accepted December 10, 2021, date of publication December 13, 2021, date of current version January 6, 2022.

Digital Object Identifier 10.1109/ACCESS.2021.3135467

Transformer Fault Diagnosis Based on Multi-Class AdaBoost Algorithm

JIFANG LI[1], GENXU LI[1], CHEN HAI[1], AND MENGBO GUO[2]
[1]College of Electricity, North China University of Water Resources and Electric Power, Zhengzhou 450000, China
[2]State Grid Henan Electric Power Company Luoyang Power Supply Company, Luoyang 471009, China

Corresponding author: Genxu Li (lxianssr@163.com)

This work was supported by the National Natural Science Foundation of China under Grant U1804149.

ABSTRACT Traditional shallow machine learning algorithms cannot effectively explore the relationship between the fault data of oil-immersed transformers, resulting in low fault diagnosis accuracy. This paper proposes a transformer fault diagnosis method based on Multi-class AdaBoost Algorithms in response to this problem. First, the AdaBoost algorithm is combined with Support Vector Machines (SVM), The SVM is enhanced through the AdaBoost algorithm, and the transformer fault data is deeply explored. Then the dynamic weight is introduced into the Particle Swarm Optimization (PSO); through the real-time update of the particle inertia weight, the search accuracy and optimization speed of the particle swarm optimization algorithm is improved, and the improved particle swarm optimization algorithm (IPSO) is used to optimize the parameters of the SVM. Finally, by analyzing the relationship between the dissolved gas in the transformer oil and the fault type, the uncoded ratio method forms a new gas group cooperation. The improved ratio method is constructed as the input feature vector. Simulations based on 117 sets of IECTC10 standard data and 419 sets of transformer fault data collected in China show that the diagnosis method proposed in this paper has strong search ability and fast convergence speed and has a significant improvement in diagnostic accuracy compared with traditional methods.

INDEX TERMS Power transformers, fault diagnosis, support vector machine, improved particle swarm optimization, DGA feature, multi-class AdaBoost algorithm.

I. INTRODUCTION

The oil-immersed transformer is a vital component of the power grid, and they are responsible for the essential functions of power transmission and conversion. Once the fault occurs, it will cause substantial economic losses. Therefore, transformer fault diagnosis is carried out to find hidden faults in time and perform maintenance according to the type of fault. It is of great significance to reduce the loss and harm caused by transformer failure and improve the stable and reliable operation of the power grid [1]. The oil-immersed transformer in normal operation produces a minimal amount of gas due to the ageing and cracking of the insulation and is dissolved in the transformer oil. The main components of these gases are hydrogen (H_2), methane (CH_4), ethane (C_2H_4), and ethylene. (C_2H_2), acetylene (C_2H_6), carbon monoxide (CO) and carbon dioxide (CO_2), etc. [2]. When transformers have different faults, specific gas components will increase rapidly. For example, when insulating oil is overheated, the ratio of CH_4 and C_2H_4 will increase rapidly; during high-energy discharge, the contents of H_2 and C_2H_2 will increase. This shows the type and type of transformer faults. The change of gas composition shows a strong correlation. Dissolved Gas Analysis (DGA) technology uses non-electrical quantities as diagnostic indicators, is not affected by electromagnetics, has good versatility, and is widely used in the online diagnosis of oil-immersed transformers.

Some experts proposed that The three-ratio [3], Rogers ratio method [4], Duval triangle method [5], [6], and other rules formed based on DGA are simple and have played a significant role. Still, they all have incomplete state coding and over-absolute coding limits. Such problems have certain limitations in practical applications [7]. In recent years, with the continuous development of artificial intelligence technology, diagnostic methods have also developed from the traditional IEC three-ratio method and improved three-ratio method to machine learning and other artificial intelligence methods, such as neural networks, Support Vector Machines (SVM),

The associate editor coordinating the review of this manuscript and approving it for publication was Tallha Akram.

Fuzzy Algorithms, Bayesian theory, Normal cloud model [8]–[13], etc. Although these methods have achieved specific diagnostic effects, they have also solved some of the problems of traditional algorithm boundaries that are too absolute and easy to overfit. However, due to transformer failure, The complex characteristics of the gas production mechanism, the small amount of sample data, and the low dimensionality of the sample data. The above-mentioned single machine learning method cannot fully dig out the connection between the transformer fault gas data, resulting in its mediocre effect in transformer fault diagnosis.

In response to this problem, the ensemble learning AdaBoost [14], [15] algorithm builds multiple weak classifiers through multiple iterations, and affects the weight of the next-generation classifier samples according to the classification results, and performs deep mining of the samples by assigning different weights to the samples, and finally, weights Voting to produce a strong classifier for transformer fault diagnosis. Zhou and others [16]–[18] used cloud diagnosis model, decision tree algorithm, extreme learning machine, etc., as weak classifiers, and then used AdaBoost algorithm for transformer fault diagnosis. Although the AdaBoost algorithm increases the diversity of samples, due to the small amount of fault samples of large oil-immersed transformers, the accuracy of cloud models, decision trees and other algorithms is related to the number of training samples, which restricts the further fault diagnosis performance. JI [19] proposes a BP-PSO-AdaBoost algorithm, which can effectively improve neural networks' accuracy and convergence speed. The SVM classifier is made up of a small number of support vectors. Its complexity depends on the number of support vectors rather than the dimension of the sample space. Even if the sample size is small, it can have a good classification effect, so it is widely used in transformer fault data classification. Application [20]. However, the setting of SVM hyperparameters requires prior empirical knowledge, and the optimal hyperparameter selection is still an open issue in related research fields. Zhang [8] used the improved krill algorithm and genetic algorithm to set the hyperparameters of SVM and achieved good results. Still, the optimization efficiency and accuracy of hyperparameters need to be improved. On this basis, this paper proposes an Improved Particle Swarm Optimization (IPSO) to optimize SVM core parameters and penalty factors. It combines the AdaBoost algorithm with SVM to obtain multiple IPSO-SVM weak classifiers through iteration, And other conduct in-depth mining of transformer fault data to improve the classification effect of SVM.

II. OIL-IMMERSED TRANSFORMER FAULT DIAGNOSIS MODEL

AdaBoost and SVM algorithms are machine learning algorithms suitable for classification. SVM is used as a weak classifier to pre-classify transformer fault data. However, it isn't easy to have a good effect due to transformer fault data's complex gas generation mechanism. We use the AdaBoost algorithm to enhance SVM. The basic principle of the AdaBoost algorithm is to train multiple weak classifiers and assign a weight to each classifier. The classification results of all classifiers are combined by weighting to form A strong classifier. The key to this algorithm is how to train and assign weights to each weak classifier fully.

A. ADABOOST ALGORITHM

AdaBoost trains the first weak classifier by assigning initial weights to the training samples. According to the classification results of the weak classifier samples after training, the sample weights are dynamically updated to make misclassified samples receive more attention, and then based on the overall weak classifier The test error is used to adjust the weight of the weak classifier, so as to change the training process of the latter classifier, and train all the classifiers one by one, and obtain the strong classifier according to the final weight of the weak classifier. For a binary classification model with T weak classifiers, when the number of training samples is n, the strong classifier obtained by the integration is:

$$F(x) = sign\left(f(x)\right) = sign\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right) \tag{1}$$

In the formula, α_t is the weight of the weak classifier, and $h_t(x)$ is the classification result of the weak classifier.

B. WEAK CLASSIFIER MODEL BASED ON SVM ALGORITHM

The sample size of transformer fault data is small, the fault gas collection is complex, and the complicated gas production mechanism. Multi-layer machine learning algorithms such as neural networks have achieved good results in many fields. Still they require a large amount of sample data and are unsuitable for transformers fault diagnosis, and SVM has a good classification effect on small samples of multi-dimensional data. The AdaBoost algorithm model based on the SVM weak classifier is shown in Figure 1.

Standard SVM is a typical two-class classifier, and transformer fault diagnosis is a linear and inseparable multi-classification problem, so the nonlinear, multi-class transformation of SVM is required. SVM hopes to find a hyperplane that can maximize classification. While ensuring the correct classification, it also ensures that each sample point can be far enough away from the hyperplane. For this reason, the objective function of the SVM nonlinear model is:

$$\begin{aligned} &\min \phi(\omega, \xi) = \frac{1}{2}\|\omega\|^2 + c\sum_{i=1}^{l} \xi_i \\ &s.t. \begin{cases} y_i\left[\omega^T \varphi(x_i) + \lambda\right] \geq 1 - \xi_i \\ \xi_i \geq 0, \quad i = 1, 2, \ldots, l \end{cases} \end{aligned} \tag{2}$$

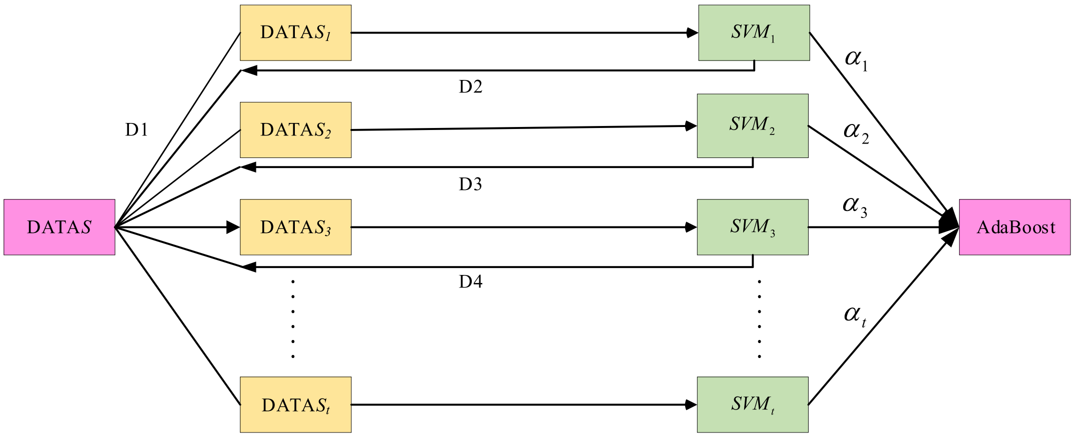

FIGURE 1. AdaBoost model.

where ξ_i is the relaxation vector and c is the penalty factor, the Lagrange function is as follows:

$$L(\omega, \lambda, \xi, \alpha, \beta) = \phi(\omega, \xi) - \sum_{i=1}^{l} \alpha_i \times \left\{ y_i \left[\omega^T \varphi(x_i) + \lambda \right] - 1 + \xi_i \right\} - \sum_{i=1}^{l} \beta_i \xi_i \tag{3}$$

The decision function of SVM is as follows:

$$f(x) = sign\left[\sum_{i=1}^{l} \alpha_i y_i K(x, x_i) + \lambda \right] \tag{4}$$

where $K(x, x_i)$ is the kernel function. The commonly used kernel functions are radial basis function (RBF), polynomial function and so on. The RBF function only needs to determine one parameter, which is conducive to optimising of parameters. Therefore, if RBF function is selected as the kernel function of SVM, the following results can be obtained:

$$K(x, x_i) = \exp\left(-g \|x - x_i\|^2 \right), \quad g > 0 \tag{5}$$

There are five types of transformer fault, and SVM is only a two classification classifier, so it is necessary to extend SVM to multi-classification. Considering the complexity and diversity of transformer fault data, it is difficult to distinguish all faults by one classification. Multiple classifications are needed to distinguish all kinds of faults accurately. The expansion mode of multi-classification is shown in Figure 2.

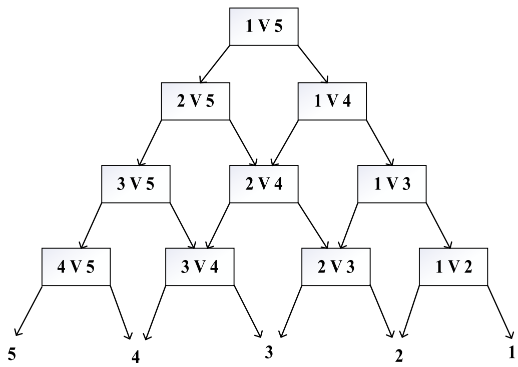

FIGURE 2. Multi-class extension.

C. FAULT DIAGNOSIS MODEL OF OIL IMMERSED TRANSFORMER BASED ON BASIC ON MULTI-CLASS ADABOOST

In the AdaBoost model, each sample is given the same initial weight and iterated many times. According to the error of each weak classifier, the weight coefficient of the weak classifier and the weight of the test sample is dynamically adjusted to increase the weight proportion of the classification error samples. The traditional AdaBoost model mainly gives weight to the two classification samples. Due to the diversity of transformer faults, the traditional AdaBoost model is improved, and a multi- classification AdaBoost model is proposed to adapt to transformer faults.

Set the transformer fault type as k, after training the t-th weak classifier SVM$_t$, first verify the classifier to ensure that the classification accuracy can meet the requirements of AdaBoost, and judge whether the sum of the weights of all samples with correct classification in each type of fault is greater than the sum of the weights of other types of fault samples.

$$\eta\left[SVM_t(x_i) = j\right] \geq \forall \eta\left[SVM_t(x_i) \neq j\right] \tag{6}$$

When RBF is selected as the kernel function, only two parameters of SVM need to be determined at this time, the penalty factor c in formula (3) and the kernel parameter g in formula (6).

Where η is the sum of weights and η is:

$$\eta = \sum_{i=1}^{n} D_t(i) \tag{7}$$

where $D_t(i)$ is the weight of training samples in the i-generation SVM. If formula (7) is not true, then retrain the weak classifier. If formula (7) is true, Calculate the classification error rate of SVM_t to training set S_t:

$$err_t = \sum_{i=1}^{n} D_t^i I\left[SVM_t\left(x_i\right) \neq y_i\right] \tag{8}$$

where $I[*]$ is the logical value, and then update the weak classifier weight α_t according to the err_t.

$$\alpha_t = \ln \frac{1 - err_t}{err_t} + \ln\left(K - 1\right) \tag{9}$$

Then the next-generation training sample weight $D_{t+1}(i)$ can be updated as:

$$D_{t+1}(i) = \frac{D_t(i)}{Z_t} \cdot \exp\left(\alpha_t \cdot I\left[SVM_t\left(x_i\right) \neq y_i\right]\right) \tag{10}$$

where Z_t is the normalization factor, which can be calculated as:

$$Z_t = \sum_{i=1}^{n} D_t(i) \cdot \exp\left(\alpha_t \cdot I\left[SVM_t\left(x_i\right) \neq y_i\right]\right) \tag{11}$$

According to formula 11, is adjusted by, so in the weak classifier, the wrongly classified sample will have a higher sample weight in the round of iterations. A weighted vote is performed on the results of each weak classifier, and the strong classifier is integrated according to the weight of the weak classifier. The final classification result is:

$$F(x) = \arg\max_{y \in Y} \sum_{t=1}^{T} \alpha_t \cdot \left[SVM_t\left(x_i\right) = y\right] \tag{12}$$

The specific algorithm flow is shown in Figure 3:

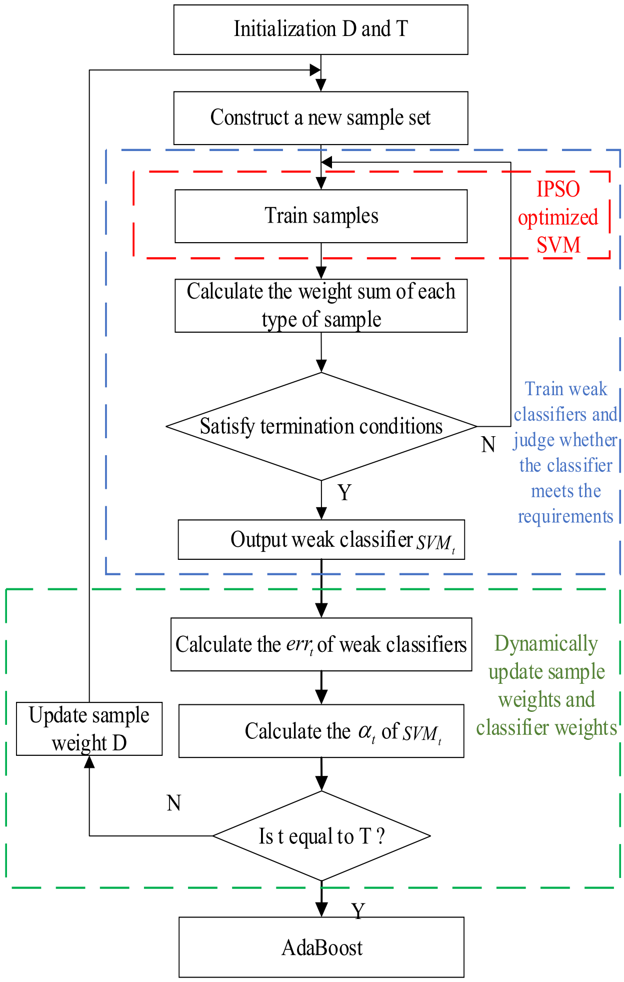

FIGURE 3. **Transformer fault diagnosis model.**

III. PARAMETER OPTIMIZATION OF WEAK CLASSIFIER BASED ON IMPROVED PARTICLE SWARM ALGORITHM

The parameter selection of SVM is related to its classification accuracy. Selecting the appropriate super parameters in an extensive range is the core of SVM model optimization. In this paper, an Improved Particle Swarm Optimization (IPSO) algorithm is proposed to optimize the parameters of SVM to obtain the IPSO-SVM model.

In the traditional particle swarm optimization algorithm, each particle searches the optimal solution individually in the optimization space, recorded as the individual extremum and shared with other particles in the whole particle swarm. The individual extremum with the optimal value is regarded as the current global optimal solution of the whole particle swarm. In the process of optimization, the search ability of the algorithm is unstable, and it is easy to fall into the local optimum. It often needs multiple iterations to find the optimal solution. Its search ability depends on the inertia weight. The larger the value is, the stronger the global search ability is, and the smaller the value is, the stronger the local search ability is. The quantitative inertia weight of the traditional particle swarm optimization algorithm is improved to time-varying inertia weight, and the inertia weight is linearly reduced from the maximum value to the minimum value by using the linearly varying weight, so as to facilitate the accurate search of the current search area and facilitate the convergence of the algorithm.

For the traditional particle swarm optimization algorithm, in each iteration, the particle updates itself through two extrema P_{best} and g_{best}, and the velocity and position of the particle can be self updated as follows:

$$\begin{aligned} v_i = \omega \times v_i + c_1 \times rand() \times \left(p_{best_i} - x_i\right) + c_2 \times rand() \\ \times \left(g_{best_i} - x_i\right) \end{aligned} \tag{13}$$

$$x_i = x_i + v_i \tag{14}$$

where $i = 1, 2, \cdots, n$ is the number of particle swarm, v_i is the speed of the i-th particle, x_i is the position of the i-th particle, c_1, c_2 are the learning factor, which determines the search ability of particles, $rand()$ is a random number between 0-1, ω is the inertia weight, which can be updated as follows:

$$\omega = \omega_{\max} - \frac{m * \left(\omega_{\max} - \omega_{\min}\right)}{m_{\max}} \tag{15}$$

where m is the number of iterations. According to the formula, with the increase of the number of iterations m, the

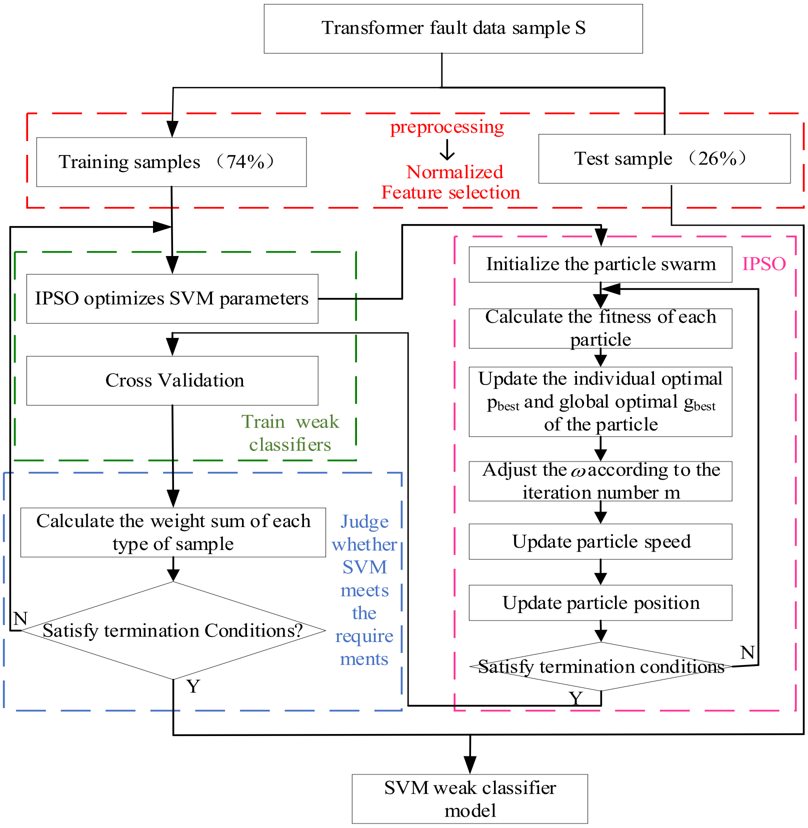

FIGURE 4. SVM weak classifier model based on IPSO.

inertia weight ω will decrease. With the increase of the number of iterations, the search center of ipso algorithm will change from global search to accurate search.

The classification ability of the weak classifier RBF-SVM depends on the penalty factor c and the kernel parameter g. The penalty factor c is the SVM's tolerance to the relaxation vector. If the value is too small, the trainer will give up the classification of the relaxation vector and pay more attention to the whole sample. If the value is too large, it will pay too much attention to the relaxation vector and lead to the wrong classification of other samples. The smaller the kernel parameter g is, the more detailed the classification will be, that is to say, the easier it will lead to overfitting; The larger the size, the more careless the classification will be, making it impossible to distinguish the data. In this paper, the parameters of RBF-SVM are optimized by the IPSO algorithm, and the IPSO-SVM weak classifier model is established, as shown in Figure 4.

IV. FEATURE VECTOR SELECTION

DGA data of oil immersed transformer includes H_2, CH_4, C_2H_6, C_2H_4, C_2H_2, CO and CO_2. In order to reduce the influence of data error on diagnosis accuracy, the selection of eigenvector is also very important. According to the power grid equipment condition based maintenance regulations and field experience, the characteristic gases produced by overheating fault are CH_4 and C_2H_6, the sum of which generally accounts for more than 80% of the total hydrocarbon [21], [22]. With the increase of fault point temperature, the proportion of C_2H_6 will increase. Generally, C_2H_2 will not be produced when the temperature is below 200 °C [23]. When the temperature is below 500 °C, the content of C_2H_2 does not exceed 2% of the total hydrocarbon. When the temperature is seriously overheated, the content of C_2H_2 does not exceed 6%. When the overheat fault involves solid insulating materials, in addition to the above gases, a lot of CO and CO_2 are also produced [24], so CO and CO_2 have a great influence on the overheat fault diagnosis results. High energy discharge fault gas production speed is fast, gas volume is large, fault characteristic gas is mainly C_2H_2 and H2, followed by a large number of C_2H_6 and CH_4 [25], C_2H_2 generally accounts for 20-70% of the total hydrocarbon, H_2 accounts for 30-90%, in most cases, C_2H_6 content is higher than CH_4. Due to the low discharge energy, the total hydrocarbon content of low-energy discharge fault is generally low, and its main component is H_2, followed by CH_4. When the discharge energy density increases, C_2H_2 will also be produced, but the proportion of C_2H_2 in total hydrocarbon is generally less than 2%, which is the main difference between high-energy discharge fault and low-energy discharge fault [25]. According to DL/T 722-2000 and IEC 60599-2015, the internal faults of transformers are divided into five types: medium and low temperature overheating (T1-T2), high temperature overheating (T3), low energy discharge (D1), high energy discharge (D2) and partial discharge (PD). Based on the collected 419 sets of domestic transformer fault data, the three-dimensional visualization of fault types and some DGA indexes is drawn, as shown in Figure 5. Fault types 1-5 in the figure represent PD, D1, D2, T1, T2 and T3 faults respectively. It can be seen from the figure that the H_2 concentration of most discharge faults is very high, while the distribution of thermal faults in the visualization diagram is relatively scattered, and there is no obvious rule. The discharge fault can be judged by the H_2 concentration, and the CH_4 concentration produced by partial discharge is relatively low. Similarly, by drawing three-dimensional visualization distribution maps of other DGA indexes, the concentrations of C_2H_6 and C_2H_2 can effectively distinguish high-energy discharge and low-energy discharge faults, the temperature range of thermal fault can be determined by C_2H_2 concentration, the concentration of C_2H_4 is higher in high-energy discharge and low-energy discharge faults, but the contents of partial discharge and thermal fault are lower. Among them, five kinds of gases H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2 are selected as eigenvectors, and their gas concentrations are recorded as $C(H_2)$, $C(CH_4)$, $C(C_2H_6)$, $C(C_2H_4)$ and $C(C_2H_2)$.

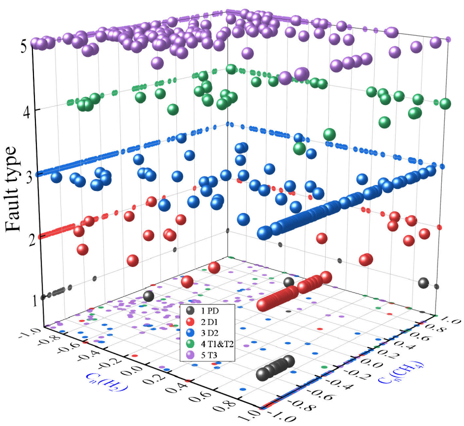

FIGURE 5. **Domestic data visualization.**

Because of the large difference of the gas production, it may affect the accuracy of fault diagnosis directly as the model input, so it is necessary to normalize the data and take the normalized gas concentration as the primary feature.

Although there is a certain relationship between different types of gas, only using the above normalized gas volume concentration as the data feature can not obtain accurate fault diagnosis results. In the field of DGA fault diagnosis, gas fraction ratio is generally used to represent more extensive feature information. Through a large number of experiments and literature review, the traditional ratio method is improved, and a new ratio method is proposed. The traditional ratio method usually needs coding, while the new ratio method only needs gas concentration ratio. The relationship between characteristics and fault types can be better reflected by the percentage of key gas in total gas or total hydrocarbon concentration, The above gases were combined in nine ratios. The DGA eigenvectors based on the new ratio method are shown in Table 1. Among them, features 1-3 are the three ratios of the three ratio method, and features 4-7 are the ratios of four kinds of carbonaceous gases CH_4, C_2H_2, C_2H_4 and C_2H_6 to total hydrocarbon respectively. The ratio of single gas to total hydrocarbon concentration can better reflect the relationship between various fault types. For example, the percentage content of C_2H_2 can determine which overheating fault the transformer is in, In feature 8-9, C_2H_4 concentration and CH_4 concentration can effectively distinguish partial discharge and other two kinds of discharge faults. The concentration of H_2 is the key to distinguish all discharge faults, Let $C_n(C_xH_x)$ be the total concentration of all carbonaceous gases.

$$C_n(C_xH_x) = C_n(CH_4) + C_n(C_2H_2) + C_n(C_2H_4) + C_n(C_2H_6) \tag{16}$$

The first part of the simulation experiment first introduces the type of transformer data used in the simulation and its source; The second part uses the global search algorithm and the IPSO algorithm to optimize the hyperparameters of the SVM. The purpose is to better exert the classification ability of the weak classifier SVM, and compare the Grey Wolf Optimizer(GWO) and PSO Algorithm illustrates the superiority of IPSO algorithm; The third part compares different input feature vectors to illustrate the superiority of the improved ratio method; the fourth part combines the AdaBoost algorithm with the SVM optimized by the IPSO algorithm to the final classification model and compares it with the model proposed in the literature.

A. EXAMPLE SAMPLE

Transformer faults are divided into internal faults and external faults. The article mainly studies the five types of internal faults defined in DL/T 722-2000 guidelines and IEC 60599-2015. The sample of the calculation example is 117 sets of IECTC10 standard data and 419 sets of transformer fault data collected in China. During the simulation, the 117 sets of IECTC10 fault data are divided into two parts, including 87 training samples and 30 test samples, which are used to test and compare AdaBoost The classification ability and fault diagnosis performance of the algorithm, and the generalization performance of the diagnosis algorithm is verified through 419 sets of transformer fault data in China. Table 2 shows the distribution of fault types in the IEC TC 10 standard data set and the domestic 419 sets of transformer fault data sets.

B. SVM PRARMETER OPTIMIZATION

In order to more clearly show the impact of penalty factors c and kernel parameters g on the accuracy of SVM

TABLE 1. DGA characteristics based on improved ratio method.

Number	DGA feature	Number	DGA feature
1	$C_n(C_2H_4)/C_n(C_2H_2)$	8	$C_n(C_2H_4+CH_4)/C_n(C_xH_x)$
2	$C_n(C_2H_4)/C_n(C_2H_6)$	9	$C_n(H_2)/C_n(H_2+C_xH_x)$
\3	$C_n(CH_4)/C_n(H_2)$	10	$C_n(CH_4)/C_n(H_2+C_xH_x)$
4	$C_n(CH_4)/C_n(C_xH_x)$	11	$C_n(C_2H_2)/C_n(H_2+C_xH_x)$
5	$C_n(C_2H_2)/C_n(C_xH_x)$	12	$C_n(C_2H_4)/C_n(H_2+C_xH_x)$
6	$C_n(C_2H_4)/C_n(C_xH_x)$	13	$C_n(C_2H_6)/C_n(H_2+C_xH_x)$
7	$C_n(C_2H_6)/C_n(C_xH_x)$		

TABLE 2. Sample data.

Fault type	T1 \T2	T3	D1	D2	PD	Total
IEC data	16	18	26	48	9	117
Domestic data	45	159	67	129	21	419

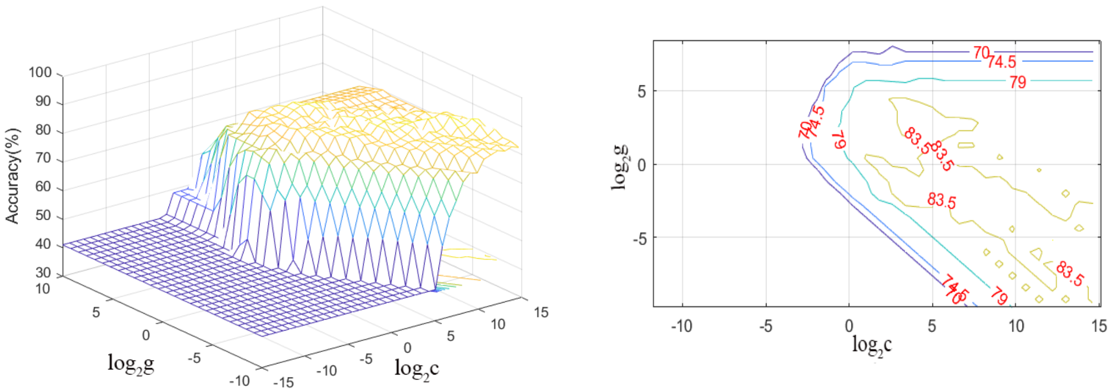

FIGURE 6. Global search algorithm of SVM in preliminary parameters optimization.

and reduce the optimization range of the IPSO algorithm, a global search algorithm is used to initially optimize the range in which the hyperparameters may have optimal solutions, and the penalty factors and kernel parameters are used for preliminary optimization. The logarithmic form of is the coordinate axis, and the curve plane and contour map of the model diagnosis accuracy rate at different parameters are shown in Figure 6. In the figure, the deeper the yellow, the higher the SVM model diagnosis accuracy rate, and the optimization of SVM parameters The better the effect, the darker the purple, the lower the diagnostic accuracy of SVM.

Based on Figure 6, it can be seen that the training effect is generally better when the penalty factor is $[2^3, 2^{10}]$, and the optimal value of the kernel parameter is $[2^{-5}, 2^5]$. This range is used as the boundary of the IPSO algorithm for optimization.

The proposed IPSO algorithm is used to accurately optimize the SVM parameters. The parameters of the IPSO algorithm during simulation are initialized as follows: the population size is 50, the maximum number of iterations is 50, the learning factor C_1 is 1.5, C_2 is 1.7, and the ω_{max} is 0.9. The ω_{min} is 0.4. The curve of the number of iterations and fitness is shown in Figure 7.

It can be seen from Figure 7 that in the initial iteration of the IPSO algorithm, the inertia weight takes the maximum value. At this time, the search center of gravity is searching for the whole world. With the increase of the number of iterations, the weight will decrease, and the local optimization ability will be gradually strengthened to benefit the algorithm. The search accuracy and convergence speed. The weight of the algorithm changes constantly in the iterative process, realizes self-update, and verifies the performance of the improved particle swarm algorithm.

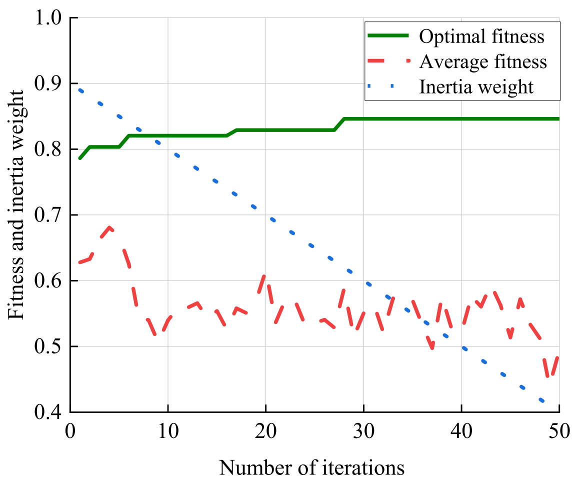

FIGURE 7. Ipso parameter optimization.

By comparing with the optimized SVM of GWO algorithm and PSO algorithm proposed by others, the result is shown in Figure 9. And use five-fold cross-validation to further increase the credibility of the results.

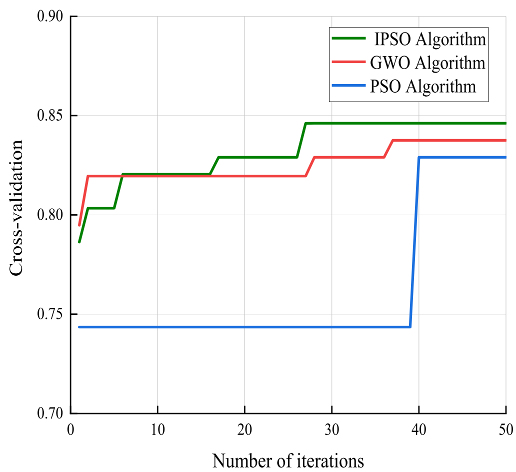

FIGURE 8. Different optimization algorithms comparison.

It can be seen from Figure 8 that the accuracy of the three algorithms are similar. Among them, The traditional PSO has the lowest diagnosis accuracy and convergence speed among the three algorithms. The IPSO algorithm has the highest accuracy, and the convergence is relatively fast, and the most stable. Therefore, it can be concluded that the IPSO algorithm is much better than the PSO algorithm.

C. DGA FEATUREVECTOR OPTIMIZATION SIMULATION

From the extensive analysis in the third chapter, it can be seen that the direct fault diagnosis of the DGA raw data will lead to the low diagnosis accuracy of the model. This article compares the influence of different feature vectors on the accuracy of fault diagnosis, and first divides the input feature vectors into three types: (1) DGA full data, including H_2, CH_4, C_2H_2, C_2H_6, C_2H_4, CO, CO_2, and total hydrocarbons; (2) the three-ratio characteristic quantity is composed of CH_4/H_2, C_2H_4/C_2H_6 and C_2H_2/C_2H_4; (3) the improved ratio method proposed in this article. Then, the fault diagnosis method based on the IPSO-SVM multi-class AdaBoost proposed in this paper is used as a model, and the fault diagnosis accuracy curve with three different input feature vectors is shown in Figure 10. Among them, the green line, red line, and blue line respectively represent the diagnosis accuracy rate of the improved ratio method, the three ratio method and the DGA gas as the input feature vector.

TABLE 3. AdaBoost algorithm validation accuracy.

Processing method	Accuracy
DGA	60%
Three ratio method	73.3%
Method of this paper	90%

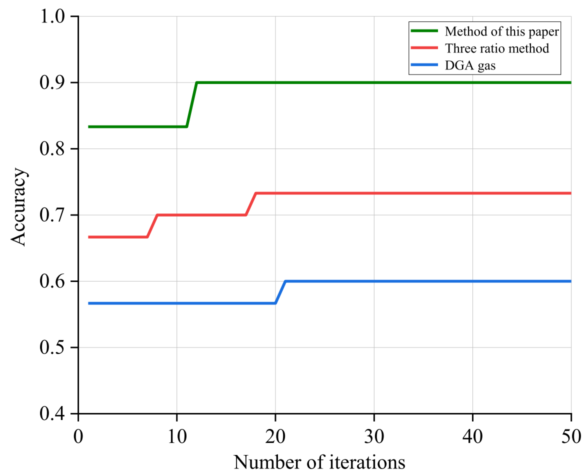

FIGURE 9. Comparison of different feature.

Figure 9 shows that with the improved ratio method as the input feature vector, the error is minimized and stabilized after several iterations of the algorithm. Compared with the traditional three ratio method and the DGA feature gas feature vector, the diagnosis is more accurate and the number of iterations is also greater. Less, the convergence speed is faster and more stable. The correct rate of transformer fault diagnosis obtained by the three methods is shown in Table 3.

It can be seen from Table 3 that the accuracy rate of the improved ratio method is 30% higher than that of the DGA full data and 16.7% higher than the three ratio method, indicating that the improved ratio method can significantly improve the accuracy of transformer fault diagnosis and has better stability for fault diagnosis, And can effectively reduce the interference caused by different data, verifying the effectiveness of the proposed improved ratio method as the input feature vector.

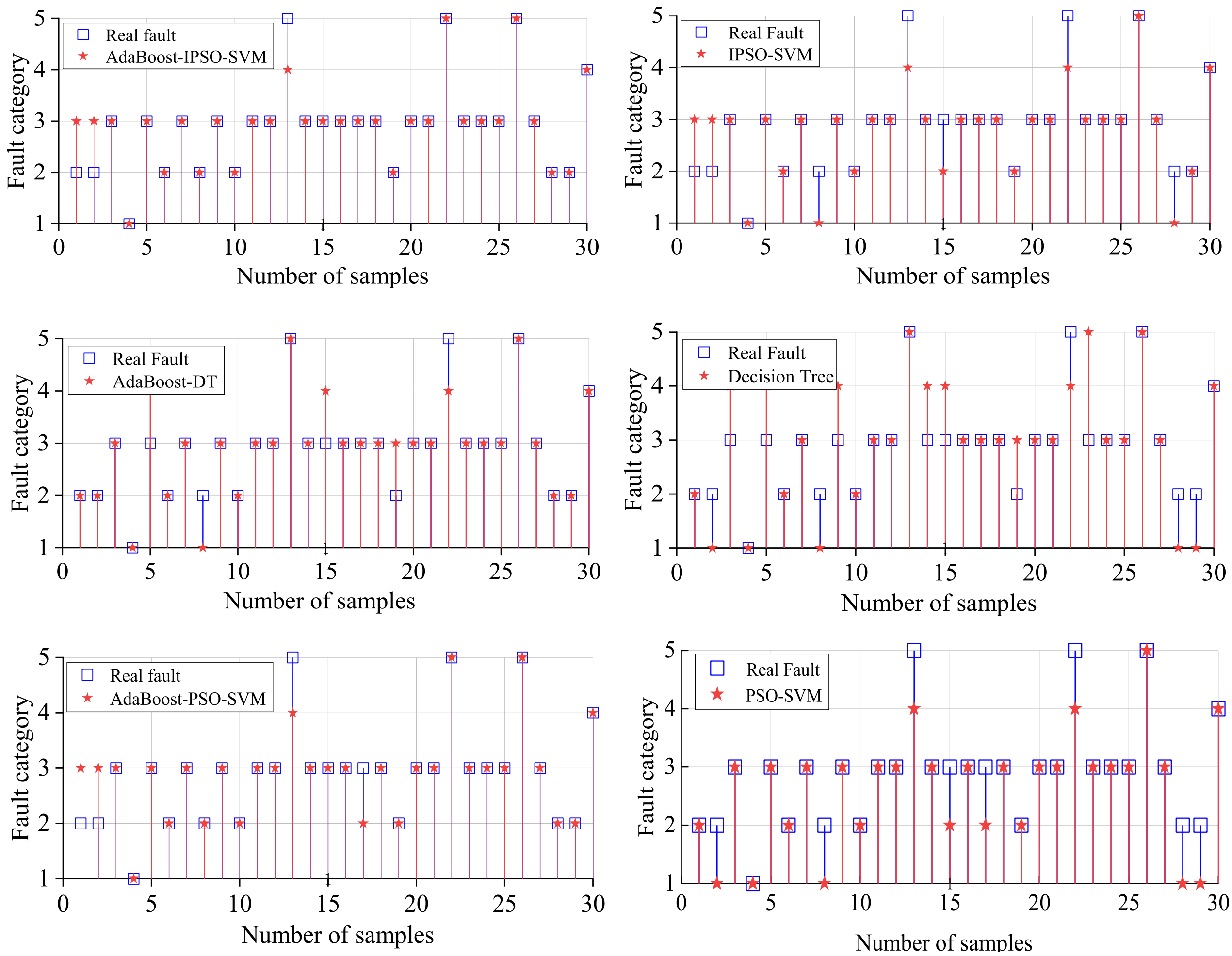

FIGURE 10. The diagnosis results of each fault model.

TABLE 4. Classification accuracy of different algorithms.

Algorithm	Accuracy	PD	D1	D2	T1	T1\T2
Decision Tree	66.7%	100%	50%	64.7%	100%	66.7%
AdaBoost-DT	86.7%	100%	87.5%	88.2%	100%	66.7%
PSO-SVM	76%	100%	50%	88.2%	100%	33.3%
IPSO-SVM	83.3%	100%	62.5%	94.12%	100%	33.3%
AdaBoost-PSO-SVM	86.7%	100%	87.5%	94.12%	100%	66.7%
AdaBoost-IPSO-SVM	90%	100%	87.5%	94.12%	100%	66.7%

D. FAULT DIAGNOSIS SIMULATION BASED ON IPASO-SVM MULTI-CLASS ADABOOST

Select the hyperparameters according to the optimization method in Section 4.2, set the number of weak classifiers to 50, and use the improved ratio method as the feature vector to train the IPSO-SVM multi-class AdaBoost model. At the same time, in order to verify the performance of the proposed method, based on the same improved ratio The method is the feature vector, and the decision tree, IPSO-SVM, AdaBoost-DT, PSO-SVM algorithm is used for fault diagnosis, and the diagnosis result is shown in Figure 10. The ordinates 1-5 in the figure indicate PD, D1, D2, T1\T2, and T3 failures respectively.

It can be seen from Figure 10 that the overall diagnosis effect of the AdaBoost model based on the SVM classifier is better than the diagnosis effect based on the decision tree model. The AdaBoost model based on the SVM classifier has a poor classification effect on high temperature faults and low energy discharge faults. The AdaBoost model is less effective in classifying high-energy discharges and partial discharges. Summarizing the correct fault diagnosis of the four methods, it can be seen that after the decision tree is enhanced by the AdaBoost algorithm, its correct rate is 86.7%, which is an increase of 20% compared with the 66.7% before the enhancement. After the IPSO-SVM algorithm proposed in this paper is enhanced by the multi-class AdaBoost algorithm, its diagnostic accuracy is up to 90%, which is 6.7% higher than the pure IPSO-SVM algorithm, and 3.3% higher than the traditional version of AdaBoost-DT and AdaBoost-PSO-SVM. It shows that the multi-class AdaBoost model based on IPSO-SVM has better fault diagnosis performance. At the same time, it can also be seen that the performance of the IPSO algorithm is better than that of the PSO algorithm.

Using the IPSO-SVM-based multi-class AdaBoost algorithm proposed in this article, the five-fold cross-validation method is used to simulate and analyze the collected domestic 419 sets of transformer fault data. and the same 419 sets are used The transformer fault data is compared with the AdaBoost-PSO-SVM algorithm, and the diagnosis results and correct rates of different fault types are shown in Figure 11.

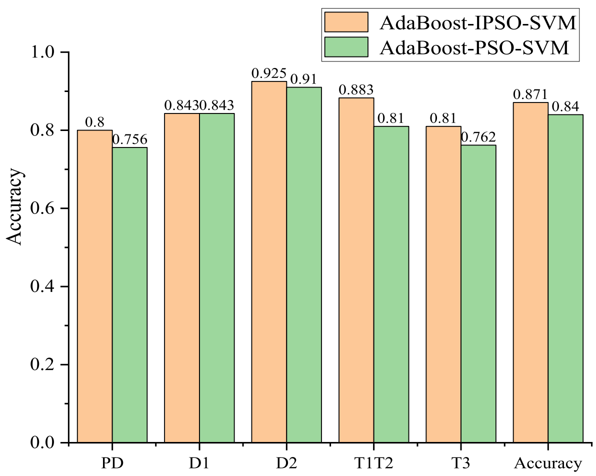

FIGURE 11. Generalization testing.

It can be seen from Figure 11 that compared with the AdaBoost-PSO-SVM algorithm, the enhanced AdaBoost-IPSO-SVM model has significantly improved the diagnostic accuracy of the five types of faults, and its diagnostic accuracy of the domestic transformer fault data is 87.1%, which is similar to the IEC TC 10 fault data, which again verifies the reliability and validity of the multi-class AdaBoost model of IPSO-SVM proposed in this paper.

V. SUMMARIZE

In this paper, the AdaBoost algorithm is used to enhance the improved particle swarm optimization (IPSO) optimized support vector machine (SVM) transformer fault diagnosis model.

And by analyzing the relationship between the dissolved gas in the transformer oil and the fault type, applying the uncoded ratio method to form a new gas combination as the characteristic parameters of the fault model, and establishing an improved ratio method as the input feature vector, through the IEC TC 10 data set Perform fault diagnosis with domestic transformer data.

(1) The IPSO-SVM fault diagnosis model enhanced by the AdaBoost algorithm can effectively and accurately identify the type of transformer fault. Compared with the traditional SVM and AdaBoost algorithm, it shows higher classification accuracy.

(2) The traditional PSO algorithm is easy to fall into the phenomenon of local optimum and premature convergence during the optimization process. By introducing linearly decreasing weights to replace the quantitative weights of the traditional PSO algorithm, the search performance of the PSO algorithm is enhanced. By comparing the traditional PSO algorithm, the IPSO algorithm shows its stronger search ability.

(3) Through the analysis of the DGA fault data, the improved ratio method proposed has a significant improvement in accuracy compared with the traditional ratio method and the DGA fault gas data.

REFERENCES

[1] Y. F. Wang, *400 Cases of Fault Diagnosis of Common Electrical and Electronic Control Equipment*. Beijing, China: China Electr. Power Press, 2011, p. 1020.

[2] Y. Liu and Y. P. Ni, "Transformer fault diagnosis method based on grey correlation analysis of three ratios," *High-Voltage Technol.*, no. 10, pp. 16–17 and 27, 2002, doi: 10.3969/j.issn.1003-6520.2002.10.007.

[3] T. F. Yang, P. Liu, J. L. Li, and Y. Hu, "New fault diagnosis method of power transformer by combination of FCM and IEC three-ratio method," *Chin. J. Anal. Chem.*, vol. 33, no. 6, pp. 66–70, Aug. 2007, doi: 10.1016/S1872-2040(07)60059-0.

[4] M. Duval, "Dissolved gas analysis: It can save your transformer," *IEEE Elect. Insul. Mag.*, vol. 5, no. 6, pp. 22–27, Nov. 1989, doi: 10.1109/57.44605.

[5] *Mineral Oil-Filled Electrical Equipment in Service–Guidance on the Interpretation of Dissolved and Free Gases Analysis*, Standard UNE-EN 60599-2016.AENOR ES,AENOR, 2016.

[6] R. Rogers, "IEEE and IEC codes to interpret incipient faults in transformers, using gas in oil analysis," *IEEE Trans. Electr. Insul.*, vol. EI-13, no. 5, pp. 349–354, Oct. 1978, doi: 10.1109/TEI.1978.298141.

[7] X. Shi, Y. L. Zhu, X. G. Ning, L. Wang, Q. G. Shun, and G. Q. Chen, "Fault diagnosis of power transformer based on deep self-encoding network," *Electr. Power Autom. Equip.*, vol. 36, no. 5, pp. 122–126, 2016, doi: 10.16081/j.issn.1006-6047.2016.05.021.

[8] Y. Zhang, X. Li, H. Zheng, H. Yao, J. Liu, C. Zhang, H. Peng, and J. Jiao, "A fault diagnosis model of power transformers based on dissolved gas analysis features selection and improved krill herd algorithm optimized support vector machine," *IEEE Access*, vol. 7, pp. 102803–102811, 2019, doi: 10.1109/ACCESS.2019.2927018.

[9] C. Y. Zhang, Z. F. Lin, D. Liu, and J. L. Huang, "Transformer fault diagnosis based on normal cloud model and improved Bayesian classifier," *Elect. Meas. Instrum.*, vol. 54, no. 4, pp. 50–56, 2017.

[10] J. Y. Wu, G. Q. Xia, T. Li, B. Gao, and G. N. Wu, "Evaluation of the aging state of transformer oil-paper insulation based on time-domain dielectric method and dynamic Bayesian network," *High Voltage App.*, vol. 55, no. 10, pp. 196–203, 2019, doi: 10.13296/j.1001-1609.hva.2019.10.032.

[11] T. Yi, Y. Xie, H. Zhang, and X. Kong, "Insulation fault diagnosis of disconnecting switches based on wavelet packet transform and PCA-IPSO-SVM of electric fields," *IEEE Access*, vol. 8, pp. 176676–176690, 2020, doi: 10.1109/ACCESS.2020.3026932.

[12] H. Keskes and A. Braham, "Recursive undecimated wavelet packet transform and DAG SVM for induction motor diagnosis," *IEEE Trans. Ind. Informat.*, vol. 11, no. 5, pp. 1059–1066, Oct. 2015, doi: 10.1109/TII.2015.2462315.

[13] H. Xu and H. Yuan, "An SVM-based adaboost cascade classifier for sonar image," *IEEE Access*, vol. 8, pp. 115857–115864, 2020, doi: 10.1109/ACCESS.2020.3004473.

[14] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp. 119–139, 1995, doi: 10.1006/jcss.1997.1504.

[15] F. Wang, Z. Li, F. He, R. Wang, W. Yu, and F. Nie, "Feature learning viewpoint of adaboost and a new algorithm," *IEEE Access*, vol. 7, pp. 149890–149899, 2019, doi: 10.1109/ACCESS.2019.2947359.

[16] X. B. Huang, W. J. Z. Li, T. Song, and Y. M. Wang, "Transformer fault diagnosis based on DGA technology and SAMME," *High Voltage App.*, vol. 52, no. 2, pp. 13–18, 2016, doi: 10.13296/j.1001-1609.hva.2016.02.003.

[17] J. Liu, L. J. Zhao, L. Huang, H. R. Zeng, X. Zhang, and H. Chen, "Transformer fault diagnosis method based on adaboost.MK and SM-SVDD," *J. Electr. Power Sci. Technol.*, vol. 32, no. 3, pp. 139–144, and 152, 2017.

[18] Q. Zhou, S. Wang, R. Liao, C. Sun, H. Xie, and J. Rao, "Power transformer fault diagnosis method based on cloud model of adaboost algorithm," *High Voltage Eng.*, vol. 41, pp. 3804–3811, Nov. 2015, doi: 10.13336/j.1003-6520.hve.2015.11.039.

[19] X. Ji, B. Yang, and Q. Tang, "Acoustic seabed classification based on multibeam echosounder backscatter data using the PSO-BP-adaboost algorithm: A case study from Jiaozhou Bay, China," *IEEE J. Ocean. Eng.*, vol. 46, no. 2, pp. 509–519, Apr. 2021, doi: 10.1109/JOE.2020.2989853.

[20] M. Singh and A. G. Shaik, "Incipient fault detection in stator windings of an induction motor using stockwell transform and SVM," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 12, pp. 9496–9504, Dec. 2020, doi: 10.1109/TIM.2020.3002444.

[21] G. Wang, A. H. Gandomi, and A. H. Alavi, "A chaotic particle-swarm krill herd algorithm for global numerical optimization," *Kybernetes*, vol. 42, no. 6, pp. 962–978, 2013, doi: 10.1108/k-11-2012-0108.

[22] S. Ghoneim, I. Taha, N. Elkalashy, and D.-E. Mansour, "Conditional probability-based interpretation of dissolved gas analysis for transformer incipient faults," *IET Gener., Transmiss. Distrib.*, vol. 11, no. 4, pp. 943–951, Mar. 2017, doi: 10.1049/iet-gtd.2016.0886.

[23] I. Hoehlein-Atanasova and R. Frotscher, "Carbon oxides in the interpretation of dissolved gas analysis in transformers and tap changers," *IEEE Elect. Insul. Mag.*, vol. 26, no. 6, pp. 22–26, Dec. 2011, doi: 10.1109/MEI.2010.5599976.

[24] Z. Qian, W. Gao, F. Wang, and Z. Yan, "A case-based reasoning approach to power transformer fault diagnosis using dissolved gas analysis data," *Eur. Trans. Elect. Power*, vol. 19, no. 3, pp. 518–530, Apr. 2009, doi: 10.1002/etep.240.

[25] C. Wei, W. Tang, and C. Wu, "Dissolved gas analysis method based on novel feature prioritisation and support vector machine," *IET Elect. Power Appl.*, vol. 8, no. 8, pp. 320–328, Aug. 2014, doi: 10.1049/iet-epa.2014.0085.

JIFANG LI was born in Kaifeng, Henan, China, in 1971. He received the M.S. degree in control technology from the Huazhong University of Science and Technology, Wuhan, in 2003, and the Ph.D. degree in power electronics and power transmission from Shanghai Maritime University, Shanghai, in 2011. He is currently a Professor with the College of Electricity, North China University of Water Resources and Electric Power. His current research interests include on-line monitoring and fault diagnosis of smart grid.

GENXU LI was born in Xinyang, Henan, China, in 1995. He is currently pursuing the master's degree with the North China University of Water Resources and Electric Power. His current research interests include on-line monitoring and fault diagnosis of smart grid.

CHEN HAI was born in Zhengzhou, Henan, China, in 1996. He is currently pursuing the master's degree with the North China University of Water Resources and Electric Power. His current research interests include on-line monitoring and fault diagnosis of smart grid.

MENGBO GUO was born in Luoyang, Henan, China, in 1993. He is graduated from the North China University of Water Resources and Electric Power. He is currently working with the State Grid Henan Electric Power Company Luoyang Power Supply Company.