Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 6610338, 12 pages
https://doi.org/10.1155/2021/6610338

Research Article

An Intelligent Fault Diagnosis Method for Transformer Based on IPSO-gcForest

Kezhen Liu,[1] Shizhe Wu,[1] Zhao Luo,[1] Zeweiyi Gongze,[2] Xianlong Ma,[2] Zhanguo Cao,[2] and Hejian Li[3]

[1]*Faculty of Electric Power Engineering, Kunming University of Science and Technology, Kunming 650500, China*
[2]*Electric Power Research Institute of Yunnan Power Co., Ltd., Kunming 650217, China*
[3]*Dali Power Supply Bureau, Yunnan Power Grid Co., Ltd., Dali 671000, China*

Correspondence should be addressed to Zhao Luo; waiting.1986@live.com

Received 14 November 2020; Revised 23 January 2021; Accepted 1 February 2021; Published 11 February 2021

Academic Editor: Michal Kunicki

Transformers are the main equipment for power system operation. Undiagnosed faults in the internal components of the transformer will increase the downtime during operation and cause significant economic losses. Efficient and accurate transformer fault diagnosis is an important part of power grid research, which plays a key role in the safe and stable operation of the power system. Existing traditional transformer fault diagnosis methods have the problems of low accuracy, difficulty in effectively processing fault characteristic information, and superparameters that adversely affect transformer fault diagnosis. In this paper, we propose a transformer fault diagnosis method based on improved particle swarm optimization (IPSO) and multigrained cascade forest (gcForest). Considering the correlation between the characteristic gas dissolved in oil and the type of fault, firstly, the noncode ratios of the characteristic gas dissolved in the oil are determined as the characteristic parameter of the model. Then, the IPSO algorithm is used to iteratively optimize the parameters of the gcForest model and obtain the optimal parameters with the highest diagnostic accuracy. Finally, the diagnosis effect of IPSO-gcForest model under different characteristic parameters and size samples is analyzed by identification experiments and compared with that of various methods. The results show that the diagnostic effect of the model with noncode ratios as the characteristic parameter is better than DGA data, IEC ratios, and Rogers ratios. And the IPSO-gcForest model can effectively improve the accuracy of transformer fault diagnosis, thus verifying the feasibility and effectiveness of the method.

1. Introduction

Transformer fault will endanger the safe and stable operation of the whole power system. Transformer fault diagnosis can analyze equipment status information to ensure reliable and efficient operation of transformer equipment. Therefore, accurate identification of transformer fault types and timely maintenance can provide an important guarantee for the normal operation of the power system [1, 2].

Since the amount of dissolved gas in the oil inside the transformer tank is closely linked to the actual operating conditions of the transformer, it is necessary to use the dissolved gas analysis (DGA) technology to evaluate the condition and monitor the early discharge, overheating, and other faults of the transformer in the oil. Dissolved gas analysis in oil is mainly used in online monitoring of oil-immersed transformers [3–5]. Based on the characteristic gas of DGA for data correlation analysis, foreign researchers have proposed the IEC ratio method, Rogers ratio method [6], Dornenburg ratio method [7], and electrical cooperative research method. However, the traditional DGA method only gives the threshold discrimination boundary of fault diagnosis, which cannot show the relationship between characteristic gases and fault types. It cannot meet the requirements of actual operation of transformer [8, 9]. With the advancement and development of artificial intelligence technology, the application of machine learning methods in transformer fault diagnosis has made remarkable

achievements. Currently, expert systems [10], deep belief networks (DBN) [11–13], random forests (RF) [14], and support vector machine (SVM) [15, 16] are commonly used in transformer fault diagnosis. Although these machine learning methods are widely used in transformer fault diagnosis, there are still certain drawbacks. For example, expert systems cannot learn to autonomously work in low efficiency, and it is hard to obtain accurate diagnosis results. DBN has strong self-learning ability, but it requires a large amount of sample data for training. The learning period of DBN is long, and it is easy to be overfit. RF is easy to be overfit when dealing with multiclassification problems of transformer fault diagnosis. SVM has outstanding performance when processing small sample data, but it is essentially a two-classifier, which is inefficient when dealing with multiclassification problems such as transformer fault diagnosis. The methods used in the above literature have improved the accuracy of transformer fault diagnosis. However, the transformer faults are diverse and complex, and the use of a single intelligent fault diagnosis method has the problems of insufficient reasoning ability and low diagnostic accuracy, which makes it difficult to obtain satisfactory diagnosis results. With the continuous development of big data technology in power system and the increase of transformer fault cases, the level of fault diagnosis needs to meet higher requirements.

Multigrained cascade forest (gcForest) is a deep integrated learning model based on decision tree proposed by Zhou Zhihua in 2017 [17, 18]. The model has the advantages of high parallel learning efficiency and strong representation learning ability. It is widely used in the fields of hyperspectral image classification [19], complex machine processing status monitoring [20], turbine fault intelligent diagnosis [21], and other fields with good results. The gcForest model consists of two parts: multigrained scanning procedure and cascade forest procedure. The multigrained scanning procedure mines the feature information of the original sample data and then supervises the learning layer by layer through the cascade forest. Therefore, the generalization ability of the model is improved. Although they perform well in much application, the rationality of the architecture and its optimization remains an unresolved problem.

Another significant but rarely studied problem in machine learning based classification and regression tasks is hyperparameter optimization. The hyperparameter settings such as the multigrained scanning window size q and the maximum cascade number allowed by the cascade forest l will have a greater impact on the model diagnostic performance. Therefore, the problem of low diagnosis accuracy can be solved by adjusting random parameters through optimization algorithms and iteratively searching for the optimal parameters of the model. There are several common optimization algorithms, such as simulated annealing algorithm [22], genetic algorithm [23], Bayesian algorithm [24], and particle swarm algorithm [25]. Particle swarm optimization (PSO) algorithm is more popular in the past few years. PSO algorithm is a group optimization algorithm that simulates the bird foraging process based on the activity of bird clusters. The PSO algorithm has fewer hyperparameters, and the parameter adjustment process is simple and easy to implement that makes it suitable for optimization under dynamic and multiobjective conditions. But the PSO algorithm tends to fall into the local optimal in the optimization process, which may cause a large error result. Therefore, the use of improved particle swarm optimization (IPSO) algorithm in transformer fault diagnosis may help it a lot.

The DGA-based transformer fault diagnosis method can analyze the equipment status information and detect the potential risks of the transformer in time, which is the key to ensuring the reliable and efficient operation of the equipment. Therefore, we proposed a transformer fault diagnosis method to improve the accuracy of transformer diagnosis, in which the key parameters of gcForest model were optimized by IPSO algorithm. Firstly, the noncode ratios of the characteristic gas dissolved in oil are determined as the characteristic parameter of the model. Then, the IPSO algorithm is used to iteratively optimize parameters q and l of the gcForest model. Under the premise of the highest diagnostic accuracy, the optimal parameters of the model are obtained through continuous iteration, and the IPSO-gcForest fault diagnosis model is established. Finally, the fault characteristic information of transformer is extracted by multigrained scanning, and the cascade forest has supervised the learning to diagnose the fault type of transformer. After that, the accurate diagnosis of transformer fault type can be got. The diagnostic performance of the IPSO-gcForest model under different characteristic parameters and size samples is analyzed through calculation examples, and the effectiveness of the method is verified. And the transformer fault diagnosis method proposed in the paper is applied to the transformer condition assessment system with good practical application results. Our contributions in this paper include the following:

(1) Different strategies are used to update the inertia weight and acceleration factor of the traditional PSO algorithm in order to improve the convergence speed and search ability of the particles.

(2) Under the premise of the highest diagnostic accuracy, the IPSO algorithm is used to iterate and update automatically to find the optimal value of the parameters in the gcForest model, which overcomes the problem of low accuracy caused by the traditional empirical selection of parameters.

(3) It is proposed that using noncode ratios as the characteristic parameter of the model can significantly improve the accuracy of transformer fault diagnosis.

(4) A new intelligent data-driven transformer fault diagnosis method is proposed. The multigrained scanning process of the gcForest model mines more transformer fault feature information. And the cascade forest process integrates multiple classifiers

for parallel training layer by layer. It can ensure that features are distinguished in different operating conditions and improve the accuracy of classification.

The rest of the paper is organized as follows: In Section 2, the principle of IPSO-gcForest model is described in detail, including the gcForest model, PSO algorithm, and its improved algorithm. In Section 3, based on IPSO-gcForest model, an intelligent transformer fault diagnosis model is built. In Section 4, the robustness of the fault diagnosis method is analyzed, and the process of parameter optimization of gcForest model by IPSO algorithm is discussed. Conclusions are presented in Section 5.

2. Principle of the IPSO-gcForest Model

2.1. PSO Algorithm. To the extent feasible, the PSO algorithm constantly adjusts each particle's speed and position based on its own search experience and that of other particles. Firstly, the state of the particle is initialized. The local extreme value and the global extreme value are iteratively searched according to the fitness function of the particle. Then, it is constantly updated in the set number of iterations. The coordinates of the particles change depending on the search velocity at each iteration, which in turn depends on the inertial weight, acceleration factor, and local and global extreme values. The formula for calculating the position and velocity of each particle is shown in

$$x_{i,d}^{t+1} = x_{i,d}^{t} + v_{i,d}^{t+1}, \tag{1}$$

$$v_{i,d}^{t+1} = w^{t+1} v_{i,d}^{t} + s_1^{t+1} r_1 \left(P_{i,d} - x_{i,d}^{t}\right) + s_2^{t+1} r_2 \left(G_d - x_{i,d}^{t}\right), \tag{2}$$

where $xt\ i,d$ represents the d-dimensional coordinate component of the t iteration of the i particle; $vt\ i,d$ represents the d-dimensional velocity component of the t iteration of the i particle; w^t represents the inertia weight at the t iteration; st 1 and st 2 represent the two acceleration factors at the t iteration; r_1 and r_2 represent random values between [0, 1]; $P_{i,d}$ represents the local extreme value of the d-dimensional component of the i particle; G_d represents the global extreme value of the d-dimensional component.

2.2. IPSO Algorithm. It can be seen from formula (1) that the main factors affecting PSO algorithm update are three parameter variables: inertial weight w and acceleration factors s_1 and s_2. This paper puts forward two improvement strategies based on the traditional PSO algorithm. First, according to the iterative process and the particle's following position, the inertia weight is varied in a nonlinear differential way to balance the overall speed of the particle search and the convergence velocity [26], as shown in equations (3) and (4). Secondly, the acceleration factor is dynamically adjusted by a cosine function to promote the coordination of the overall optimization and local optimization capabilities of the particles and improve the algorithm's optimization capability [27], as shown in

$$\frac{dw}{dt} = \frac{\left(w_{\text{ini}} - w_{\text{fin}}\right)}{T_{\max}} - \frac{4\left(w_{\text{ini}} - w_{\text{fin}}\right)}{T_{\max}^2} \times t, \tag{3}$$

$$w^t = w_{\text{ini}} + \frac{\left(w_{\text{ini}} - w_{\text{fin}}\right)}{T_{\max}} \times t - \frac{2\left(w_{\text{ini}} - w_{\text{fin}}\right)}{T_{\max}^2} \times t^2, \tag{4}$$

$$s_1^t = s_{1,\text{ini}} + \left(s_{1,\text{fin}} - s_{1,\text{ini}}\right)\left(\frac{1 - \cos\, t\pi/T_{\max}}{2}\right), \tag{5}$$

$$s_2^t = s_{2,\text{ini}} + \left(s_{2,\text{fin}} - s_{2,\text{ini}}\right)\left(\frac{1 - \cos\, t\pi/T_{\max}}{2}\right), \tag{6}$$

where w_{ini} and w_{fin} represent the initial and final values of the inertia weight, respectively; t represents the current number of iterations; $T_{\max}$ represents the maximum number of iterations, $s_{1,\text{ini}}$, $s_{1,\text{fin}}$ and $s_{2,\text{ini}}$, $s_{2,\text{fin}}$ represent the initial and final values of acceleration factors s_1 and s_2, respectively.

2.3. gcForest Model. The gcForest model is composed of multigrained scanning and cascade forest. The multigrained scanning stage can extract the features of the original sample set. The cascade forest structure can adaptively determine the number of cascading layers, and it can carry on representation learning and improve the generalization ability of the model. The complete random forest and random forest [17] in the gcForest model are integrated by CART decision trees.

2.3.1. Decision Tree. The decision tree is based on examples to realize the tasks of classification and regression. In other words, it obtains classification rules by recursively analyzing the training set of the original sample set, thereby generating a decision tree to process the testing set. The decision tree is a hierarchical structure composed of nodes containing sample attributes and branches containing attribute test conditions. Starting from the root node of the decision tree, it applies the attribute test conditions to the training set, selects the appropriate branch according to the testing results, and then follows the branch to an internal node or uses the new attribute test condition to reach the leaf node. The structure of the decision tree is shown in Figure 1.

The common algorithms of decision tree are ID3, C4.5, and CART. ID3 algorithm adopts a divide and conquer strategy and uses information gain as the selection criterion of attributes. So, all subsets only contain the same kind of information. The important improvement of C4.5 algorithm for ID3 is using information gain rate to select attributes [28, 29]. The CART algorithm is the basic decision tree algorithm of the completely random forest and random forest, which uses Gini coefficient as the attribute's selection criterion.

The CART algorithm divides the training set of the original sample set into two subsets by using category k and threshold u_k, and then it minimizes the cost function $H(k, u_k)$ to generate the purest subset. During the growth of the decision tree, we select the Gini coefficient as the best division metric for the root node and internal nodes. Then, we

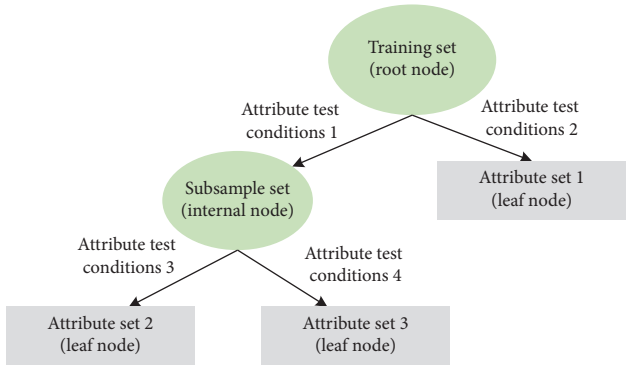

Figure 1: Structure of the decision tree.

use the Gini coefficient and the cost function to select the optimal attribute to divide the training set. After the decision tree is established, the testing set is used to prune the tree, and it can improve the generalization ability of the decision tree. The Gini coefficient and cost function are shown in

$$G_j = 1 - \sum_{k=0}^{n} {p_{j,k}}^2, \tag{7}$$

$$H(k, u_k) = \frac{y_{\text{left}}}{y} G_{\text{left}} + \frac{y_{\text{right}}}{y} G_{\text{right}}, \tag{8}$$

where $p_{j,k}$ represents the percentage of training instances in which the node j belongs to category k, $y_{left/right}$ is the number of instances of the left and right subsample sets, and $G_{left/right}$ is the measure of the impurity of the left and right subsample sets.

2.3.2. Multigrained Scanning. The multigrained scanning structure uses scan windows of different sizes to scan the original input features, which can produce many feature instances of different dimensions. Then, the feature instances corresponding to the original input features are trained by a completely random forest and a random forest to generate a class probability vector. Finally, the feature vectors are obtained by splicing to improve the representation learning ability of the model. The multigrained scanning process is shown in Figure 2.

As showed in Figure 2, the multigrained scanning phase is divided into two processes: feature scanning and feature conversion. Assume that the original input feature is of $m \times m$ dimensions, the sliding window size is of $q \times q$ dimensions, and the sliding step size is e. The scanning window extracts feature information by scanning the original input features and will generate N q-dimensional feature instances, as shown in

$$N = \left[\frac{(m-q)}{e+1} \right]^2. \tag{9}$$

If each forest outputs c-dimensional class probability vectors, after completely random forest and random forest training, all class probability vectors are connected into L-dimensional feature vectors, as shown in

$$L = 2 \times \left[\frac{(m-q)}{e+1} \right]^2 \times c. \tag{10}$$

The scale of the feature vectors obtained by the multigrained scanning is much higher than that of the original input feature vectors. Therefore, more feature information can be extracted.

2.3.3. Cascade Forest. The cascade forest is integrated deep learning based on decision trees. The cascade forest has high accuracy when processing high-dimensional data and has scalability and parallelism. The supervised learning of cascade forest layer by layer can improve the representation ability of feature information. Each layer of the cascade forest contains two completely random forest classifiers and two random forest classifiers. The combination of multiple different types of base classifiers can fully learn the feature information of the input feature vector, thereby improving the overall recognition performance of the model. The cascade forest process is shown in Figure 3.

The input feature vector of the cascade forest is the feature vector finally generated in the multigrained scanning process, and then supervised learning is carried out between cascading layers. The class vectors outputs between the cascade-forest layers are not merged before the logistic regression. The generated class vectors are spliced together with the input feature vectors as the input of the next layer. After layer-by-layer training, the final class vector is generated by logistic regression for all class vectors in the final cascade layer, from which the maximum value is taken to obtain the final classification of the original input features. In order to avoid overfitting in the cascade forest training, the completely random forest and random forest each are trained with 5-fold cross-validation to generate class vectors.

The cascade level of cascaded forest can be adaptive, and the class vector of each cascading layer is dynamically updated. The performance of the whole cascade forest is evaluated according to the testing set. If the gcForest model does not improve significantly during training within several consecutive layers, the cascade process will be terminated automatically. This process can improve the accuracy of fault diagnosis and reduce the training time, and the dynamic changes of the cascade layer can make the gcForest model suitable for different sizes of sample data. When the sample data is small, the fault feature information will be closely combined to enhance the characterization learning ability of the original input feature. When the sample data is large, the number of cascade layers will be limited to accelerate the training process of cascaded forest.

3. Transformer Fault Diagnosis Model Based on IPSO-gcForest

3.1. Characteristic Parameter Selections. When the operation and maintenance personnel analyze the abnormal

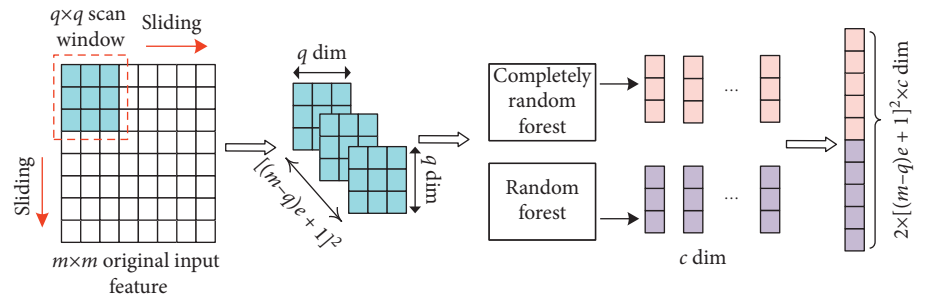

Figure 2: Multigrained scanning process.

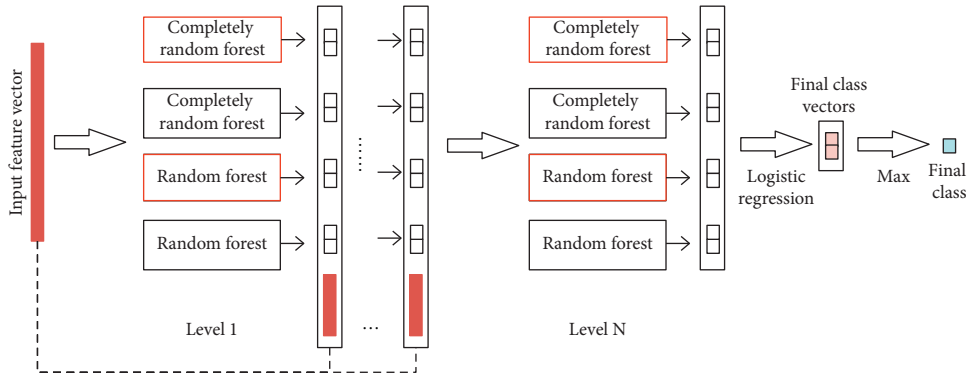

Figure 3: Cascade forest process.

conditions and failure causes of the transformer, the analysis of the dissolved gas in the oil is a vital part. Different faults in power transformers will produce different characteristic gas, but the characteristic gas content in DGA data is quite different, which has a certain impact on the diagnosis and testing of internal faults of oil-immersed transformers. Therefore, by comparing DGA data, IEC ratios (CH_4/H_2, C_2H_4/C_2H_6, C_2H_2/C_2H_4), Rogers ratios (CH_4/H_2, C_2H_2/C_2H_4, C_2H_4/C_2H_6, C_2H_6/CH_4), and noncode ratios (CH_4/H_2, C_2H_2/C_2H_4, C_2H_4/C_2H_6, $CH_4/(C_1+C_2)$, $C_2H_2/(C_1+C_2)$, $H_2/(H_2+C_1+C_2)$, $C_2H_4/(C_1+C_2)$, $C_2H_6/(C_1+C_2)$, $(CH_4+C_2H_4)/(C_1+C_2)$) as the diagnostic accuracy of the model's characteristic parameters, the input characteristic parameters of the model are determined, where C_1 is CH_4, and C_2 is the sum of C_2H_2、C_2H_4 and C_2H_6.

Since the dissolved gas content data in transformer oil is disturbed and affected by the monitoring device, ambient temperature, and personnel operations, the original data needs to be normalized. The normalization of feature quantity can reduce the impact of data on the performance of the model and improve the training speed and diagnostic accuracy of the model. In order to ensure that all feature quantities are in the same value range, it is needed to normalize the feature quantities, as shown in

$$y* = \frac{y - y_{\min}}{y_{\max} - y_{\min}}, \tag{11}$$

where $y*$ is the normalized data; $y_{\min}$ and $y_{\max}$ are the minimum and maximum of a certain dimension feature vector; and y is the original data.

3.2. IPSO-gcForest Diagnostic Model Technical Route. With its own internal structure, the gcForest model can fully mine fault feature information and accurately diagnose transformer faults. When the gcForest model is used to identify fault types, it is necessary to determine the key parameters of the model according to human experience or control variables, which may easily lead to poor diagnostic results. Thus, under the premise of satisfying the highest diagnostic accuracy, the IPSO algorithm obtains the optimal parameters of the gcForest model through continuous iterative solving, which will improve the diagnostic accuracy. The fault types of transformer can be divided into seven states: normal (N), high-energy discharge (D1), low-energy discharge (D2), partial discharge (D3), high-temperature overheating (T1), medium-temperature overheating (T2), and low-temperature overheating (T3). The fault diagnosis based on IPSO-gcForest model includes three main steps: data preprocessing, IPSO algorithm optimization parameters, and fault type identification. The whole process is shown in Figure 4, and the specific steps are shown as follows:

Step 1: the noncode ratios of the characteristic gas dissolved in the oil are determined as the characteristic parameter of the model, and then the characteristic

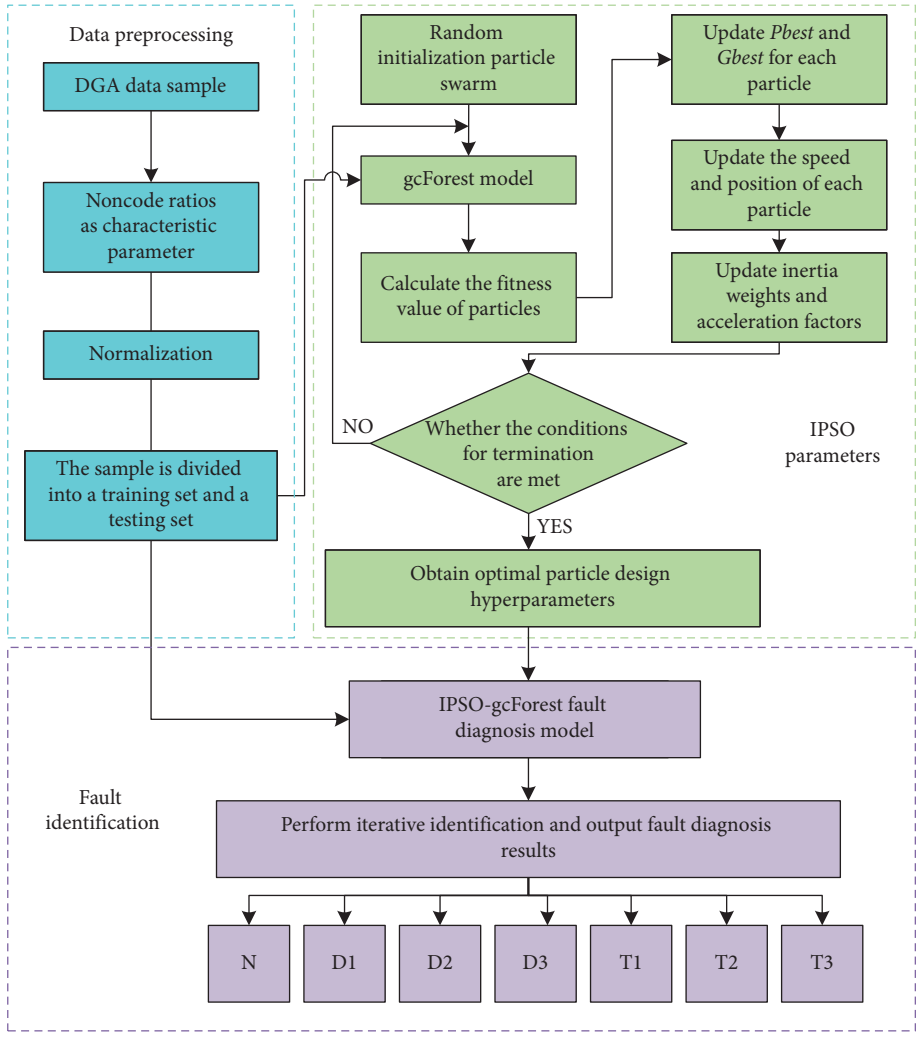

Figure 4: IPSO-gcForest model diagnostic roadmap.

parameter is normalized. According to the model testing requirements, the original sample was randomly divided into a training set and a testing set at a ratio of 8 : 2.

Step 2: initialize the population particles randomly, and set the value range and search range of q and l. Then, the number of particles and the maximum number of iterations can be determined.

Step 3: build the gcForest model based on the values of initialized q and l. The training set and the testing set are used to train and diagnose gcForest, respectively, and then the diagnostic accuracy of the training set is used as the fitness value of the particles.

Step 4: the local extreme values and global extreme values of particles are determined according to the initial fitness of particles, and the velocity and position of particles are updated by using equations (1) and (6). The corresponding particle fitness values are calculated and compared with local extreme value and global extreme value. The new local extreme value and global extreme value are determined to achieve the highest diagnosis and recognition accuracy.

Step 5: when the particle fitness value tends to be stable or the number of iterations reaches a preset value, the particle iteration optimization is stopped to obtain the optimal parameters. Otherwise, return to step 4.

Step 6: the IPSO-gcForest fault diagnosis model is constructed based on the optimal parameters obtained from the IPSO algorithm, and the diagnosis results are analyzed comprehensively with the evaluation index.

3.3. Model Evaluation Index. To validate the diagnostic performance of the IPSO-gcForest model, the diagnostic accuracy, precision, and recall rate were used as evaluation indexes to analyze the diagnosis results of the model.

The diagnostic accuracy represents the ratio of the number of correct fault samples to the total number of samples, which can directly evaluate the generalization ability of the model.

The diagnostic precision refers to the proportion of correctly identifying class A fault samples and all the fault samples identified as class A fault samples, indicating the precise detection of class A fault samples. The definition of diagnostic precision is shown in

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{12}$$

TP is the number of true positives, and FP is the number of false positives.

The diagnostic recall rate refers to the percentage of correctly identified class A fault samples and actual class A fault samples, indicating whether all class A fault samples have been checked. The definition of diagnostic recall rate is shown in

$$\text{recall rate} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{13}$$

FN is of course the number of false negatives.

4. Transformer Fault Diagnosis Model Based on IPSO-gcForest

This paper collects fault sample data of transformer voltage level from 35 kV to 500 kV, from the transformer online monitoring data and historical fault data of China Southern Power Grid Corporation, the transformer fault oil chromatographic data in published papers, the "Typical Cases of Application of Power Grid Equipment Detection Technology" published by the State Grid and IEC TC 10 database. All the above data samples comprise 1601 cases of transformer fault data. In this paper, the training set and testing set are divided at the proportion of 8 : 2. Among them, 1280 cases received supervised training to adjust the parameters of the model to improve the fitting degree of the model. 321 cases were used to evaluate the performance and generalization ability of the model. Thus, the transformer fault diagnosis is realized. The sample data distribution for each fault type is shown in Figure 5.

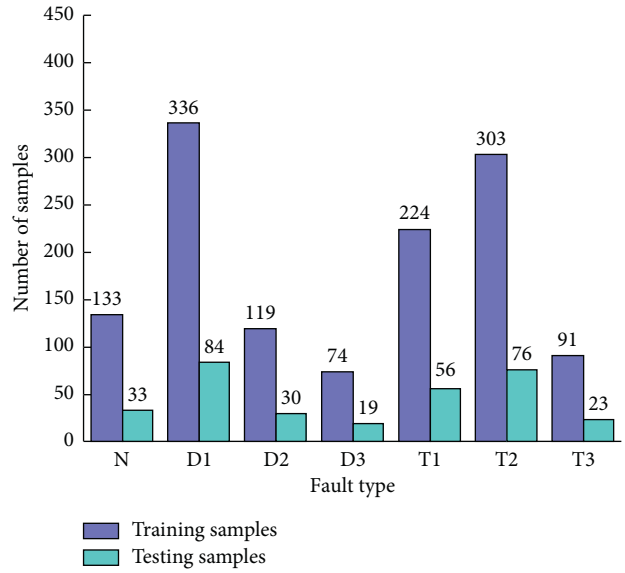

Figure 5: The distribution of transformer fault sample data.

4.1. IPSO-Gcforest Model Parameter Selection and Optimization Results. After normalizing the data in Figure 5, the noncode ratios of the characteristic gas dissolved in the oil are determined as the characteristic parameter of the model. In the process of q and l optimization of gcForest model parameters by IPSO, the diagnostic accuracy of training set is taken as the particle fitness value. After adjusting the model parameters and comparative analysis of the diagnosis results, the model parameters are determined as follows: the number of decision trees in a random forest during multigrained scanning is 500, and the decision tree growth rule is that the purity of the leaf node reaches the optimal or the depth reaches 50. The number of decision trees in a single random forest of cascade layer is 101, and the decision tree growth rule is that the purity of the leaf node reaches the optimal or the depth reaches 50. The parameters are set during the optimization process as shown in Table 1. The fitness change of the particles during the optimization process is shown in Figure 6.

As can be seen from Figure 6, the parameters q and l of the gcForest model go through five rounds of 100 iterations each. The accuracy of transformer fault diagnosis reaches the best in the 68, 49, 54, 65, and 52 iterations, respectively. At the same time, the IPSO algorithm optimization process is improved from 93.15% or 93.46% through 3 to 4 steps to the optimal fitness value of 94.70%. Finally, when q is 4 and l is 5, the particle fitness is the best, reaching 94.70%.

4.2. Comparison of Different Characteristic Parameters. According to the data distribution in Figure 5, the noncode ratios were used as input characteristic parameters to test the IPSO-gcForest model. In order to verify the effectiveness of the proposed method, the DGA data, IEC ratios, and Rogers ratios are used as input characteristic parameters in contrast with the results obtained from noncode ratios. In order to diagnose and analyze transformer fault types, the above four different types of characteristic parameters were, respectively, input into RF model, DBN model, gcForest model, PSO-gcForest model, and IPSO-gcForest model for diagnosis. The RF model adopts bootstrap resampling method. The number of subtrees is 100, and the number of split features is 7. The activation function of the DBN model uses the sigmoid function, and the learning rate is 0.001. The momentum is 0.9, and the number of hidden layers is 3. The default parameter setting of gcForest model is that the number of decision trees in a random forest during multigrained scanning is 500, and the window size q is 2. The number of decision trees in a single random forest in the cascade layer is 101, and the maximum number of allowed cascades l is 7. The results are shown in Table 2.

As can be seen from Table 2, the diagnostic accuracy of the same characteristic parameter increased in the order of RF model, DBN model, gcForest model, PSO-gcForest model, and IPSO-gcForest model. The diagnostic accuracy of the same method is improved according to the characteristic parameters of DGA data, IEC ratios, Rogers ratios, and noncode ratios. With noncode ratios as the characteristic parameter, IPSO-gcForest has the highest diagnostic accuracy, reaching 94.70%. Compared with RF model, DBN

Table 1: The parameters of IPSO-gcForest model.

Parameters	Value
T_{max}	100
Population number of particles	50
q value range	[1, 9]
q search range	[−1, 1]
l value range	[1, 20]
l search range	[−2, 2]
w_{ini}	0.9
w_{fin}	0.4
$s_{1,ini}$	2.5
$s_{1,fin}$	0.5
$s_{2,ini}$	2.5
$s_{2,fin}$	0.5

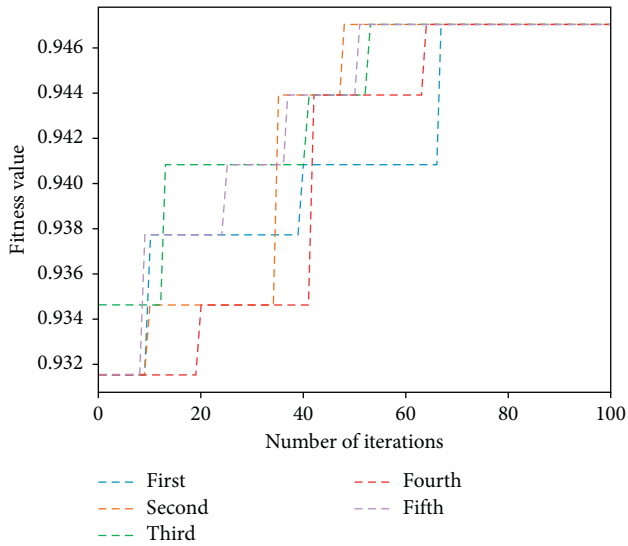

Figure 6: The change of particle fitness.

model, gcForest model, and PSO-gcForest model diagnostic results, the accuracy of IPSO-gcForest fault diagnosis is improved by 10.90%, 9.03%, 5.91%, and 1.87%, respectively. Compared with characteristic parameters of DGA data, IEC ratios, and Rogers ratios, the diagnostic accuracy of IPSO-gcForest was improved by 10.59%, 7.16%, and 3.11%, respectively. It shows that noncode ratios can provide more characteristic information as the input characteristic parameter of the transformer fault diagnosis model.

4.3. Comparison of Different Diagnostic Models. Due to the unbalanced distribution of the samples of each fault type in the collected transformer fault data, the performance of the model cannot be effectively verified only by the diagnostic accuracy. Therefore, the precision, recall rate, and receiver operating characteristic (ROC) curve are used to measure the generalization ability of the model. The noncode ratios are used as the input characteristic parameter of different diagnostic models, and the diagnostic result is shown in Table 3.

As can be seen from Table 3, the precision and recall rate of IPSO-gcForest method are all above 84%, and the average precision and average recall rate are 94.00% and 92.77%, respectively. The results show that IPSO-gcForest model has obvious advantages in the classification performance of each fault type. For the RF model to diagnose transformer fault types, the partial discharge fault diagnosis accuracy is the highest, reaching 88.24%. However, the recall rate of low energy discharge fault is the lowest, which is only 50.00%. The reason is that the fault types of transformers are related to each other, and different fault superpositions may occur. The recall rate of low energy discharge fault identified by IPSO-gcForest model is the highest, reaching 93.33%. The results show that it can effectively identify the actual fault types of transformer.

The ROC curve draws the trend chart by the real case rate (vertical axis) and false positive case rate (horizontal axis) under different discriminant probability thresholds. The ROC curve can comprehensively evaluate the classification performance of fault diagnosis methods, especially for unbalanced sample. By calculating the area under the ROC curve, it can measure the learning effect of the model better on a few cost-sensitive samples that need attention. Moreover, the classification performance and the overall trend of the ROC curve can be intuitively evaluated. The ROC curves of different diagnostic models are shown in Figure 7.

As can be seen from Figure 7, the area under the ROC curve of IPSO-gcForest diagnostic method is the highest, reaching 0.9873. Compared with the area under the ROC curve of other transformer fault diagnosis, it has increased by 13.67%, 11.71%, 6.77%, and 4.63% in turn. The results show that the proposed method has good classification ability for unbalanced sample.

4.4. Comparison of Samples of Different Sizes. In order to further analyze the robustness of IPSO-gcForest diagnostic models under different size samples, according to the proportions of 25%, 50%, 75%, and 100%, the fault samples in Figure 5 are divided into sample 1 (400 cases), sample 2 (800 cases), sample 3 (1201 cases), and sample 4 (1601 cases). Each sample is divided into training set and testing set according to the proportion of 8 : 2, and the diagnostic accuracy is shown in Figure 8.

As can be seen from Figure 8, the IPSO-gcForest model achieves high accuracy in fault diagnosis under different size samples. The results show that the performance of IPSO-gcForest model is better than that of the other three fault diagnosis methods. Compared with sample 1, sample 2, and sample 3, the diagnostic accuracy of IPSO-gcForest model in sample 4 increased by 9.51%, 5.88%, and 3.03%, respectively. It indicates that the larger the sample size, the more the feature information extracted. When the size of samples

Table 2: Diagnosis accuracy of different characterizing parameters in percentage.

Characteristic parameter	RF	DBN	gcForest	PSO-gcForest	IPSO-gcForest
DGA data	76.64	79.44	81.31	81.62	84.11
IEC ratios	78.82	80.37	82.55	84.42	87.54
Rogers ratios	80.37	81.31	83.49	89.09	91.59
Noncode ratios	83.80	85.67	88.79	92.83	94.70

Table 3: Comparison of diagnostic results of different models.

Evaluation index	Fault type	Diagnostic results of different models /%				
		RF	DBN	gcForest	PSO-gcForest	IPSO-gcForest
Precision	N	80.00	80.56	85.29	93.75	96.88
	D1	80.85	82.11	90.80	98.77	98.80
	D2	71.43	80.00	82.14	90.00	96.55
	D3	88.24	87.50	73.68	71.43	84.21
	T1	88.68	90.57	91.38	89.66	91.38
	T2	87.65	88.89	92.41	94.87	94.94
	T3	85.00	85.00	87.50	95.24	95.24
Recall rate	N	84.85	87.88	87.88	90.91	93.94
	D1	90.48	92.86	94.05	95.24	97.62
	D2	50.00	53.33	76.67	90.00	93.33
	D3	78.95	73.68	73.68	78.95	84.21
	T1	83.93	85.71	94.64	92.86	94.64
	T2	93.42	94.74	96.05	97.37	98.68
	T3	73.91	73.91	60.87	86.96	86.96

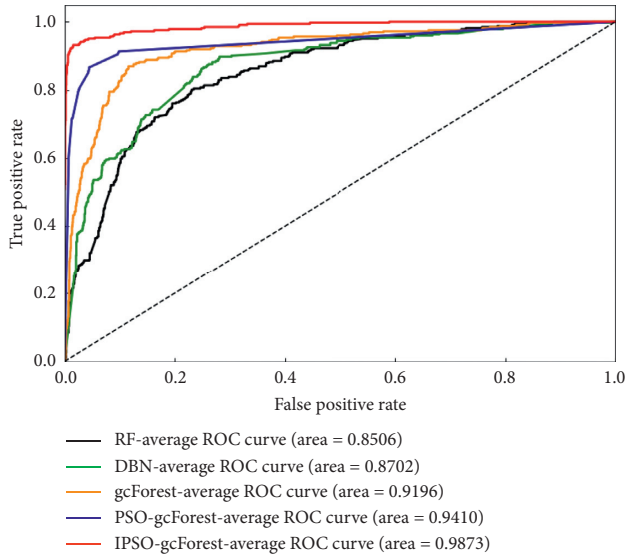

Figure 7: ROC curves of different diagnostic models.

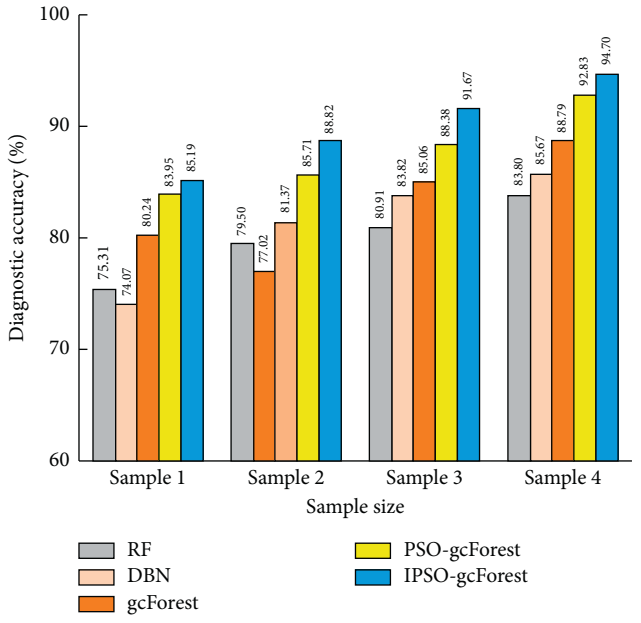

Figure 8: Diagnostic accuracy of different sample sizes.

decreases, the diagnostic accuracy of each method will decrease. However, the reduction of sample size has little effect on fault diagnosis accuracy of IPSO-gcForest model. This indicates that IPSO-gcForest model has better model performance and strong robustness under small size samples.

4.5. Case Study. Table 4 shows the oil chromatographic data of a transformer with SFSZ9-50000/110 in a substation after its failure on January 21, 2020.

Table 4: Transformer DGA data.

Gas type	H_2	CH_4	C_2H_2	C_2H_4	C_2H_6
Gas content (μL/L)	16.08	2.75	55.70	25.21	0.70

By selecting the noncode ratios as the input characteristic parameter of the IPSO-gcForest model, the oil chromatographic data is diagnosed and identified. The result of

Figure 9: Burning loss between turns of A phase low voltage coil.

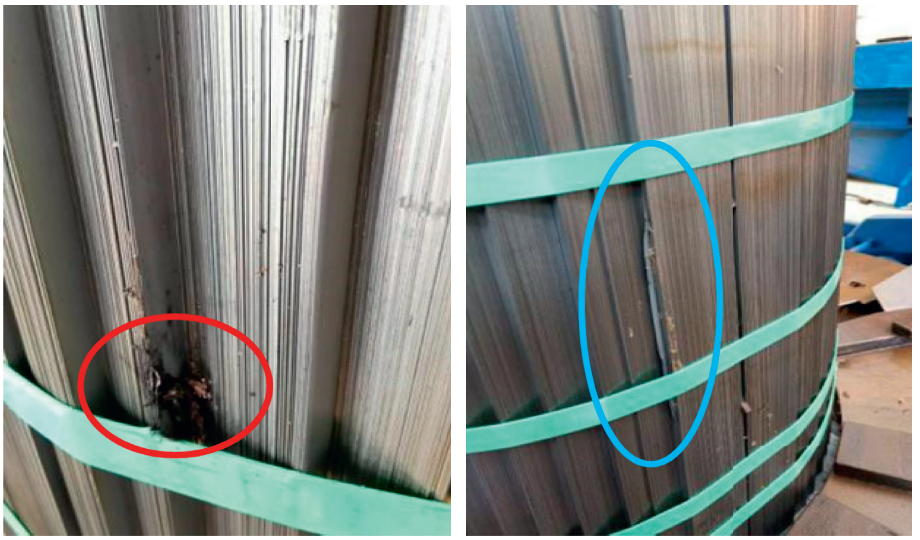

Figure 10: Discharge point and deformation diagram of A phase iron core column.

Figure 11: B phase low voltage coil bulge.

Figure 12: C phase low voltage coil and silicon steel sheet deformation.

the diagnosis is high-energy discharge with a probability of 87.63%. But the code determined by the three-ratio method is "202," and the corresponding fault type cannot be determined.

The maintenance personnel found that the A phase low voltage coil of the transformer was burned in a large area. From the bottom 38 to 54 and 68 to 71, severe short-circuit and interturn short-circuit occurred between the cakes. The windings were melted and twisted in many places, and the upper coil had radial deformation. The sinking of the whole coil reaches about 40 mm, and the cushion block has dislocated and fallen off. There are a lot of melted copper and carbonization marks of insulating material in the fault position, as shown in Figure 9. There are obvious discharge traces between the low voltage coil and the iron core, and the silicon steel sheet deformed slightly, as shown in Figure 10.

The B phase low voltage coil of the transformer is obviously bulged. The insulating paper is damaged, and the axial height of the winding sinks about 15 mm, as shown in Figure 11. There is deformation between the lower end of the transformer C phase low voltage coil and the core, and there is a slight loosening when pressing by hand. There are no obvious changes in the axial height of the winding, as shown in Figure 12.

From the analysis of the field situation, there is a high-energy discharge problem existing in the transformer A phase low voltage winding, which is consistent with the diagnosis result of the transformer fault diagnosis method proposed in this paper.

5. Conclusions

This paper combines the current artificial intelligence technology and machine learning algorithms; thus, a transformer fault diagnosis method based on IPSO-gcForest model is proposed. The following conclusions are obtained from the example analysis results: (1) by improving the location update strategy of the traditional PSO algorithm, the key parameters of the gcForest model are optimized by using the IPSO algorithm, which overcomes the random fluctuation of the output of the gcForest model and makes the diagnosis model have better generalization performance. (2) Compared with the RF, DBN, gcForest, and PSO-gcForest models, the IPSO-gcForest model has higher diagnostic accuracy in the diagnostic model with the noncode ratios, DGA data, IEC ratios, and Rogers ratios as characteristic parameters. Among them, the model with noncode ratios as characteristic parameter has higher diagnostic accuracy than the other three characteristic parameters. (3) The proposed IPSO-gcForest transformer fault method has higher identification accuracy and higher recall rate than other compared methods. Moreover, its AUC value is also the highest, which improves the classification ability of unbalanced sample data. (4) With the increasing sample size, the IPSO-gcForest model achieves improved diagnostic accuracy and more stable diagnostic performance. In the future, it will be possible to increase the collection of discharge and overheat mixed fault cases to verify the effectiveness of the proposed method. And further research on optimization model structure will be conducted.

Data Availability

The data were obtained from the transformer online monitoring data and historical fault data of China Southern Power Grid Corporation, the transformer fault oil chromatographic data in published papers, and the paper entitled the "Typical Cases of Application of Power Grid Equipment Detection Technology" published by the State Grid and IEC TC 10 database.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by National Natural Science Foundation of China (grant no. 51907084), Yunnan Provincial Talents Training Program (grant no. KKSY201704027), Science and Technology Fund of Yunnan Power Grid Corporation (grant no. YNKJXM20180736), and Scientific Research Foundation of Yunnan Provincial Department of Education (grant no. 2018JS032).

References

[1] H. Cong, S. Du, Q. Li, and J. Liu, "Electro-Thermal fault diagnosis method of RAPO vegetable oil transformer based on characteristic gas and ratio criterion," *IEEE Access*, vol. 7, pp. 101147–101159, 2019.

[2] B. Qi, P. Zhang, Z. Rong, J. Wang, C. Li, and J. Chen, "Rapid transformer health state recognition through canopy cluster-merging of dissolved gas data in high-dimensional space," *IEEE Access*, vol. 7, pp. 94520–94532, 2019.

[3] J. Fan, Z. Liu, A. Meng et al., "Characteristics of tin oxide chromatographic detector for dissolved gases analysis of transformer oil," *IEEE Access*, vol. 7, pp. 94012–94020, 2019.

[4] C. Xiang, Z. Huang, J. Li, Q. Zhou, and W. Yao, "Graphic approaches for faults diagnosis for Camellia insulating liquid filled transformers based on dissolved gas analysis," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 25, no. 5, pp. 1897–1903, 2018.

[5] J. Faiz and M. Soleimani, "Assessment of computational intelligence and conventional dissolved gas analysis methods for transformer fault diagnosis," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 25, no. 5, pp. 1798–1806, 2018.

[6] I. B. M. Taha, S. S. M. Ghoneim, and A. S. A. Duaywah, "Refining DGA methods of IEC code and rogers four ratios for transformer fault diagnosis," in *Proceedings of the 2016 IEEE Power and Energy Society General Meeting (PESGM)*, pp. 1–5, Boston, MA, USA, July 2016.

[7] J. Faiz and M. Soleimani, "Dissolved gas analysis evaluation in electric power transformers using conventional methods a review," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 24, no. 2, pp. 1239–1248, 2017.

[8] J. Dai, H. Song, G. Sheng, and X. Jiang, "Dissolved gas analysis of insulating oil for power transformer fault diagnosis with deep belief network," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 24, no. 5, pp. 2828–2835, 2017.

[9] X. Wang, Z. D. Wang, Q. Liu et al., "Dissolved gas analysis (DGA) of mineral oil under thermal faults with tube heating method[C]," in *Proceedings of the international conference on dielectric liquids*, pp. 1–4, Manchester, UK, June 2017.

[10] M. Žarkovic and Z. Stojkovic, "Analysis of artificial intelligence expert systems for power transformer condition monitoring and diagnostics," *Electric Power Systems Research*, vol. 149, pp. 125–136, 2017.

[11] M. Ou, H. Wei, Y. Zhang et al., "A dynamic adam based deep neural network for fault diagnosis of oil-immersed power transformers," *Energies*, vol. 12, no. 6, 2019.

[12] T. Wang, Y. He, T. Shi, and B. Li, "Transformer incipient hybrid fault diagnosis based on solar-powered RFID sensor and optimized DBN approach," *IEEE Access*, vol. 7, pp. 74103–74110, 2019.

[13] Y. Liang, Y. Xu, X. Wan et al., "Dissolved gas analysis of transformer oil based on Deep Belief Networks," in *Proceedings of the International Conference on Properties and Applications of Dielectric Materials*, Xi'an, China, May 2018.

[14] C. Dai, Z. Liu, K. Hu, and K. Huang, "Fault diagnosis approach of traction transformers in high-speed railway combining kernel principal component analysis with random forest," *IET Electrical Systems in Transportation*, vol. 6, no. 3, pp. 202–206, 2016.

[15] Y. Zhang, X. Li, H. Zheng et al., "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.

[16] 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.

[17] G. Zhu, Q. Hu, R. Gu, C. Yuan, and Y. Huang, "ForestLayer: efficient training of deep forests on distributed task-parallel platforms," *Journal of Parallel and Distributed Computing*, vol. 132, pp. 113–126, 2019.

[18] Y. Zhang, J. Zhou, W. Zheng et al., "Distributed deep forest and its application to automatic detection of cash-Out fraud," 2018, https://arxiv.org/abs/1805.04234.

[19] X. Liu, R. Wang, Z. Cai, Y. Cai, and X. Yin, "Deep multigrained cascade forest for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 10, pp. 8169–8183, 2019.

[20] Z. Lu, M. Wang, W. Dai, and J. Sun, "In-process complex machining condition monitoring based on deep forest and process information fusion," *The International Journal of Advanced Manufacturing Technology*, vol. 104, no. 5-8, pp. 1953–1966, 2019.

[21] X. Liu, Y. Tian, X. Lei et al., "Deep forest based intelligent fault diagnosis of hydraulic turbine," *Journal of Mechanical Science and Technology*, vol. 33, no. 5, pp. 2049–2058, 2019.

[22] S. Tang, M. Peng, G. Xia et al., "Optimization design for supercritical carbon dioxide compressor based on simulated annealing algorithm," *Annals of Nuclear Energy*, vol. 140, Article ID 107107, 2020.

[23] E. Abiri, A. Darabi, M. R. Salehi et al., "Optimized gate diffusion input method-based reversible magnitude arithmetic unit using non-dominated sorting genetic algorithm II," *Circuits Systems and Signal Processing*, vol. 39, no. 9, pp. 4516–4551, 2020.

[24] J. Zhu, L. Han, X. Meng et al., "An AMP-based low complexity generalized sparse bayesian learning algorithm," *IEEE Access*, vol. 7, pp. 7965–7976, 2019.

[25] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in *Proceedings of the international conference on networks*, pp. 1942–1948, Paris, France, November 2002.

[26] Q. Feng, "Improved particle swarm optimizer based on nonlinear inertia weight dynamic changing," *Computer Science*, vol. 35, pp. 146–148, 2008.

[27] J. Baizhuang, "Improved PSO algorithm based on cosine functions and its simulation," *Journal of Computer Applications*, 2013.

[28] F. N. Budiman and E. S. Wahyuni, "Discrimination of particle-initiated defects in gas-insulated system using C4.5 algorithm," in *Proceedings of the International Conference on Information Technology Computer and Electrical Engineering*, pp. 191–196, Semarang, Indonesia, October 2016.

[29] S. Yang, J. Guo, J. Jin et al., "An improved Id3 algorithm for medical data classification," *Computers & Electrical Engineering*, pp. 474–487, 2017.