IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 22, 2019, accepted June 22, 2019, date of publication July 5, 2019, date of current version August 12, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2927018

A Fault Diagnosis Model of Power Transformers Based on Dissolved Gas Analysis Features Selection and Improved Krill Herd Algorithm Optimized Support Vector Machine

YIYI ZHANG[1], XIN LI[1], HANBO ZHENG[1], HUILU YAO[2], JIEFENG LIU[1], CHAOHAI ZHANG[1], HONGBO PENG[1], AND JIAN JIAO[3]

[1]Guangxi Key Laboratory of Power System Optimization and Energy Technology, Guangxi University, Nanning 530004, China
[2]School of Physical Science and Technology, Guangxi University, Nanning 530004, China
[3]Electric Power Research Institute, Guangxi Power Grid Company Ltd., Nanning 530023, China

Corresponding authors: Hanbo Zheng (hanbozheng@163.com) and Huilu Yao (huiluy@qq.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61473272 and Grant 51867003, in part by the Basic Ability Promotion Project for Yong Teachers in Universities of Guangxi under Grant 2019KY0046 and Grant 2019KY0022, in part by the Natural Science Foundation of Guangxi under Grant 2018JJB160056, Grant 2018JJB160064, and Grant 2018JJA160176, in part by the Guangxi Thousand Backbone Teachers Training Program, in part by the Boshike Award Scheme for Young Innovative Talents, and in part by the Guangxi Bagui Young Scholars Special Funding.

ABSTRACT In this paper, a set of dissolved gas analysis (DGA) new feature combinations is selected as input from the mixed DGA feature quantity, and an improved krill herd (IKH) algorithm optimized support vector machine (SVM) transformer fault diagnosis model is established to solve the problem that the single characteristic gas or characteristic gas ratio, which are utilized as the DGA feature quantity cannot fully reflect the transformer fault classification. The following work has been done in this paper: 1) IEC TC 10 fault data and other 117 sets of fault data in China are preprocessed in order to reduce the influence on the diagnosis results causing by the edge data in the fuzzy area; 2) the SVM parameters and 11 features are encoded by a binary code technique; 3) a preferred DGA feature set for fault diagnosis of power transformers is selected by genetic algorithm (GA) and SVM, and; 4) IKH is utilized to optimize the parameters of SVM. Combining with cross-validation principle, a transformer fault diagnosis model based on IKH algorithm to optimize SVM is established. The fault diagnosis results based on the new fault sample show that the proposed DGA feature set to increase the accuracy by 26.78% and 10.83% over the DGA full data and IEC ratios. Moreover, the accuracy of IKHSVM is better than the GASVM, back-propagation neural network (BPNN), and particle swarm optimization optimized support vector machine (PSOSVM), the accuracy rates are 85.71%, 75%, 64.29%, and 71.43%, which proves the validity of the proposed fault diagnosis model.

INDEX TERMS Power transformers, fault diagnosis, support vector machine, improved krill herd algorithm, DGA feature.

I. INTRODUCTION

The oil-immersed transformer is a vital component of the power grid, so it is of great significance to timely detect the potential troubles to improve the stability and security of the power system. The highly reliable data lead to the extensive use of the Dissolved Gas Analysis (DGA) technology in the field of transformer fault diagnosis. The content of dissolved gases in the oil has a significant correspondence with the fault type and the fault severity of transformer. In this case, the dissolved gases include CH_4, C_2H_6, C_2H_4, C_2H_2, CO, CO_2 and H_2 are commonly selected as feature gases [1].

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

According to the research [2], the type of transformer fault is relative to the content ratio of feature gases as well. Some experts proposed that Rogers Ratio [3], IEC Ratio [4] and Doermenbur Ratio [5] could be used as the feature quantity to reflect transformer anomalies. However, Rogers Ratio can just reflect the thermal decomposition temperature range only [6], and IEC Ratio has problems such as incomplete coding and excessive coding boundaries that will cause the judgment errors. Thus, a single gas or the content ratio of feature gases cannot show the relationship between the transformer fault and gases completely, and the limitation may affect the accuracy of the diagnosis result [7]–[9].

Aiming at making up for the shortcomings of the existing DGA feature selection, in this paper, a set of optimal DGA feature combinations from the set of characteristic gas and characteristic gas ratios will be screened out. There have been several methods for feature selection of data, such as genetic algorithm [10], binary particle swarm algorithm [11], neural network [12], imperialist competitive algorithm [13], and tabu search [14]. Binary particle swarm algorithm is a population-based optimizer similar to GA. BPSO algorithm has the strong global search ability, but it cannot converge to the global optimal position of particles [15]. Moreover, with the iteration of the algorithm, the randomness of BPSO becomes stronger, but it lacks the local search ability in the later period [16], [17]. In the iterative process of the imperialist competitive algorithm, the number of empire is continuously reduced, resulting in a decrease in population diversity, which is unfavorable for solving high-dimensional multi-mode optimization problems, and the algorithm is easy to fall into the local optimal solution [18]. GA can handle large-scale complex data well due to its adaptability, and is especially suitable for solving multi-objective optimization problems [18], [19]. Therefore, genetic algorithm is used for feature selection in this paper.

In recent years, the diagnosis method using the dissolved gas analysis technology in oil and the fuzzy theory [20], artificial neural network [21] and support vector machine [22] has significantly improved the accuracy of fault diagnosis. The fuzzy theory is simple in structure and fast in diagnosis, but its learning ability is insufficient [23]; the artificial neural network has strong self-learning and parallel processing ability, but it is easy to fall into local optimum [24], [25]. Compared with other methods, SVM can solve ''dimensionality disaster'', ''over-fitting'' and local minimum point problems [26], [27], but the performance of SVM mainly depends on its kernel function and its parameters [28], [29], so it needs to adopt an intelligent optimization algorithm to find the appropriate parameters to improve the generalization and robustness of the SVM [30]. Krill Herd (KH) [31], [32] idealizes the communication and foraging behavior of krill group. KH algorithm has excellent local and global optimization performance and has only one variable parameter—so it is efficient and straightforward. Further, KH algorithm is helpful in getting better convergence speed [33]–[35]. However, the standard KH algorithm cannot jump out of the local optimum in the later stage, which leads to poor solution accuracy. Therefore, IKH algorithm is utilized to optimize the parameters of SVM in this paper.

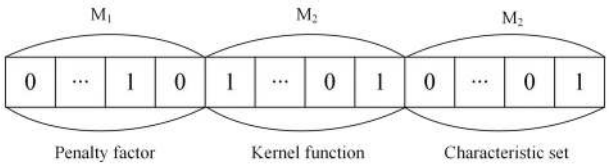

FIGURE 1. The binary encoding of chromosomes.

TABLE 1. Dissolved gas in oil.

Number	DGA Feature	Number	DGA Feature
1	H_2	7	CO_2
2	CH_4	8	TH
3	C_2H_2	9	CH_4/H_2
4	C_2H_4	10	C_2H_4/C_2H_6
5	C_2H_6	11	C_2H_2/C_2H_4
6	CO		

In this paper, the genetic algorithm is used to binary code the SVM parameters and DGA feature quantities. The GA combined with SVM is used to optimize the input features of the fault diagnosis model. The SVM parameters can be optimized by IKH algorithm and it is possible to build the transformer fault diagnosis model based on IKH algorithm optimized SVM by the cross-validation principle. The fault data of 150 sets of IEC TC 10 and of 117 sets in China are screened. The validity and superiority of the proposed method are verified by the diagnosis results based on the processed 113 sets of fault samples of IEC TC10, and using the fault data in China to test the accuracy of the method again.

II. OPTIMIZATONG OF TRANSFORMER FAULT FEATURE RATIO BASED ON GA AND SUPPORT VECTOR MACHINE

A. GAS FEATURE DISSOLVED IN OIL

Considering the advantages of characteristic gas and DGA ratio, CH_4, C_2H_6, C_2H_4, C_2H_2, CO, CO_2, H_2, total hydrocarbon (TH) and the gas content ratios of three gas groups (CH_4/H_2, C_2H_4/C_2H_6, and C_2H_2/C_2H_4) are treated as the gas feature dissolved in oil. The specific numbers are shown in Table 1.

In order to reduce the impact of data error on diagnostic accuracy, 150 sets of IEC TC 10 fault data and 117 sets of fault data in China were screened. According to the grid equipment condition maintenance rules and field experience, the characteristic gases produced by overheating faults are mainly CH_4 and C_2H_6, and the sum of the two generally accounts for more than 80% of the total hydrocarbon [36], [37]. As the temperature at the fault point increases, the proportion of C_2H_6 will increase. Usually, C_2H_2 is not produced in the event of overheating [38]. In the case of overheating faults generally below 500°, the C_2H_2 content will not exceed 2% of the total hydrocarbon, while in severe overheating, the maximum C_2H_2 content will not exceed 6%. When the

TABLE 2. Attention value of dissolved gas in oil.

Characteristic gas	Specification	Attention value (μL/L)
C_2H_2	330kV or more	≤1
	Other	≤5
H_2	-	≤150
TH	-	≤150

overheating fault involves solid insulating materials, in addition to producing the above gases, a large amount of CO and CO_2 are produces as well [39]. High-energy discharge faults produce gas rapidly and large amounts of gas are produced. The fault characteristic gases are mainly C_2H_2 and H_2, followed by a large number of C_2H_6 and CH_4. C_2H_2 generally accounts for 20-70% of total hydrocarbons, H_2 accounts for 30-90%, and in most cases, C_2H_6 content is higher than CH_4. Low-energy discharge faults generally have low total hydrocarbon content because of the low fault energy. And the main component is H_2, followed by CH_4. When the discharge energy density is increased, C_2H_2 may be produced as well, but the proportion of C_2H_2 in the total hydrocarbons is ordinarily less than 2%. This is also the major indicator of the difference between the two types of discharge faults [40]. Theoretically, the gas content should be less than the attention value under normal conditions [41]. According to the regulations, the attention values of the characteristic gas components are shown in Table 2.

However, among the original fault data, some data fall into a fuzzy area, and the characteristic gas concentration satisfies a plurality of fault conditions at the same time, which may cause the classification boundary to be unclear, thereby leading to the lower accuracy of fault diagnosis. Therefore, 37 sets of IEC TC 10 and 12 sets of the data from China in the fuzzy area are deleted. The remaining fault samples will be used for the diagnosis of the model proposed in this paper.

B. SVM MODEL FOR TRANSFORMER FAULT DIAGNOSIS

The standard SVM is a typical two-class classifier, and the transformer fault diagnosis is a linear and inseparable multi-classification problem, so nonlinear and multi-class transformation of SVM is needed [15].

The SVM nonlinear model is shown in equation (1).

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

where ξ_i is a relaxation variable and parameter C is a penalty factor.

Its Lagrangian function is shown as follows:

$$L(\omega, \lambda, \xi, \alpha, \beta) = \Phi(\omega, \xi) - \sum_{i=1}^{l}\alpha_i\left\{y_i[\omega^T\varphi(x_i) + \lambda] - 1 + \xi_i\right\} - \sum_{i=1}^{l}\beta_i\xi_i \tag{2}$$

The decision function is:

$$f(x) = sign[\sum_{i=1}^{l}\alpha_i y_i K(x, x_i) + \lambda] \tag{3}$$

Commonly used kernel functions include radial basis functions (RBF), polynomial functions, etc. [16]. The RBF function only needs to determine one parameter, so it is beneficial to the optimization of the parameters. Therefore, RBF is used as the kernel function of the SVM:

$$K(x_i, x_j) = \exp(-\gamma\left\|x_i - x_j\right\|^2), \quad \gamma > 0 \tag{4}$$

Typical combinations are one against all (OAA), one against one (OAO), minimum output coding (MOC), etc. [17]. OAO has the best effect in transformer fault diagnosis, so this paper uses OAO method to expand the two-class SVM into multi-class SVM.

C. OPTIMIZATION OF TRANSFORMER FAULT FEATURE BASED ON GA

1) CHROMOSOME CODING

Due to the encoding and decoding operations of binary coding are simple, the genetic operations such as crossover and mutation are easy to implement, the chromosome of GA is generated by binary coding. Each chromosome consists of SVM penalty factor c, kernel parameter and DGA features. Each chromosome is made up of three genes. The first two genes are composed of binary codes with length of 10, representing SVM penalty factor c and nuclear parameters σ respectively. The third gene represents the selection of DGA feature quantities, with a length of 11, representing 11 kinds of DGA feature quantities. Among them, the coding of DGA feature quantities corresponds to the ordinal order of the feature quantities in Table 1, 1 indicates that the corresponding DGA feature quantities have been selected, and 0 indicates that the corresponding DGA feature quantities have not been selected.

2) FITNESS CALCULATION

After encoding the chromosome, the k-fold cross-classification accuracy of the transformer failure training sample is taken as the individual fitness f:

$$f(M_1, M_2, M_3) = \frac{1}{k}\sum_{i=1}^{k}\frac{m_T^i}{m^i} \times 100\% \tag{5}$$

where M_1, M_2 and M_3 represent the combined coding of c, σ and DGA features of support vector machine, respectively. m_i represents the number of samples in the i-th verification set when using the SVM algorithm; $mi\ T$ denotes the number of correct classifications in the verification set when using SVM algorithm; k represents the number of folds for cross-validation and is set to 5.

In order to obtain the f of the chromosome, the M_1, M_2 and M_3 segments need to be decoded in segments. The decimal numbers obtained after decoding the M_1 segment and the M_2 segment are c and of the SVM; the selected gas ratios

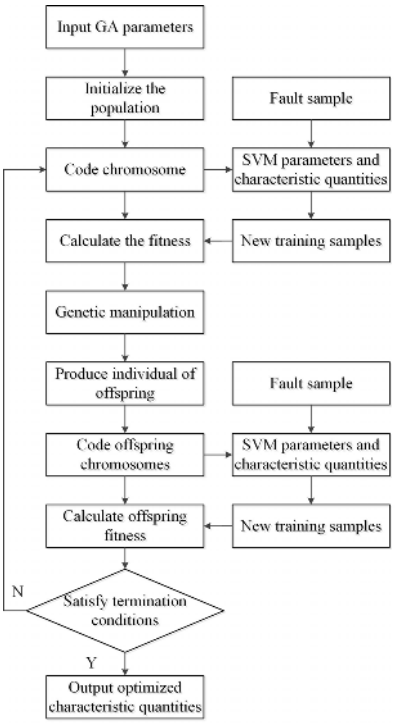

FIGURE 2. Flowchart of DGA feature selection.

according to the coding value of each bit in M_3 segment 0 or 1, form a new training sample.

3) GENETIC MANIPULATION

Genetic operations include selection operations, cross operations, and mutation operations.

- The selection operation can increase the average fitness value of the group. In this paper, the ''steady-state selection'' is utilized to preserve individuals with higher fitness in the father.
- Cross operations are used to generate new individuals. The ''single point crossover'' is utilized in this paper.
- The mutation operation is used to assist the generation of new individuals, which determines the local search ability of the genetic algorithm. The ''basic bit variation'' is utilized in this paper.

4) ALGORITHM FLOW

The preferred flowchart of DGA feature based on GA and SVM is shown in Figure 2.

III. TRANSFORMER FAULT DIAGNOSIS MODEL BASED ON IMPROVED KRILL HERD (IKH) ALGORITHM

In the KH algorithm, each krill individual represents a potential solution in the solution space, and the food is the optimal global solution for the KH algorithm [19].

As what is shown in equation (11), the location update of krill individuals is determined by the combination of induced exercise, foraging movement, and random diffusion.

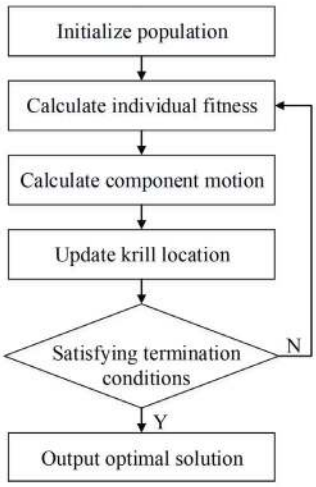

FIGURE 3. Flowchart of the improved krill herd algorithm.

Active and reactive currents I_{ai} and I_{ri} of Z can be calculated as:

$$Z_k = R_k + S_k + T_k \tag{6}$$

where Z_k represents the total movement of the krill; R_k represents the guided movement; S_k represents the foraging movement (i.e., the movement of each guided by the food); T_k represents the physical random spread of each.

A. INDUCED MOVEMENT

Induced motion is divided into target guidance, local influence, and guiding inertia:

$$R_k = R^{\max}\alpha_k + \omega_r R_k^{old} \tag{7}$$

where $R^{\max}$ indicates the maximum induction velocity; ω_r indicates the induction weight; α_k indicates the direction of induction.

B. FORAGING MOVEMENT

The foraging movement is divided into foraging experience and food guidelines. The speed of foraging is defined as follows:

$$S_k = V_s\beta_k + \omega_s R_k^{old} \tag{8}$$

where V_s is the maximum speed of foraging; δ is the weight of foraging; β_k is the direction of foraging.

C. RANDOM DIFFUSION

The random diffusion speed is defined as follows:

$$T_K = T^{\max}(1 - t/t_{\max})\delta \tag{9}$$

where T^{max} represents the maximum random diffusion speed, and δ represents the direction of random diffusion. The speed of krill individual is determined by the speed of induced movement, the speed of foraging movement and the speed of random diffusion.

$$\frac{dx_k}{dt} = R_k + S_k + T_k \tag{10}$$

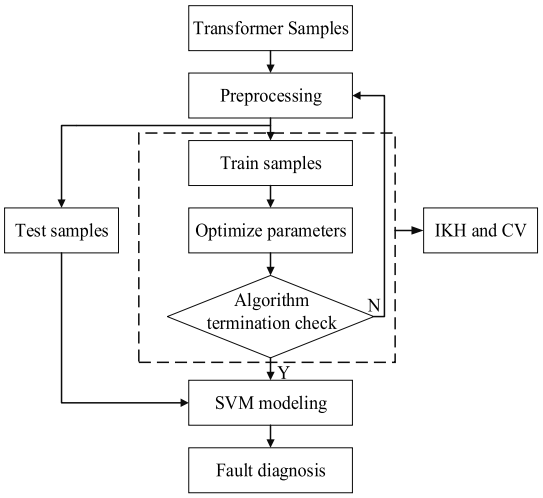

FIGURE 4. Flowchart of IKHSVM transformer fault diagnosis model.

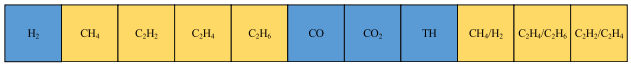

FIGURE 5. The new DGA feature set.

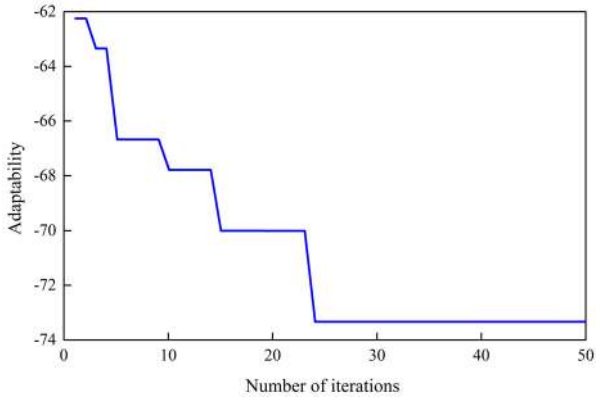

FIGURE 6. Evolution process of IKH.

The location update of Krill individual during the Δt period is expressed as follows:

$$x_k(t + \Delta t) = x_k(t) + \Delta t \frac{dx_k}{dt} \tag{11}$$

$$\Delta t = C_t \sum_{i=1}^{NV} (U_i - L_i) \tag{12}$$

where C_t represents the step size scaling factor; NV represents the variable dimension; U_i represents the upper bound of the variable; L_i represents the lower bound of the variable.

When using the standard KH algorithm to solve the global optimization problem, as the number of iterations increases, most of the krill moves to the same direction, which tends to make the problem solution fall into local optimum. The perturbation operator β is introduced in [19] to avoid KH algorithm falling into local optimum, so as to improve the performance of krill swarm optimization algorithm.

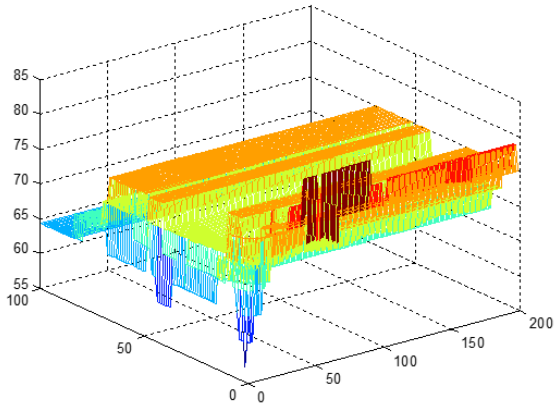

FIGURE 7. Testing accuracies for sample points.

TABLE 3. Transformer fault samples classification.

Fault type	LE-D	HE-D	LM-T	H-T	N-C
Sample quantities	22	43	9	14	25

Perturbation operator β is determined by equation (13)

$$\beta = \eta \times fitness \tag{13}$$

In the formula, $\eta \in [0, 1]$, gradually decreases from 1 to 0 as the number of iterations increases. The *fitness* represents the fitness of individual krill, and the greater the fitness is, the worse the solution to the problem represented by the individual position will be.

Redefine random diffusion based on the perturbation operator β:

$$T_K = T^{\max}(1 - t/t_{\max} + \beta)\delta \tag{14}$$

In the early stage of the algorithm iteration, since the β value of the disturbance operator is relatively large, the disturbance is large; The perturbation operator β gradually decreases with the increase of the number of iterations, thus reducing the random diffusion range of the krill population, which is beneficial to the krill's more detailed search in its own neighborhood, so the algorithm has stronger local optimization ability. In the later stage of the KH algorithm iteration, the krill individuals move closer to the global optimal individuals and are prone to fall into local optimum values. After introducing the perturbation operator β, the larger the fitness value, the larger the perturbation of krill individuals, which enables the krill to carry out a larger range of random diffusion motion, thus expanding the optimization range of the algorithm and avoiding the KH algorithm falling into local optimum. The flow chart of the improved KH is shown in Figure 3.

The parameters of SVM are optimized by IKH algorithm to construct a transformer diagnostic model based on IKH optimized SVM. The flowchart of IKHSVM fault diagnosis model is shown in Figure 4.

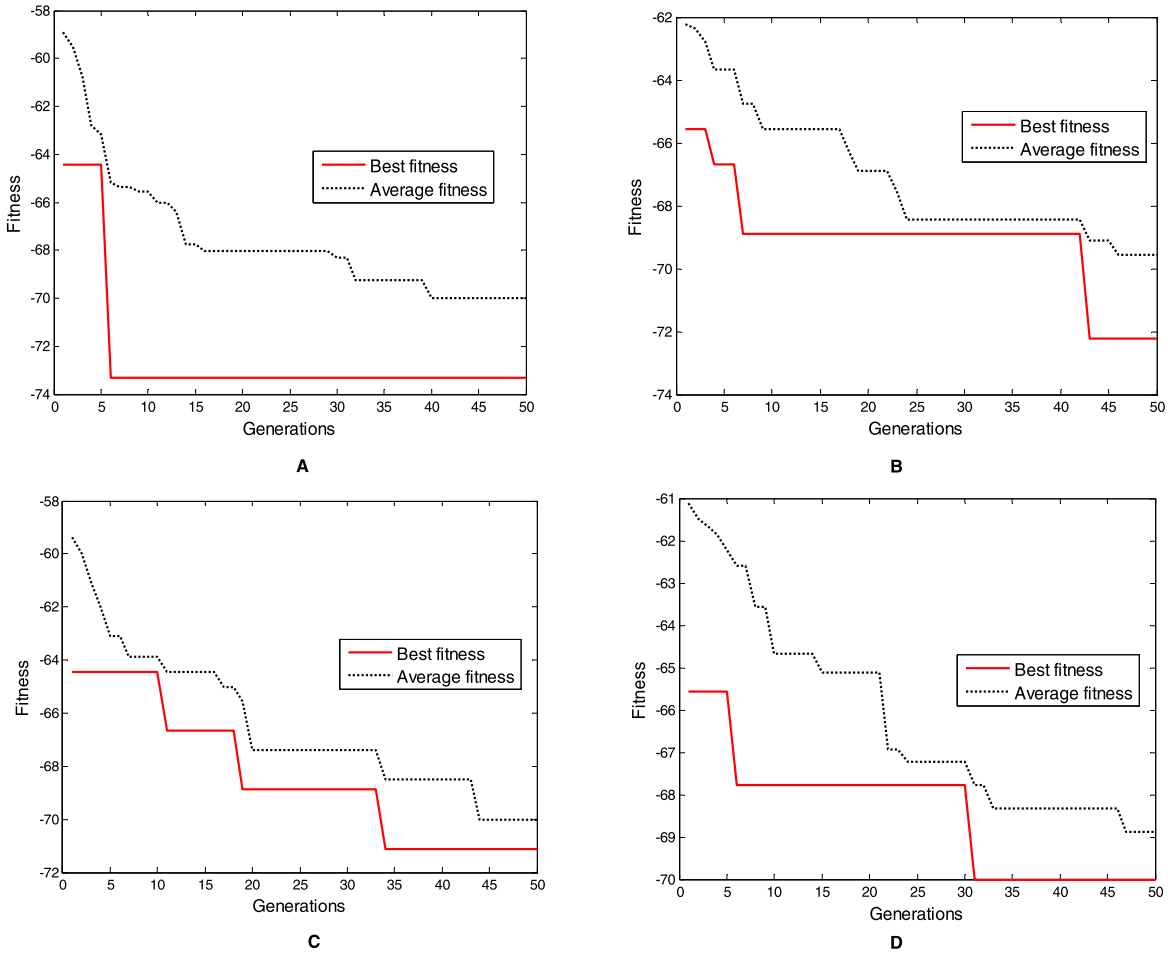

FIGURE 8. Diagnostic iterative diagram.

The flow of the transformer fault diagnosis model is as follows:

- Normalize preprocessing of DGA ratios to eliminate the effect of numerical values between different ratios.
- Establish IKHSVM fault diagnosis model by combining the IKH algorithm and Cross Validation (CV) principle.
- Diagnose the test sample and training sample by using the IKHSVM fault diagnosis model and the result can be obtained.

IV. RESULT ANALYSIS

A. FAULT SAMPLE SELECTION DATA SOURCE AND PARAMETER SETTING

The 113 groups of IEC TC 10 fault data have been divided into two parts, of which 85 samples are training samples, and the other 28 samples are test samples. Table 3 shows the classification of 113 fault samples.

When GA is used to optimize both SVM parameters and DGA feature quantities, GA has a maximum number of iterations of 50 and an initially generated particle swarm of 10. The SVM parameters are set as follows: the range of the penalty factor c and the kernel parameter σ are set to [0, 200] and [0, 100] respectively, the number of DGA feature quantities is 11.

TABLE 4. Three Sets of feature data.

DGA feature quantity coding	CV accuracy	Feature quantities
0 0 1 0 1 0 0 1 1 1 1	88.67%	6
0 0 1 1 0 1 0 0 0 1 1	88.67%	5
0 1 1 1 0 0 0 0 1 1 1	88.67%	6

In the IKHSVM fault diagnosis model, the krill population size is set to 100, the maximum induction speed $\boldsymbol{R}^{\max} = 0.02$, the maximum foraging speed $V_S = 0.01$, the maximum random diffusion speed $= 0.01$, the maximum number of iterations $K = \boldsymbol{T}^{\max}$ 100, and the number of CV is 9. The SVM parameters are set as follows: the range of the penalty factor c and the kernel parameter σ are set to [0, 200] and [0, 100] respectively.

B. DGA FEATURE QUANTITY OPTIMIZATION RESULT ANALYSIS

Three sets of DGA feature combinations are selected from 30 repeated calculations. The cross-validation accuracy and test sample accuracy can be seen in Table 4.

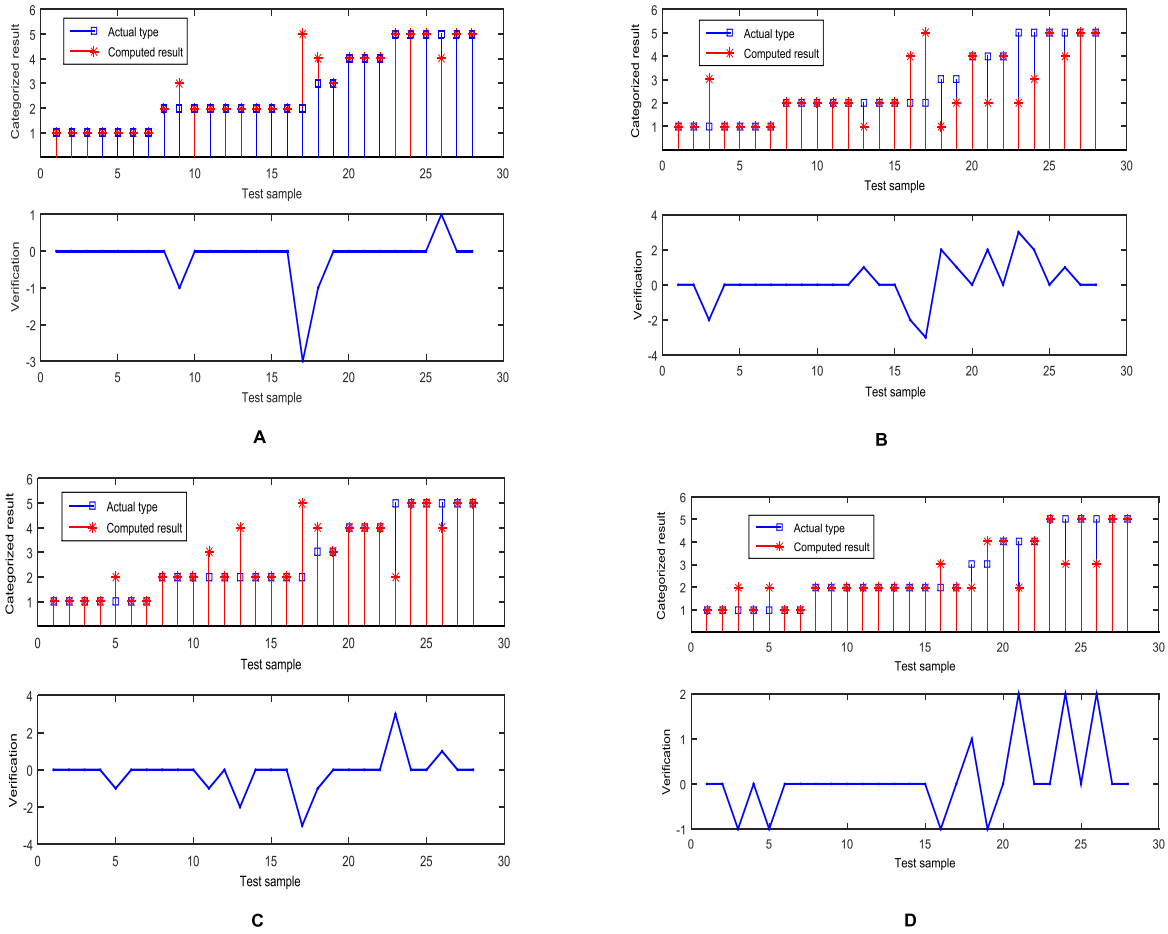

FIGURE 9. Fault diagnosis results.

The SVM is used to diagnose the test sample - the set of features with the highest accuracy is selected from the three sets of feature combinations and defined as the new DGA feature combination, as shown in Figure 5. The yellow regions are the selected features.

In order to compare the impact of different feature quantities on the fault diagnosis rate, the input features of IKHSVM are divided into three categories: (1) DGA full data, including H_2, CH_4, C_2H_2, C_2H_4, C_2H_6, CO, CO_2 and total Hydrocarbon; (2) The three-ratio feature quantity consists of three gas ratios of CH_4/H_2, C_2H_4/C_2H_6, and C_2H_2/C_2H_4; (3) The preferred DGA feature combinations selected in this paper. Using the feature sets from three categories respectively as the input of SVM and using the five fault types as outputs. The IKHSVM fitness curve when the optimal DGA feature combination is input can be seen in Figure 6.

Table 4 gives the statistical results of the average accuracy of the IKHSVM testing samples based on the three feature sets. It can be seen from Table 5 that the fault diagnosis accuracy rates of the DGA full data and the three-ratio feature quantity test samples are 58.93% and 74.27% respectively, and the accuracy of the three-ratio feature quantity is better than the DGA full data. The diagnostic accuracy rate of the new DGA feature set is 26.78% higher than the DGA full data, and 10.83% higher than the three-ratio feature quantity, indicating that the new DGA feature set can significantly improve the accuracy of transformer fault diagnosis. At the same time, there is not much difference between the accuracy of the new DGA feature set, the test samples and training samples of the three-ratio feature quantity. It indicates that the new DGA feature set is stable for fault diagnosis and can reduce the interference caused by different data.

TABLE 5. Average accuracy with different feature quantities.

Feature	Accuracy（%）
	Testing
DGA full data	58.93
Three-ratio feature	74.27
New DGA feature set	85.71

C. COMPARISON OF DIAGNOSTIC METHOD

In order to show more clearly the relationship between the test sample and the accuracy of penalty factor c and nuclear parameter σ, the search range of c and σ is set to

TABLE 6. Average accuracy of new DGA feature.

Diagnosis Method	Accuracy（%）
	Testing
BPNN	64.29
GASVM	75
PSOSVM	71.43
IKHSVM	85.71

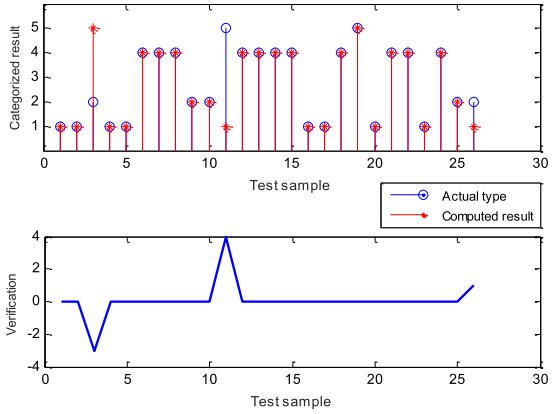

FIGURE 10. Fault diagnosis results.

500 and 250 equal parts respectively, totaling 125,000 sampling points (parameter combination), and each sampling point is used as the input of SVM to obtain the diagnosis results. Figure 7 shows the diagnostic results at each sampling point. In Figure 7, the deeper the yellow color is, the better the optimization effect of SVM parameters and the higher the accuracy of fault diagnosis will be; the deeper the blue color is, the lower the accuracy of diagnosis and the lower the accuracy of fault diagnosis will be.

Based on the same combination of new DGA features, transformer fault diagnosis is carried out using IKHSVM, GASVM, BPNN and PSOSVM. The ABCD in Figure 8 show the diagnostic iteration diagrams of the above four methods respectively. It can be seen that the best fitness of IKHSVM is better than the other three methods.

The ABCD charts of Figure 9 show the classification accuracy of IKHSVM, GASVM, BPNN and PSOSVM. The accuracy of the four methods is 85.71% (24/28), 75% (21/28), 64.29% (18/28) and 71.43% (20/28), respectively.

Table 6 shows the results of transformer fault diagnosis. It can be seen from Table 6 that the average testing accuracy of IKHSVM reaches 85.71% when the new feature set is input, which is higher than that of BPNN, GASVM and PSOSVM as 64.29%, 75% and 71.43% respectively. Compared with GASVM and PSOSVM, the average test accuracy of IKHSVM is higher which reflects that IKH algorithm has good local and global optimization performance. It effectively improves the accuracy of transformer fault diagnosis.

The remaining 105 groups of China fault data have been divided into two parts, of which 79 samples are training samples, and the other 26 samples are test samples. Figure 10 shows the accuracy of the IKHSVM model based the above fault data is 88.46% (23/26). It can be seen the result is similar to that obtained by IEC TC 10 and it verifies the reliability and validity of the IKHSVM model again.

V. CONCLUSION

In this paper, GA and SVM are used to optimize a set of new DGA features to establish an IKHSVM transformer fault diagnosis model for fault diagnosis. Fault data is preprocessed to reduce the influence of fuzzy boundary on diagnosis results. The conclusions are as follows:

- Using GA algorithm and SVM, a new DGA feature set of transformer fault diagnosis is selected from 11 kinds of DGA feature quantities. The average test accuracy of the new DGA feature set is 26.78% and 10.83% higher than that of the DGA full data and the three-ratio feature quantity respectively. The new feature set proposed in this paper can accurately reflect the faults of the transformer, which is better than other feature quantities.
- By using the new DGA set as input, the accuracy of the IKHSVM model of transformer fault diagnosis reaches 85.71% which is higher than the standard BPNN, GASVM model and PSOSVM model. In this case, IKH algorithm is more suitable for parameter optimization of support vector machines. It proves the effectiveness of the proposed method.
- In the study, some shortcomings still existence, such as the insufficient data samples. In the next research, we should not only consider the existing small sample data, but should consider more about how to be versatile and scalable in the context of ubiquitous power Internet of things and big data.

REFERENCES

[1] M. Dong, H. Zheng, Y. Zhang, K. Shi, S. Yao, X. Kou, G. Ding, and L. Guo, "A novel maintenance decision making model of power transformers based on reliability and economy assessment," *IEEE Access*, vol. 7, pp. 28778–28790, 2019.
[2] J. Liu, H. Zheng, Y. Zhang, X. Li, J. Fang, Y. Liu, C. Liao, Y. Li, and J. Zhao, "Dissolved gases forecasting based on wavelet least squares support vector regression and imperialist competition algorithm for assessing incipient faults of transformer polymer insulation," *Polymers*, vol. 11, no. 1, p. 85, Jan. 2019.
[3] E. Li, L. Wang, and B. Song, "Fault diagnosis of power transformers with membership degree," *IEEE Access*, vol. 7, pp. 28791–28798, Feb. 2019.
[4] *Mineral Oil-Impregnated Electrical Equipment in Service: Guide to the Interpretation of Dissolved and Free Gases Analysis*, Standard IEC 60599, AMD, 2007, pp. 1–8.
[5] *IEEE Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers*, IEEE Standard C57.104-2008 (Revision of IEEE Standard C57.91-1991), 2009, pp. 1–28.
[6] J. Liu, H. Zheng, Y. Zhang, H. Wei, and R. Liao, "Grey relational analysis for insulation condition assessment of power transformers based upon conventional dielectric response measurement," *Energies*, vol. 10, no. 10, p. 1526, 2017.
[7] Y. Zhang, J. Liu, H. Zheng, H. Wei, and R. Liao, "Study on quantitative correlations between the ageing condition of transformer cellulose insulation and the large time constant obtained from the extended Debye model," *Energies*, vol. 10, no. 11, p. 1842, Nov. 2017.
[8] H. Zheng, J. Liu, Y. Zhang, Y. Ma, Y. Shen, X. Zhen, and Z. Chen, "Effectiveness analysis and temperature effect mechanism on chemical and electrical-based transformer insulation diagnostic parameters obtained from PDC data," *Energies*, vol. 11, no. 1, p. 146, Jan. 2018.

[9] J. Liu, H. Zheng, Y. Zhang, T. Zhou, J. Zhao, J. Li, J. Liu, and J. Li, "Comparative investigation on the performance of modified system poles and traditional system poles obtained from PDC data for diagnosing the ageing condition of transformer polymer insulation materials," *Polymers*, vol. 10, no. 2, p. 191, Feb. 2018.

[10] A. Vani and P. S. R. C. Murthy, "A hybrid neuro genetic approach for analyzing dissolved gases in power transformers," *Int. J. Adv. Res. Electr. Electron. Instrum. Eng.*, vol. 3, no. 11, pp. 13101–13107, 2014.

[11] L. Y. Chuang, S. W. Tsai, and C. H. Yang, "Improved binary particle swarm optimization using catfish effect for feature selection," *Expert Syst. Appl.*, vol. 38, no. 10, pp. 12699–12707, 2011.

[12] W. Bin, P. Qinke, Z. Jing, and C. Xiao, "A binary particle swarm optimization algorithm inspired by multi-level organizational learning behavior," *Eur. J. Oper. Res.*, vol. 219, no. 2, pp. 224–233, Jun. 2012.

[13] M. Ghasemi, S. Ghavidel, and S. Rahmani, "A novel hybrid algorithm of imperialist competitive algorithm and teaching learning algorithm for optimal power flow problem with non-smooth cost functions," *Eng. Appl. Artif. Intell.*, vol. 29, pp. 54–69, Mar. 2014.

[14] R. A. Gallego, R. Romero, and A. J. Monticelli, "Tabu search algorithm for network synthesis," *IEEE Trans. Power Syst.*, vol. 15, no. 2, pp. 490–495, May 2000.

[15] J. Z. Li, Q. G. Zhang, K. Wang, and J. Y. Wang, "Optimal dissolved gas ratios selected by genetic algorithm for power transformer fault diagnosis based on support vector machine," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 23, no. 2, pp. 1198–1206, Apr. 2016.

[16] P. Ghamisi and J. A. Benediktsson, "Feature selection based on hybridization of genetic algorithm and particle swarm optimization," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 2, pp. 309–313, Feb. 2015.

[17] T. Kari, W. Gao, and D. Zhao, "Hybrid feature selection approach for power transformer fault diagnosis based on support vector machine and genetic algorithm," *IET Gener. Transmiss. Distrib.*, vol. 12, no. 21, pp. 5672–5680, 2018.

[18] J. Huang, Y. Cai, and X. Xu, "A hybrid genetic algorithm for feature selection wrapper based on mutual information," *Pattern Recognit. Lett.*, vol. 28, no. 13, pp. 1825–1844, May 2007.

[19] A. Bouraoui, S. Jamoussi, and Y. Benayed, "A multi-objective genetic algorithm for simultaneous model and feature selection for support vector machines," *Artif. Intell. Rev.*, vol. 50, no. 2, pp. 261–281, Feb. 2017.

[20] H. B. Zheng, R. J. Liao, S. Grzybowski, and L. J. Yang, "Fault diagnosis of power transformers using multi-class least square support vector machines classifiers with particle swarm optimisation," *IET Electr. Power Appl.*, vol. 5, no. 9, pp. 691–696, Nov. 2011.

[21] J. Fang, H. Zheng, J. Liu, J. Zhao, Y. Zhang, and K. Wang, "A transformer fault diagnosis model using an optimal hybrid dissolved gas analysis features subset with improved social group optimization-support vector machine classifier," *Energies*, vol. 11, no. 8, p. 1922, Jul. 2018.

[22] K. Bacha, S. Souahlia, and M. Gossa, "Power transformer fault diagnosis based on dissolved gas analysis by support vector machine," *Electr. Power Syst. Res.*, vol. 83, no. 1, pp. 73–79, Sep. 2011.

[23] Y. Zhang, J. Liu, H. Zheng, and K. Wang, "Feasibility of a universal approach for temperature correction in frequency domain spectroscopy of transformer insulation," *IEEE Trans. Dielectr. Electr. Insul*, vol. 25, no. 5, pp. 1766–1773, Oct. 2018.

[24] X. Huang, Y. Zhang, J. Liu, H. Zheng, and K. Wang, "A novel fault diagnosis system on polymer insulation of power transformers based on 3-stage GA-SA-SVM OFC selection and ABC-SVM classifier," *Polymers*, vol. 10, no. 10, p. 1096, Oct. 2018.

[25] C. Pan, W. Chen, and Y. Yun, "Fault diagnostic method of power transformers based on hybrid genetic algorithm evolving wavelet neural network," *IET Electr. Power Appl.*, vol. 2, no. 1, pp. 71–76, Jan. 2008.

[26] F. Jawad and S. Milad, "Assessment of computational intelligence and conventional dissolved gas analysis methods for transformer fault diagnosis," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 25, no. 5, pp. 1798–1806, Oct. 2018.

[27] S.-W. Fei and Y. Sun, "Forecasting dissolved gases content in power transformer oil based on support vector machine with genetic algorithm," *Electr. Power Syst. Res.*, vol. 78, no. 3, pp. 507–514, Mar. 2008.

[28] Z. Chen, T. Lin, N. Tang, and X. Xia, "A parallel genetic algorithm based feature selection and parameter optimization for support vector machine," *Sci. Program.*, vol. 2016, no. 2, pp. 1–10, Jun. 2016.

[29] H. Wei, Y. Wang, L. Yang, C. Yan, Y. Zhang, and R. Liao, "A new support vector machine model based on improved imperialist competitive algorithm for fault diagnosis of oil-immersed transformers," *J. Elect. Eng. Technol.*, vol. 12, no. 2, pp. 830–839, 2017.

[30] H. Zheng, Y. Zhang, J. Liu, H. Wei, J. Zhao, and R. Liao, "A novel model based on wavelet LS-SVM integrated improved PSO algorithm for forecasting of dissolved gas contents in power transformers," *Electr. Power Syst. Res.*, vol. 155, pp. 196–205, Feb. 2018.

[31] M. Duval and A. dePabla, "Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases," *IEEE Electr. Insul. Mag.*, vol. 17, no. 2, pp. 31–41, Mar. 2001.

[32] Z. Li, Q. Cao, Y. Zhao, P. Tao, and R. Zhuo, "Krill herd algorithm for signal optimization of cooperative control with traffic supply and demand," *IEEE Access*, vol. 7, pp. 10776–10786, Jan. 2019.

[33] J. Li, Y. Tang, C. Hua, and X. Guan, "An improved krill herd algorithm: Krill herd with linear decreasing step," *Appl. Math. Comput.*, vol. 234, pp. 356–367, May 2014.

[34] G. Wang, A. H. Gandomi, and A. H. Alavi, "An effective krill herd algorithm with migration operator in biogeography-based optimization," *Appl. Math. Model.*, vol. 38, nos. 9–10, pp. 2454–2462, May 2014.

[35] A. Mukherjee and V. Mukherjee, "Solution of optimal reactive power dispatch by chaotic krill herd algorithm," *IET Gener. Transmiss. Distrib.*, vol. 9, no. 15, pp. 2351–2362, Nov. 2015.

[36] 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.

[37] I. B. M. Taha, D.-E. A. Mansour, S. S. M. Ghoneim, and N. I. Elkalashy, "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.

[38] 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.

[39] 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, Oct. 2009.

[40] 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.

[41] J. Liu, X. Fan, H. Zheng, Y. Zhang, C. Zhang, B. Lai, J. Wang, G. Ren, and E. Zhang, "Aging condition assessment of transformer oil-immersed cellulosic insulation based upon the average activation energy method," *Cellulose*, vol. 26, no. 6, pp. 3891–3908, Mar. 2019.