

Received August 24, 2021, accepted September 8, 2021, date of publication September 13, 2021, date of current version September 20, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3112478

# A Transformer Fault Diagnosis Method Based on Parameters Optimization of Hybrid Kernel Extreme Learning Machine

JIFANG LI<sup>1</sup>, CHEN HAI<sup>1</sup>, ZHEN FENG<sup>1</sup>, AND GENXU LI<sup>1</sup>

College of Electricity, North China University of Water Resources and Electric Power, Zhengzhou, Henan 450000, China

Corresponding author: Chen Hai (haichen00@126.com)

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

**ABSTRACT** Dissolved gas analysis (DGA) is a widely used method for diagnosing internal transformer defects. The traditional single intelligent diagnostic method cannot efficiently process large amounts of incomplete defect information with DGA, which affects the accuracy of fault diagnosis. To this end, this paper proposes a transformer fault diagnosis method based on the optimization of kernel parameters and weight parameters of a kernel extreme learning machine (KELM). First, based on Mercer's theorem, we combine the radial basis kernel function with the polynomial kernel function to construct a new hybrid kernel function. Then, the gray wolf optimization (GWO) algorithm and the differential evolution (DE) algorithm are combined to improve the diversity of the gray wolf population, enhance the searchability of GWO, and prevent GWO from falling into a local optimum during the iterative process. Finally, the kernel parameters and weight parameters of the hybrid kernel function are optimized by using the modified grey wolf optimization (MGWO) algorithm. The International Electrotechnical Commission Technical Committee (IEC TC) 10 transformer fault data and constructed hybrid feature set is used as the input set of the model, the model is simulated and analyzed, and the transformer fault data collected at a site are used for training and verification. The simulation results on the two sets of data show that the method can accurately and effectively diagnose transformer faults, and has a higher fault diagnosis accuracy rate than traditional methods.

**INDEX TERMS** Power transformers, fault diagnosis, extreme learning machine, hybrid kernel function, modified gray wolf optimization.

## I. INTRODUCTION

A power transformer is an important part of a power system. It undertakes the tasks of power conversion and power transmission, and plays a vital role in the power system. Since transformer failures may induce very large economic losses, formulating corresponding measures in advance according to the transformer status, and timely detecting and accurately determining latent transformer faults, are of great significance for extending the life of a transformer and improving the safety, reliability, and economy of a power system.

Dissolved gas analysis (DGA) is an important tool for detecting early faults of oil-filled transformers. Once an oil-immersed transformer fails, insulators in the transformer, such as the internal solid and liquid materials, may chemically

decompose due to electrical and thermal stress, leading to gas release [1].

Therefore, we can detect abnormal states of a transformer by identifying the composition and content of the dissolved gas in the oil, and further determine the fault type, severity and development trend. Traditional fault diagnosis methods mainly include the ratio method [2], [3], key gas method [4], triangle method [5], and pentagon method [6]–[8]. Although they are simple and effective, these methods still have many problems, such as inconsistent diagnostic results and low accuracy, which reduce the reliability of fault analysis. In recent years, with the continuous development of artificial intelligence (AI) technology, machine learning and pattern recognition methods have been widely used in transformer fault diagnosis. Artificial neural networks [9], [10], support vector machines (SVM) [11], fuzzy logic [12], Bayesian neural networks [13], adaptive network-based fuzzy inference

The associate editor coordinating the review of this manuscript and approving it for publication was Huaqing Li<sup>1</sup>.

systems [14], deep belief networks [15] and other technologies have been applied to transformer fault tasks. These methods compensate for the shortcomings of traditional DGA methods, directly or indirectly improve the accuracy of fault diagnosis, and provide new ideas for transformer fault diagnosis.

The extreme learning machine (ELM) proposed by Guang-Bin Huang *et al.* in 2004 provides a unified learning platform with a wide range of feature mapping types, and thus can approximate any continuous objective function, making classification of disjoint areas possible. Since they have better scalability, stronger generalization and faster learning speeds than SVMs [16], ELMs have been widely used in the fields of fault diagnosis, such as transmission line fault diagnosis [17], rolling bearing fault diagnosis [18]–[20], and analog circuit fault diagnosis [21], [22]. In the field of transformer fault diagnosis, Malik and Mishra [23] applied a principal component analysis on International Electrotechnical Commission Technical Committee (IEC TC) 10 DGA data to find the most relevant variables first and then used an ELM to classify early transformer faults. Compared with fuzzy logic and artificial neural networks, their approach has better diagnostic accuracy. Li [24] *et al.* first used an adaptive evolutionary ELM to optimize the weights and bias of a network and then utilized an arctangent transformation to change the structure of the experimental data, leading to an improved generalization ability of the algorithm to improving. Huang [25] *et al.* proposed ELM power transformer fault diagnosis technology based on multiscale information fusion. Specifically, multiple ELM models with different numbers of nodes are employed to generate the initial diagnosis results and then multiscale information fusion is used to achieve a final fault diagnosis result by fusing the initial diagnosis results. Although it has advantages, including fast calculation speeds, strong generalization abilities, and anti-overfit abilities, Wei's method selects the optimal number of neurons in the hidden layer of the ELM according to trial and error, and thus may not fully reveal the sample information, which will result in overfitting when the number of hidden layers is too large. The ELM based kernel function may be an effective method for addressing the above problem. However, the performance of the ELM based kernel function is sensitive to the parameters and regularization coefficients of the kernel function. This paper combines the modified gray wolf optimization (MGWO) algorithm with the hybrid kernel ELM (KELM) to construct a transformer fault diagnosis method. In the proposed method, a hybrid kernel function with a linear combination of a global kernel function and a local kernel function is proposed to improve the learning ability and generalization ability of the KELM. Moreover, modified GWO is proposed by combining differential evolution (DE) [26] and gray wolf optimization (GWO) to optimize the parameters of the hybrid kernel; thus, the optimized network structure of the hybrid KELM is achieved, improving the accuracy of transformer fault diagnosis. Finally, the experimental results on IEC TC 10 DGA data and our data show that the KELM algorithm has a better

classification effect on the five types of transformer faults than the SVM algorithm.

## II. TRANSFORMER FAULT DIAGNOSIS BASED ON A KERNEL EXTREME LEARNING MACHINE

### A. KERNEL EXTREME LEARNING MACHINE

The ELM output function in the case of a single output node is:

$$f(x) = \sum_{i=1}^L \beta_i G(a_i, b_i, x) = \boldsymbol{\beta} \cdot \mathbf{h}(x) \quad (1)$$

where  $\beta_i$  is the output weight between the  $i$ -th node of the hidden layer and the output layer and  $\boldsymbol{\beta} = [\beta_1, \dots, \beta_L]^T$  is the output weight vector.  $G(a_i, b_i, x)$  is the output of the  $i$ -th hidden layer node, and the node parameter is randomly generated.  $\mathbf{h}(x) = [G(a_1, b_1, x), \dots, G(a_L, b_L, x)]^T$  is the output vector of the hidden layer relative to the input. The output vector  $\mathbf{h}(x)$  is a feature map that maps data from the  $n$ -dimensional input space to the  $L$ -dimensional hidden layer feature space  $\mathbf{H}$ .

After introducing the kernel function, the kernel matrix of the KELM can be defined as:

$$\Omega_{\text{ELM}} = \mathbf{H}\mathbf{H}^T : \Omega_{\text{ELM}} = \mathbf{h}(x_i) \cdot \mathbf{h}(x_j) = K(x_i, x_j) \quad (2)$$

The output function of the ELM classifier can be further written as:

$$\begin{aligned} f(x) &= \mathbf{h}(x) \mathbf{H}^T \left( \frac{\mathbf{I}}{\lambda} + \mathbf{H}\mathbf{H}^T \right)^{-1} \mathbf{T} \\ &= \begin{bmatrix} K(x, x_1) \\ \vdots \\ K(x, x_N) \end{bmatrix}^T \left( \frac{\mathbf{I}}{\lambda} + \Omega_{\text{ELM}} \right)^{-1} \mathbf{T} \end{aligned} \quad (3)$$

where  $\mathbf{I}$  is the identity matrix,  $\lambda$  is the regularization coefficient, and  $\mathbf{T}$  is the training set label.

After we use this method, we do not need to know the specific form of the feature map  $\mathbf{h}(x)$  but use the kernel function for the output calculation. Thus, the random generation of weights and bias is avoided, and there is no need to set the number of hidden layer neurons  $L$ .

### B. HYBRID KERNEL FUNCTIONS

Traditional kernel functions include global kernel functions and local kernel functions. A global kernel function emphasizes the influence of the data as a whole on the kernel function; hence, the generalization performance is stronger, but the learning ability is weak. A local kernel function emphasizes the influence of the data near the key point on the kernel function; hence it has a strong learning ability, but the generalization performance is weak. According to Mercer's theorem, the nonnegative linear combination of Mercer nuclei is still a Mercer nucleus [27]. To obtain a strong learning ability and generalization ability, two different types of kernel functions are merged by means of variable weights to construct a hybrid kernel function.

The Gaussian kernel function  $K(x,y) = \exp(-\gamma \cdot \|x - y\|^2)$  can map the input space to an infinite-dimensional feature space. The function structure is simple, the convergence speed is fast, and the learning ability is strong. Therefore, the Gaussian kernel function is selected as the local kernel function [28]. In the function,  $\gamma$  is the width control parameter, which can control the radial range of the function. When the input sample changes in a large range, the polynomial kernel function  $K(x,y) = ((x \cdot y) + \eta)^d$  still has a greater impact on the sample and has a good generalization performance. Therefore, a polynomial function is used as the global kernel function, and the hybrid kernel function after fusion is:

$$K_H(x,y) = (1 - \omega) \exp(-\gamma \cdot \|x-y\|^2) + \omega ((x \cdot y) + \eta)^d \quad (4)$$

where  $\omega$  is an adjustable global kernel function weight parameter, which defines the relative contribution of the global kernel to the hybrid kernel function. It can assign different linear weights to the global kernel function and the local kernel function.

### C. FAULT DIAGNOSIS MODEL

Transformer fault diagnosis is essentially a multiclassification problem. This paper uses a hybrid kernel function ELM as a classifier to extract the internal features of transformer fault data. The constructed fault diagnosis model mainly includes the following:

- (1) Sample collection: Transformer DGA data containing various fault types are collected to form a fault sample set.
- (2) Feature selection: The hybrid feature set is used as the input of the fault diagnosis model.
- (3) Normalization processing: To eliminate the difference in the data size of different features, the feature data are normalized preprocessing with formula (5), and the normalized sample value is in the range of [0, 1]. Hence, the calculation speed of the model is increased.

$$x^* = \frac{x_i - x_{i \min}}{x_{i \max} - x_{i \min}} \quad (5)$$

where  $x^*$  is the value after normalization and  $x_{i \max}$  and  $x_{i \min}$  are the maximum and minimum values of the data before normalization, respectively.

- (4) Sample division: Five-fold cross validation is used to randomly divide the samples into a training set and a test set. Cross validation can be used to effectively evaluate the performance of the training model and improve the stability and generalization ability of the model.
- (5) Model training: The training set is used to train the model and to construct the optimal fault diagnosis model.
- (6) Statistical results: The test set is used to test the effect of the model, cross validation is applied and the training results are counted.

The flowchart of the fault diagnosis model is shown in Fig.1.

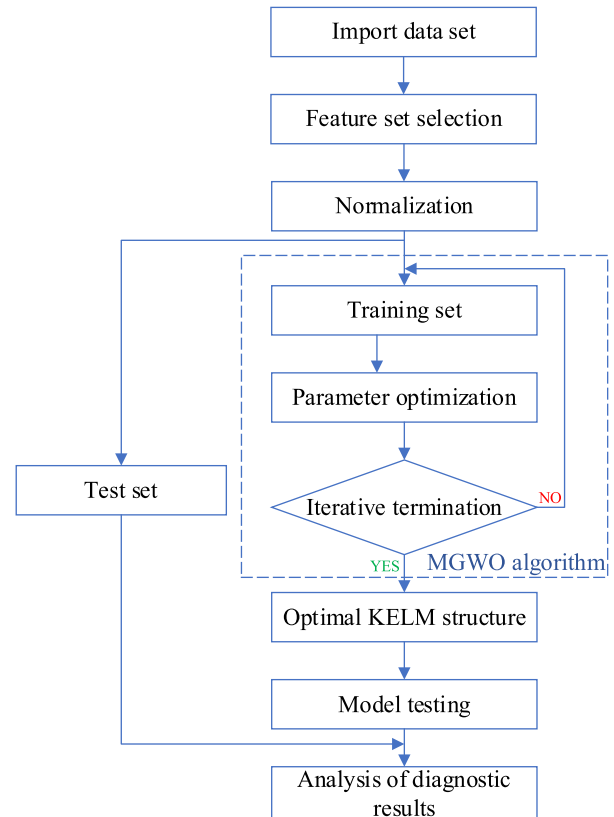


FIGURE 1. Fault diagnosis flowchart of the hybrid kernel extreme learning machine.

### III. PARAMETER OPTIMIZATION BASED ON THE MODIFIED GRAY WOLF ALGORITHM

#### A. MODIFIED GRAY WOLF OPTIMIZATION ALGORITHM

GWO is a group intelligence algorithm based on gray wolves' social hierarchy and hunting behaviors [29]. In traditional GWO, each wolf pack is generated by the initial wolf pack, which means that the GWO algorithm's accuracy depends largely on the initial wolf pack. If the randomly generated initial wolf pack selection fails, the algorithm will converge prematurely, and the accuracy will be reduced. In the later stages of the iteration, when each individual is close to the prey, the algorithm can become easily stuck in a state of stagnation and can lack diversity. The DE algorithm is introduced to improve GWO, and GWO is forced to jump out of the local optimum through mutation, crossover, and selection.

In the GWO algorithm, each gray wolf represents a candidate solution in the population. The optimal solution is defined as alpha ( $\alpha$ ), the second-best solution is defined as beta ( $\beta$ ), the third-best solution is delta ( $\delta$ ), and the remaining candidate solutions are assumed to be omega ( $\omega$ ) wolves. In GWO, the optimization process is guided by  $\alpha$ ,  $\beta$ , and  $\delta$ . The location update process is shown in

formula (6)-formula (13):

$$\vec{D} = \left| \vec{C} \cdot \vec{X}_p(t) - \vec{X}(t) \right| \tag{6}$$

$$\vec{X}(t+1) = \vec{X}_p(t) - \vec{A} \cdot \vec{D} \tag{7}$$

$$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a} \tag{8}$$

$$\vec{C} = 2 \cdot \vec{r}_2 \tag{9}$$

$$\vec{a} = 2 - 2 \left( \frac{t}{t_{\max}} \right) \tag{10}$$

$$\begin{cases} \vec{D}_\alpha = \left| \vec{C}_1 \cdot \vec{X}_\alpha - \vec{X} \right| \\ \vec{D}_\beta = \left| \vec{C}_2 \cdot \vec{X}_\beta - \vec{X} \right| \\ \vec{D}_\delta = \left| \vec{C}_3 \cdot \vec{X}_\delta - \vec{X} \right| \end{cases} \tag{11}$$

$$\begin{cases} \vec{X}_1 = \vec{X}_\alpha - \vec{A}_1 \cdot (\vec{D}_\alpha) \\ \vec{X}_2 = \vec{X}_\beta - \vec{A}_2 \cdot (\vec{D}_\beta) \\ \vec{X}_3 = \vec{X}_\delta - \vec{A}_3 \cdot (\vec{D}_\delta) \end{cases} \tag{12}$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \tag{13}$$

where  $t$  represents the current iteration number;  $\vec{A}$  and  $\vec{C}$  are coefficient vectors;  $\vec{X}_\alpha$ ,  $\vec{X}_\beta$  and  $\vec{X}_\delta$  represent the current positions of  $\alpha$ ,  $\beta$  and  $\delta$  respectively;

After each position update, the gray wolf population generates a mutated individual  $M_{i,t}$  with formula (14):

$$M_{i,t} = X_{best,t} + F \cdot (X_{r_1,t} - X_{r_2,t}) \tag{14}$$

Among them, the parameters  $r_1^i$  and  $r_2^i$  are mutually exclusive integers randomly generated within the range of  $[1, NP]$ ;  $F$  is the scale factor; and  $X_{best,t}$  is the wolf pack individual with the best fitness value in the  $t$ -th generation group.

After the mutation stage is over, the test individual  $U_{i,t} = [u_{i,t}^1, u_{i,t}^2, \dots, u_{i,t}^D]$  is obtained through the crossover operation, and the crossover operation is shown in formula (15):

$$u_{i,t}^j = \begin{cases} m_{i,t}^j, & \text{if } (rand(0, 1) \leq CR) \\ & \text{or } (j = j_{rand}) \\ x_{i,t}^j, & \text{otherwise} \end{cases} \quad j = 1, 2, \dots, D \tag{15}$$

where  $CR$  is the crossover rate; and  $j_{rand}$  is a randomly generated integer to ensure that test individual is different from  $X_{i,t}$  by at least one parameter.

Finally, the individuals who enter the next generation are selected through the greedy algorithm. The selection operation is shown in formula (16):

$$X_{i,t+1} = \begin{cases} U_{i,t}, & \text{if } f(X_{i,t}) \leq f(U_{i,t}) \\ X_{i,t}, & \text{otherwise} \end{cases} \tag{16}$$

The pseudo process of the modified gray wolf algorithm is shown in Table 1.

### B. MGWO EFFICIENCY TEST

To verify the optimized performance of MGWO, 13 benchmark functions [30], [31] are used to test the algorithm. To verify the results, the MGWO algorithm is compared to the particle swarm optimization (PSO) and traditional GWO algorithms. The test function is shown in Table 2. Among them, the function dimension, variable search range and function optimal value are abbreviated as Dim value, Range and  $f_{\min}$  respectively. In addition, the Dim, the population size, and the maximum number of iterations of the 13 benchmark functions are set to 30, 30, and 1000, respectively. To avoid random errors, each test function was run independently 30 times, and its average results and standard deviations were recorded. The test results are shown in Table 3.

As seen from the table, MGWO is significantly better than PSO for 13 out of 10 benchmark functions. Moreover, the results of MGWO are significantly better than those of traditional GWO in dealing with all the selected functions. In short, the modified GWO algorithm has a more powerful search performance and can avoid local optima very well.

### C. PARAMETER OPTIMIZATION BASED ON MGWO

The parameter setting has a great influence on the classification accuracy of the KELM. As the parameters change, the correct rate of the fault diagnosis model also changes. The parameters of the hybrid KELM are optimized by using the modified GWO algorithm. The specific optimization steps are as follows:

- (1) The initial parameter settings include the number of gray wolf individuals  $NP$ , the maximum number of iterations  $t_{\max}$ , the parameter dimension  $D$ , the scale factor  $F$ , and the crossover rate.
- (2) Set the value range of each parameter  $\lambda, \gamma, \eta, \omega$  of the hybrid KELM. Initialize the gray wolf population, and each gray wolf individual corresponds to a set of parameters.
- (3) Calculate and rank the fitness of the gray wolf population. The top three individuals are the  $\alpha$  wolf,  $\beta$  wolf and  $\delta$  wolf.
- (4) Initialize random numbers  $\vec{r}_1$  and  $\vec{r}_2$ , and use formulas (8), (9) and (10) to calculate  $\vec{A}$ ,  $\vec{C}$ , and  $\vec{a}$ , respectively. Formulas (11), (12), and (13) are used to update the gray wolf population positions.
- (5) Perform the mutation operation to generate mutant individuals, and then perform crossover operations between the mutant individuals and the gray wolf population according to formula (15) to obtain test individuals.
- (6) According to formula (16), judge whether each test individual is retained.
- (7) Calculate the fitness of the updated population and sort the population. If the termination condition is met, the optimal individual and its fitness are output. Otherwise, return to step 4.

TABLE 1. Pseudo code of MGWO.

<b>Algorithm: MGWO</b>	
<b>Input:</b>	Number of gray wolves ( $NP$ ), the maximum number of iterations ( $t_{\max}$ ), parameter dimension ( $D$ ), scale factor ( $F$ ), crossover rate ( $CR$ )
<b>Output:</b>	The location and fitness of the best individual.
1	Randomly generate the first generation of wolves with ( $NP$ ) individuals.
2	$t = 0$ ;
3	<b>while</b> $t < t_{\max}$ <b>do</b>
4	- Calculate the fitness of each wolf;
5	- Record the current three individuals with the best fitness and update the positions of $\alpha$ , $\beta$ and $\delta$ ;
6	- Use formulas (8), (9) and (10) to calculate $\bar{A}$ , $\bar{C}$ and $\bar{a}$ respectively;
7	<b>for</b> $i = 1$ to $NP$ <b>do</b>
8	Use formula (13) to update the gray wolf position;
9	<b>end</b>
10	<b>for</b> $i = 1$ to $NP$ <b>do</b>
11	Use formula (14) to generate a mutated individual $M_{i,t}$ ;
12	<b>end</b>
13	<b>for</b> $i = 1$ to $NP$ <b>do</b>
14	$j_{rand} = \lceil rand[0,1) \times D \rceil$ ;
15	<b>for</b> $j = 1$ to $D$ <b>do</b>
16	$u_{i,t}^j = \begin{cases} m_{i,t}^j, & \text{if } (rand(0,1) \leq CR) \text{ or } (j = j_{rand}) \\ x_{i,t}^j, & \text{otherwise} \end{cases} \quad j = 1, 2, \dots, D$ ;
17	<b>end</b>
18	<b>end</b>
19	<b>for</b> $i = 1$ to $NP$ <b>do</b>
20	<b>if</b> $f(X_{i,t}) \leq f(U_{i,t})$ <b>then</b>
21	$X_{i,t} = U_{i,t}$ ;
22	$f(X_{i,t}) = f(U_{i,t})$ ;
23	<b>end</b>
24	<b>end</b>
25	$t = t + 1$ ;
26	<b>end</b>
27	Return to the optimal gray wolf position;

## IV. SIMULATION EXPERIMENT ANALYSIS

### A. TRANSFORMER FAULT TYPES AND SIMULATION

#### EXPERIMENT DATA

According to IEC 60599, transformer fault types are divided into five types: partial discharge (PD), discharges of low energy (D1), discharges of high energy (D2), thermal faults of low and medium temperatures (T1 and T2), and thermal faults of high temperatures (T3). The IEC TC 10 transformer fault data is used to train and test the model, and the transformer fault data collected in China is used to test the generalization performance of the diagnostic model. One hundred seventeen sets of IEC TC 10 transformer fault data and three hundred seventy sets of domestic transformer fault data were collected. The distribution of the fault samples is shown in Table 4.

### B. FEATURE SET SELECTION

Feature selection is the key to classification models. For transformer fault diagnostic models, different diagnosis methods use different feature combinations. AI methods often use DGA gas content as the input of the diagnostic model, whereas traditional DGA diagnostic methods often use the dissolved gas ratio as a feature combination. To obtain the core attributes contained in the sample data, based on the feature combination corresponding to the above methods, this paper proposes a hybrid feature set. The feature combinations corresponding to different diagnostic methods are shown in Table 5.

The five traditional feature sets in Table 5 are used to perform fault diagnosis on the IEC TC 10 transformer fault data, and they are with the proposed hybrid feature sets. The



TABLE 2. Benchmark functions.

Function	Dim	Range	$f_{\min}$
$f_1(x) = \sum_{i=1}^n x_i^2$	30	[-100,100]	0
$f_2(x) = \sum_{i=1}^n  x_i  + \prod_{i=1}^n  x_i $	30	[-10,10]	0
$f_3(x) = \sum_{i=1}^n \left(\sum_{j=1}^i x_j\right)^2$	30	[-100,100]	0
$f_4(x) = \max_i \{ x_i , 1 \leq i \leq n\}$	30	[-100,100]	0
$f_5(x) = \sum_{i=1}^{n-1} \left[100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2\right]$	30	[-30,30]	0
$f_6(x) = \sum_{i=1}^n ([x_i + 0.5])^2$	30	[-100,100]	0
$f_7(x) = \sum_{i=1}^n ix_i^4 + \text{random}[0,1)$	30	[-1.28,1.28]	0
$f_8(x) = \sum_{i=1}^n -x_i \sin(\sqrt{ x_i })$	30	[-500,500]	-418.98 29*n
$f_9(x) = \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i) + 10]$	30	[-5.12,5.12]	0
$f_{10}(x) = -20 \exp\left(-0.2 \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}\right) - \exp\left(\frac{1}{n} \sum_{i=1}^n \cos(2\pi x_i)\right) + 20 + e$	30	[-32,32]	0
$f_{11}(x) = \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	30	[-600,600]	0
$f_{12}(x) = \frac{\pi}{n} \left\{10 \sin(\pi y_1) + \sum_{i=1}^{n-1} (y_i - 1)^2 [1 + 10 \sin^2(\pi y_{i+1})] + (y_n - 1)^2\right\} + \sum_{i=1}^n \mu(x_i, 10, 100, 4)$	30	[-50,50]	0
$y_i = 1 + \frac{x_i + 1}{4} \quad \mu(x_i, a, k, m) = \begin{cases} k(x_i - a)^m & x_i - a \\ 0 & -a < x_i < a \\ k(-x_i - a)^m & x_i < -a \end{cases}$			
$f_{13}(x) = 0.1 \left\{\sin^2(3\pi x_1) + \sum_{i=1}^n (x_i - 1)^2 [1 + \sin^2(3\pi x_i + 1)] + (x_n - 1)^2 [1 + \sin^2(2\pi x_n)]\right\} + \sum_{i=1}^n \mu(x_i, 5, 100, 4)$	30	[-50,50]	0

SVM method is used to carry out the simulation test, and the correct average rate of the cross-validation fault diagnosis task is shown in Table 6. Table 6 shows that the diagnostic accuracy is the highest when the mixed feature set is used as the input for fault diagnosis.

C. SEARCH SPACE

The generalization performance of the KELM is closely related to the parameters of the kernel function. To obtain a great generalization performance and improve the convergence speed, it is necessary to select an appropriate parameter optimization space. Performance tests are performed on the radial basis kernel function parameter  $(\lambda, \gamma)$  and the polynomial kernel function parameter  $(\lambda, \eta)$ . Based on the same data set  $\lambda, \gamma, \eta \in \{2^{-24}, 2^{-23}, \dots, 2^{24}, 2^{25}\}$ , each pair of  $(\lambda, \gamma)$  and  $(\lambda, \eta)$  has 2500 different combinations, and the

polynomial kernel parameter  $d = 3$ . The simulation results are shown in Fig. 2 to Fig. 5. Fig. 2 is a grid diagram of the KELM classification accuracy that uses the radial basis kernel function. Fig.3 is a contour map of KELM classification accuracy using the radial basis kernel function. Fig. 4 is a grid diagram of the KELM classification accuracy using the polynomial kernel function. Fig. 5 is a contour map of the KELM classification accuracy using the polynomial kernel function.

In Fig. 2 to Fig. 5, the deeper the yellow dots are, the higher the accuracy of the diagnosis is. The deeper the purple dots are, the lower the fault diagnostic accuracy is, and the red dots are the points with the maximum fault diagnosis accuracies. Fig. 2 to Fig. 5 show that the selection of parameter combinations A and B has a significant impact on the accuracy of fault diagnosis. Only in a very narrow range can the fault

TABLE 3. Results of the benchmark functions.

Function	MGWO		GWO		PSO	
	Ave	Std	Ave	Std	Ave	Std
$f_1$	<b>1.6461E-35</b>	2.1483E-35	1.3298E-27	2.8560E-27	1.7731E-04	1.9882E-04
$f_2$	<b>3.3316E-21</b>	3.6715E-21	1.4194E-16	1.7594E-16	0.0488	0.0854
$f_3$	<b>1.0377E-07</b>	2.7501E-07	1.3878E-05	3.4465E-05	87.4931	40.4881
$f_4$	<b>1.3409E-08</b>	1.2912E-08	8.4069E-07	1.1217E-06	1.1879	0.2134
$f_5$	<b>26.5754</b>	0.5024	27.0346	0.7513	110.2371	108.8880
$f_6$	0.5519	0.3477	0.8195	0.4855	<b>2.6167E-04</b>	7.2517E-04
$f_7$	<b>0.0016</b>	7.0528E-04	0.0018	8.9172E-04	0.1748	0.0746
$f_8$	<b>-1.0145E+04</b>	1.4857E+03	-6.1026E+03	865.8568	-4.9076E+03	1.1984E+03
$f_9$	<b>1.2522</b>	2.1886	2.6197	3.2580	58.8756	16.9085
$f_{10}$	<b>2.2086E-14</b>	4.6165E-15	9.9417E-14	2.0009E-14	0.3222	0.5714
$f_{11}$	<b>0.0020</b>	0.0053	0.0030	0.0055	0.0077	0.0074
$f_{12}$	0.0378	0.0160	0.0465	0.0220	<b>0.0138</b>	0.0450
$f_{13}$	0.4399	0.1942	0.6577	0.2086	<b>0.0081</b>	0.0197

TABLE 4. Distribution of transformer fault samples.

Fault type	PD	D1	D2	T1&T2	T3
IEC TC 10 transformer fault data	9	26	48	16	18
Domestic transformer fault data	17	47	120	31	155

TABLE 5. Diagnostic methods and feature sets.

Feature set	Feature combination
DGA gas	$H_2, CH_4, C_2H_2, C_2H_4, C_2H_6, CO, CO_2$
Doernenberg ratios	$CH_4 / H_2, C_2H_2 / C_2H_4, C_2H_6 / C_2H_2$
Rogers ratios	$CH_4 / H_2, C_2H_2 / C_2H_4, C_2H_4 / C_2H_6, C_2H_6 / CH_4$
IEC 60599 ratios	$CH_4 / H_2, C_2H_2 / C_2H_4, C_2H_4 / C_2H_6$
CIGRE ratios	$C_2H_2 / C_2H_6, H_2 / CH_4, C_2H_4 / C_2H_6, C_2H_2 / H_2, CO / CO_2$
Hybrid feature set	$CH_4, C_2H_2, C_2H_4, CH_4 / H_2, C_2H_4 / C_2H_6$

TABLE 6. SVM validation accuracy.

Feature set	Accuracy
DGA gas	71.79%
Doernenberg ratios	71.79%
Rogers ratios	70.09%
IEC 60599 ratios	70.09%
CIGRE ratios	79.49%
Hybrid feature set	81.2%

diagnosis accuracy rate reach the best. If the parameters are selected improperly, the accuracy rate will decrease sharply. The classification accuracy contour and the maximum value

points of the classification accuracy rate are combined to select an appropriate parameter range. The radial basis kernel function parameter search range from Fig. 3 is selected as  $\lambda \in [2^{-5}, 2^5]$  and  $\lambda \in [2^{-5}, 2^5]$ , and the polynomial kernel function parameter search range from Fig. 5 is selected as  $\lambda \in [2^0, 2^{10}]$  and  $\eta \in [2^{-10}, 2^0]$ . Then, the optimal parameter search ranges of the two kernel functions are shown in Table 7.

#### D. SIMULATION ANALYSIS

Based on the IEC TC 10 transformer fault data in Table 2, the hybrid feature set is used as input to test the MGWO-KELM model. The 117 groups of fault data are divided into

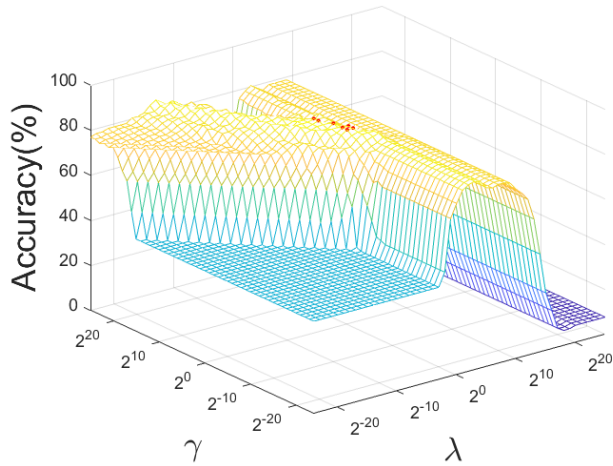


FIGURE 2. Classification accuracy of radial basis kernel function.

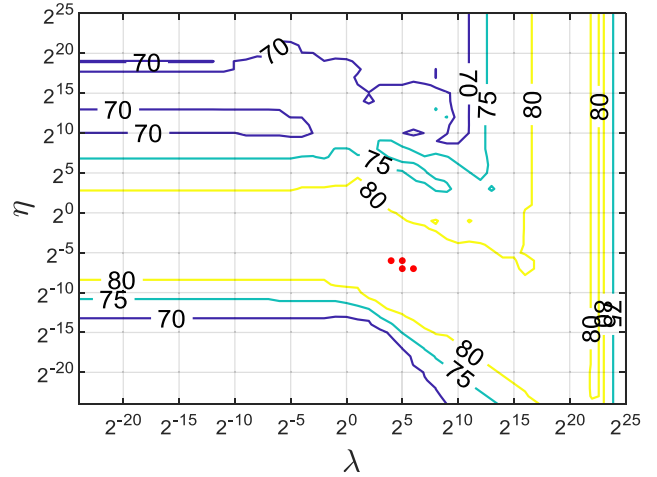


FIGURE 5. Classification contours of polynomial kernel function.

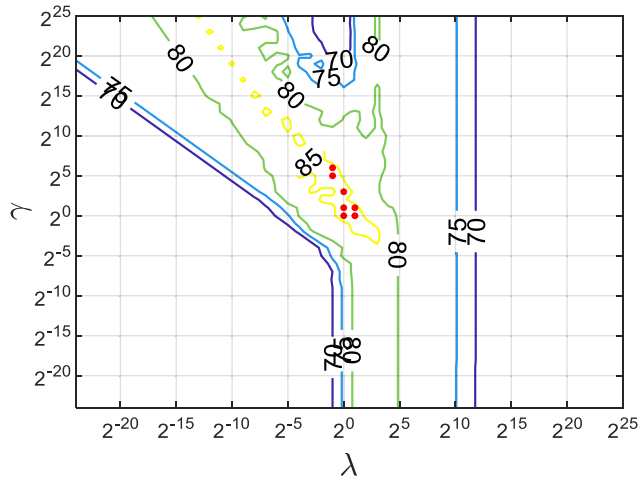


FIGURE 3. Classification contours of radial basis kernel function.

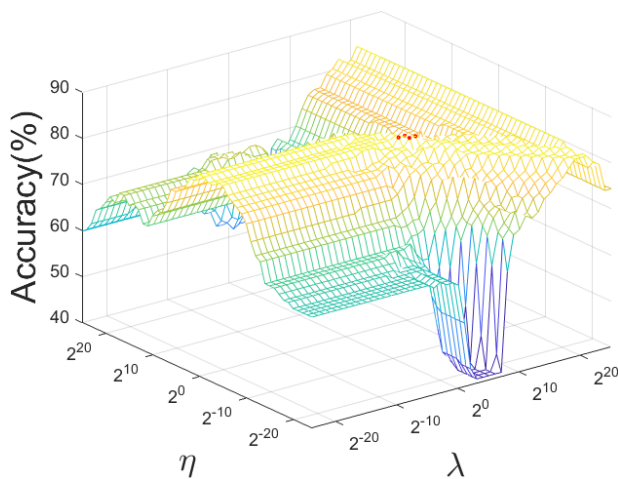


FIGURE 4. Classification accuracy of polynomial kernel function.

5 groups randomly. Four groups are used as the training set each time and the remaining group is used as the test set. The calculation is performed five times in a loop. Finally, the

TABLE 7. Search range of parameter.

Parameter	Search scope
Regularization factor $\lambda$	$[2^0, 2^5]$
Kernel parameter $\gamma$	$[2^{-3}, 2^8]$
Kernel parameter $\eta$	$[2^{-10}, 2^0]$
Weight parameter $\omega$	$[0, 1]$

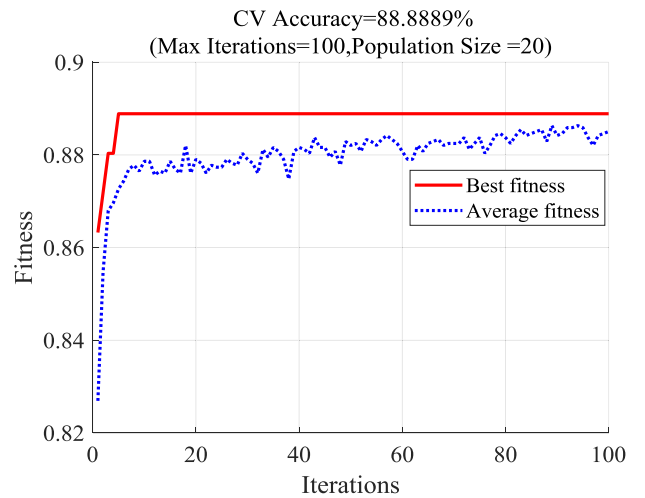


FIGURE 6. Iteration diagram of the MGWO-KELM.

correct average rate of the five test groups is used as the fitness value. The kernel function parameters and weight search range during training are shown in Table 7. The relevant initialization parameters of the modified gray wolf algorithm are set as follows: the population size is 20, the maximum number of iterations is 100, the variable dimension is 4, the scaling factor  $F = 0.5$ , and the crossover rate  $CR = 0.5$ . The iterative diagram of the MGWO-KELM model is shown in Fig. 6.



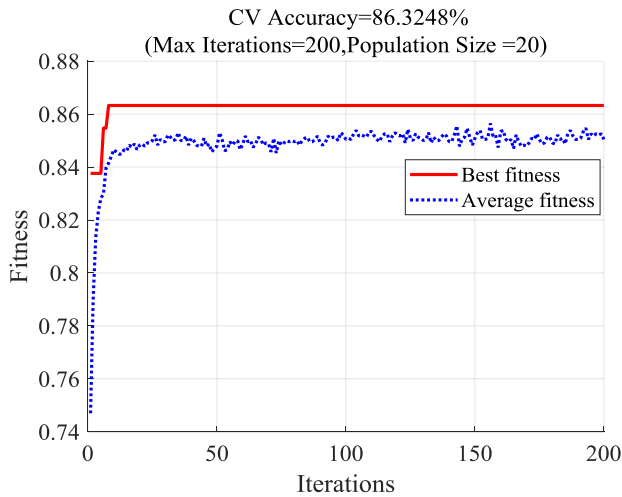


FIGURE 7. Iteration diagram of the PSO-KELM.

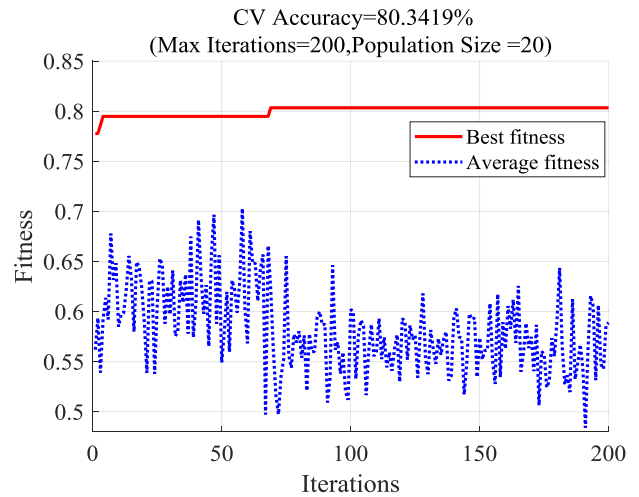


FIGURE 9. Iteration diagram of the PSO-SVM.

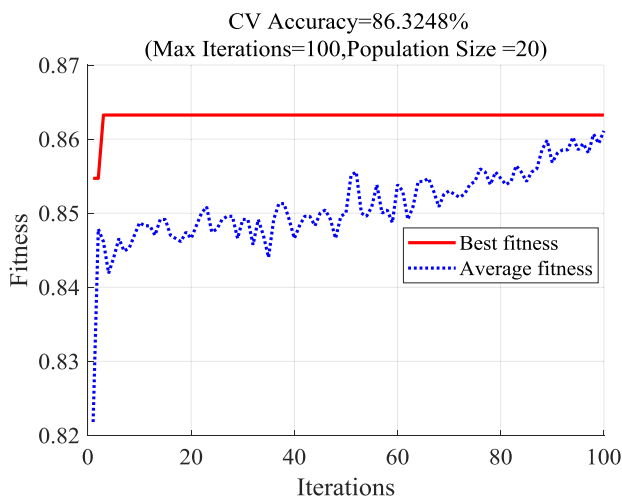


FIGURE 8. Iteration diagram of the MGWO-SVM.

Based on the same combination of DGA features, the PSO-KELM, MGWO-SVM, and PSO-SVM models are used to diagnose transformer faults. The fitness curve of the obtained PSO-KELM algorithm is shown in Fig. 7, the fitness curve of the MGWO-SVM algorithm is shown in Fig. 8, and the fitness curve of the PSO-SVM algorithm is shown in Fig. 9. As shown in Fig. 8 and Fig. 9, during the iterative process, the average fitness of the SVM algorithm frequently fluctuates, indicating that it is too sensitive to the parameters and that subtle parameter fluctuations will seriously affect the classification effect of the model. This model can increase the difficulty of optimization. Fig.7 and Fig.9 show that the PSO algorithm has poor searchability and is prone to falling into a local optimum. In addition, the algorithm stalled early in the iteration and could not optimize the model parameters. Comparing Fig. 6 with Figs. 7-9 shows that the hybrid core ELM has a strong learning ability and generalization performance leading to a higher classification accuracy rate during the training process. The modified GWO algorithm has a strong

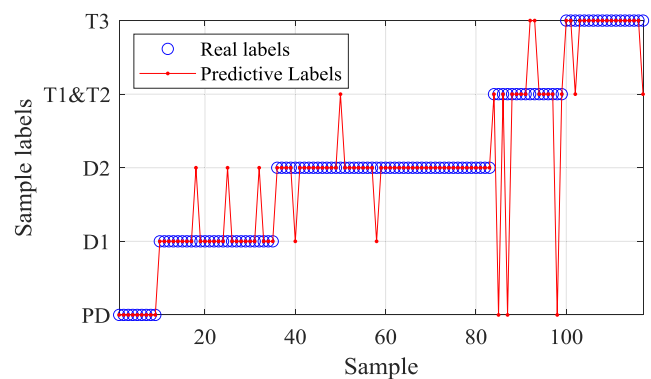


FIGURE 10. Diagnosis results of the MGWO-KELM.

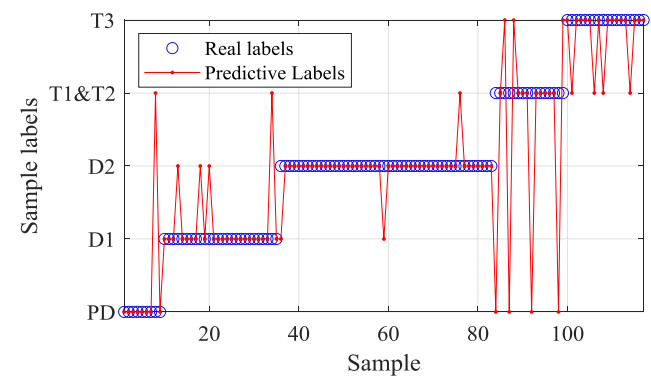


FIGURE 11. Diagnosis results of the PSO-KELM.

search ability and faster convergence speed, and it only needs a few iterations to achieve the best network structure.

Fig. 10 to Fig. 13 show the classification results on the IEC TC 10 transformer fault data of the trained model. Fig. 10 shows the fault classification result of the MGWO-KELM model proposed in this paper, Fig. 11 shows the fault classification result of the PSO-KELM model, Fig.12 shows the fault classification result of the MGWO-SVM model, and

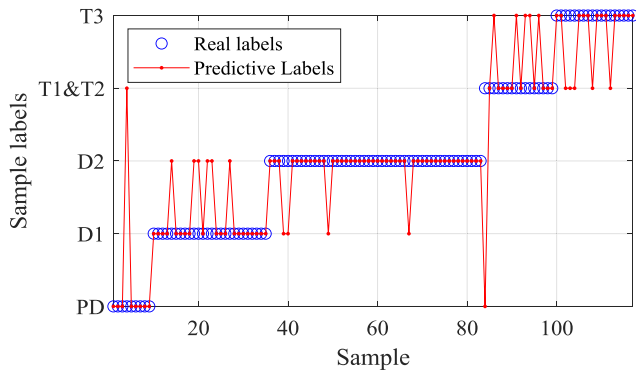


FIGURE 12. Diagnosis results of the MGWO-SVM.

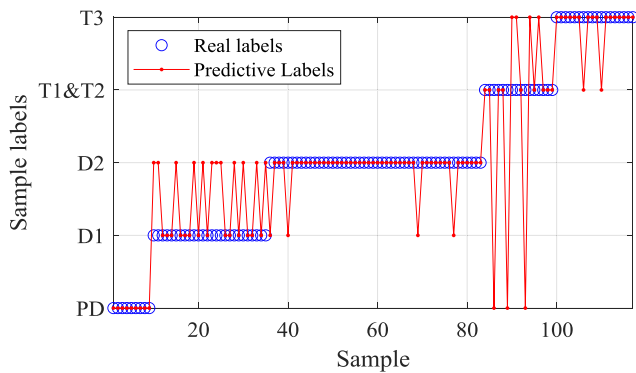


FIGURE 13. Diagnosis results of the PSO-SVM.

TABLE 8. Fault diagnosis accuracy.

Fault diagnosis model	Accuracy
MGWO-KELM	88.89%
PSO-KELM	83.76%
MGWO-SVM	81.2%
PSO-SVM	78.63%
IKHSVM	85.71%

Fig. 13 shows the fault classification result of the PSO-SVM model. The correct fault diagnosis rate of each method is shown in Table 8.

Fig.10 to Fig.13 and Table 8 show that the correct rate of the MGWO-KELM algorithm proposed in this article is 88.89%, which is higher than 83.76% obtained by the PSO-KELM, 81.2% obtained by the MGWO-SVM, and 78.63% obtained by the PSO-SVM. This shows that the hybrid KELM optimized by the modified gray wolf algorithm has a better fault diagnosis performance. The proposed algorithm is further compared with the methods in the literature. In [11], the author established an improved Krill-herd (IKH) algorithm to optimize the SVM transformer fault diagnosis model. Based on the same IEC TC 10 data set, the average testing accuracy of IKHSVM in [11] reaches 85.71%. Compared with the results in [11], the test accuracy of the MGWO-KELM is higher, which verifies the validity of the model.

TABLE 9. MGWO-KELM paired t-test.

Contrast model	$\mu$	$\sigma^2$	$\tau_t$
PSO-KELM	0.0517	0.0014	3.0995
MGWO-SVM	0.0691	0.0007	6.0355
PSO-SVM	0.0996	0.006	2.8714

A paired t-test is used to determine whether there are significant differences between the MGWO-KELM algorithm and the other three algorithms. First, the error rate of the 5-fold cross-validation test set of the different algorithms is calculated. According to Difference  $\Delta_i$ , the t-test is carried out. The mean  $\mu$ , variance  $\sigma^2$  and t-statistic  $\tau_t$  of the differences between the different algorithms are calculated. The results are shown in Table 9. When  $\alpha = 0.05$ , the critical value  $t_{\alpha/2, k-1} = 2.776$ . The t-statistic values  $\tau_t$  of the three-component paired t-tests are all greater than 2.776. This shows that the MGWO-KELM algorithm is significantly better than the compared algorithms.

The MGWO-KELM model is used to simulate and analyze the collected domestic 370 sets of transformer fault data. The diagnosis results and correct rates of the different fault types are shown in Table 10. For T3 faults, 147 groups were correctly diagnosed, and 8 groups had diagnostic errors, including the PD 2 group, D2 3 group, and T1 and T2 3 groups. The correct rate is the highest among all failures, and it is 94.8%. For D1 faults, 29 groups were correctly diagnosed, and the correct rate was 61.7%.

The MGWO-SVM algorithm is used to diagnose the same 370 sets of transformer fault data, and a comparison of the simulation results is shown in Fig. 14. Compared with the SVM, the hybrid KELM algorithm has a significant improvement in the diagnostic accuracy of the five types of faults. The diagnostic accuracy of the MGWO-KELM model for the 370 sets of transformer fault data is 87.3%, which is similar to the results on the IEC TC 10 fault data; this further verifies the reliability and validity of the MGWO-KELM model.

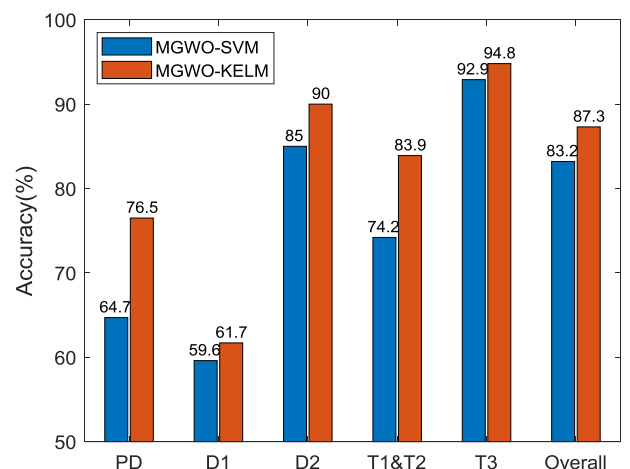


FIGURE 14. Comparison of fault diagnosis accuracy.

TABLE 10. Diagnostic results on the domestic transformer fault data.

Fault type	MGWO-KELM model diagnosis results					Accuracy
	PD	D1	D2	T1&T2	T3	
PD	13	1	1	0	2	76.5%
D1	1	29	12	5	0	61.7%
D2	2	4	108	4	2	90%
T1&T2	1	2	0	26	2	83.9%
T3	2	0	3	3	147	94.8%

## V. CONCLUSION

In this paper, a transformer fault diagnosis model is established based on the KELM algorithm. The model is optimized from two aspects: the hybrid kernel function and the MGWO algorithm, which further improves the model effect. By comparison with other algorithms, the conclusions are as follows:

- (1) The fault diagnosis model based on the hybrid KELM algorithm can accurately and effectively identify the type of transformer fault. Compared with the traditional SVM, it has higher classification accuracy.
- (2) The traditional GWO algorithm tends to fall into local minima or prematurely converge during the optimization process. The DE algorithm is used to improve the GWO algorithm and enhance the search performance of the gray wolf algorithm. Compared to the conventional PSO algorithm, the MGWO algorithm has a stronger search ability and faster convergence speed.
- (3) There are still some shortcomings in the research. For example, when the number of samples is small, the proposed model has difficulty distinguishing D1 and D2 faults, resulting in misclassification. In fact, transformer fault are is also related to other factors, such as voltage levels, insulating oil types [32], oil temperatures, loads, and operating years [33]. In future work, the relationship between multisource fault information and transformer faults will be comprehensively considered to further enhance the accuracy and reliability of transformer fault diagnosis.

## REFERENCES

- [1] J. Faiz and M. Soleimani, "Dissolved gas analysis evaluation in electric power transformers using conventional methods a review," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 24, no. 2, pp. 1239–1248, Apr. 2017, doi: 10.1109/tdei.2017.005959.
- [2] *IEEE Approved Draft Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers*, Standard PC57.104/D6.2, Mar. 2019, pp. 1–103.
- [3] *Mineral Oil-Filled Electrical Equipment in Service—Guidance Interpretation of Dissolved and Free Gases Analysis*, Spanish Assoc. Standardization Certification, Madrid, Spain, 2016.
- [4] *IEEE Draft Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers*, Standard PC57.104/D3.2, Mar. 2008, p. 1.
- [5] M. Duval, "A review of faults detectable by gas-in-oil analysis in transformers," *IEEE Elect. Insul. Mag.*, vol. 18, no. 3, pp. 8–17, May 2002, doi: 10.1109/mei.2002.1014963.
- [6] M. Duval and L. Lamarre, "The duval pentagon new complementary tool for the interpretation of dissolved gas analysis in transformers," *IEEE Electr. Insul. Mag.*, vol. 30, no. 6, pp. 9–12, Dec. 2014, doi: 10.1109/mei.2014.6943428.
- [7] D.-E.-A. Mansour, "A new graphical technique for the interpretation of dissolved gas analysis in power transformers," in *Proc. Annu. Rep. Conf. Electr. Insul. Dielectr. Phenomena*, Oct. 2012, pp. 195–198, doi: 10.1109/CEIDP.2012.6378754.
- [8] D. A. Mansour, "Development of a new graphical technique for dissolved gas analysis in power transformers based on the five combustible gases," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 22, no. 5, pp. 2507–2512, Oct. 2015, doi: 10.1109/Tdei.2015.004999.
- [9] F. R. Barbosa, O. M. Almeida, A. P. S. Braga, M. A. B. Amora, and S. J. M. Cartaxo, "Application of an artificial neural network in the use of physicochemical properties as a low cost proxy of power transformers DGA data," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 19, no. 1, pp. 239–246, Feb. 2012, doi: 10.1109/TDEI.2012.6148524.
- [10] S. M. S. Ghoneim, B. M. T. Ibrahim, and I. E. Nagy, "Integrated ANN-based proactive fault diagnostic scheme for power transformers using dissolved gas analysis," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 23, no. 3, pp. 1838–1845, Jun. 2016, doi: 10.1109/tdei.2016.005301.
- [11] 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.
- [12] A. Abu-Siada and S. Hmood, "A new fuzzy logic approach for consistent interpretation of dissolved gas-in-oil analysis," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 20, no. 6, pp. 2343–2349, Dec. 2013, doi: 10.1109/tdei.2013.6678888.
- [13] A. J. Q. Carita, L. C. Leite, A. P. P. Medeiros, R. Barros, and L. Sauer, "Bayesian networks applied to failure diagnosis in power transformer," *IEEE Latin Amer. Trans.*, vol. 11, no. 4, pp. 1075–1082, Jun. 2013, doi: 10.1109/ta.2013.6601752.
- [14] T. Kari, W. Gao, D. Zhao, Z. Zhang, W. Mo, Y. Wang, and L. Luan, "An integrated method of ANFIS and dempster-shafer theory for fault diagnosis of power transformer," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 25, no. 1, pp. 360–371, Feb. 2018, doi: 10.1109/tdei.2018.006746.
- [15] D. Jiejie, S. Hui, and S. Gehao, "Dissolved gas analysis of insulating oil for power transformer fault diagnosis with deep belief network," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 24, no. 5, pp. 2828–2835, Oct. 2017, doi: 10.1109/tdei.2017.006727.
- [16] G. B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Trans. Syst., Man, Cybern. B. Cybern.*, vol. 42, no. 2, pp. 513–529, Apr. 2012, doi: 10.1109/TSMCB.2011.2168604.
- [17] Y. Q. Chen, O. Fink, and G. Sansavini, "Combined fault location and classification for power transmission lines fault diagnosis with integrated feature extraction," *IEEE Trans. Ind. Electron.*, vol. 65, no. 1, pp. 561–569, Jan. 2018, doi: 10.1109/tie.2017.2721922.
- [18] Q. Tong, J. Cao, B. Han, X. Zhang, Z. Nie, J. Wang, Y. Lin, and W. Zhang, "A fault diagnosis approach for rolling element bearings based on RSGWPT-LCD bilayer screening and extreme learning machine," *IEEE Access*, vol. 5, pp. 5515–5530, 2017, doi: 10.1109/Access.2017.2675940.
- [19] S. S. Udmale and S. K. Singh, "Application of spectral kurtosis and improved extreme learning machine for bearing fault classification," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 11, pp. 4222–4233, Nov. 2019, doi: 10.1109/Tim.2018.2890329.
- [20] F. Ding, X. Zhang, W. Wu, and Y. Wang, "A fault feature extraction method of motor bearing using improved LCD," *IEEE Access*, vol. 8, pp. 220973–220979, 2020, doi: 10.1109/access.2020.3043803.

- [21] G. Sheng, Q. Hong, and M. Wei, "Research on ELM soft fault diagnosis of analog circuit based on KSLPP feature extraction," *IEEE Access*, vol. 7, pp. 92517–92527, 2019, doi: [10.1109/access.2019.2923242](https://doi.org/10.1109/access.2019.2923242).
- [22] L. Wang, F. Lyu, Y. Su, and J. Yue, "Kernel entropy-based classification approach for superbuck converter circuit fault diagnosis," *IEEE Access*, vol. 6, pp. 45504–45514, 2018, doi: [10.1109/Access.2018.2864138](https://doi.org/10.1109/Access.2018.2864138).
- [23] H. Malik and S. Mishra, "Extreme learning machine based fault diagnosis of power transformer using IEC TC10 and its related data," in *Proc. Annu. IEEE India Conf. (INDICON)*, Dec. 2015, pp. 1–5, doi: [10.1109/INDICON.2015.7443245](https://doi.org/10.1109/INDICON.2015.7443245).
- [24] S. Li, G. Wu, B. Gao, C. Hao, D. Xin, and X. Yin, "Interpretation of DGA for transformer fault diagnosis with complementary SaE-ELM and arctangent transform," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 23, no. 1, pp. 586–595, Feb. 2016, doi: [10.1109/tdei.2015.005410](https://doi.org/10.1109/tdei.2015.005410).
- [25] W. Huang, Z. Huang, S. Long, and Y. Long, "Multiscale information fusion-based extreme learning machine for power transformer fault diagnosis," in *Proc. IEEE 3rd Conf. Energy Internet Energy Syst. Integr. (EI2)*, Nov. 2019, pp. 2182–2186, doi: [10.1109/EI247390.2019.9061786](https://doi.org/10.1109/EI247390.2019.9061786).
- [26] A. K. Qin, V. L. Huang, and P. N. Suganthan, "Differential evolution algorithm with strategy adaptation for global numerical optimization," *IEEE Trans. Evol. Comput.*, vol. 13, no. 2, pp. 398–417, Apr. 2009, doi: [10.1109/tevc.2008.927706](https://doi.org/10.1109/tevc.2008.927706).
- [27] T. Liu, H. Zhu, M. Wu, and W. Zhang, "Rotor displacement self-sensing method for six-pole radial hybrid magnetic bearing using mixed-kernel fuzzy support vector machine," *IEEE Trans. Appl. Supercond.*, vol. 30, no. 4, Jun. 2020, Art. no. 3603404, doi: [10.1109/tasc.2020.2990415](https://doi.org/10.1109/tasc.2020.2990415).
- [28] L. Sun, R. Liu, J. Xu, S. Zhang, and Y. Tian, "An affinity propagation clustering method using hybrid Kernel function with LLE," *IEEE Access*, vol. 6, pp. 68892–68909, 2018, doi: [10.1109/access.2018.2880271](https://doi.org/10.1109/access.2018.2880271).
- [29] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, Mar. 2014, doi: [10.1016/j.advengsoft.2013.12.007](https://doi.org/10.1016/j.advengsoft.2013.12.007).
- [30] X.-S. Yang, "Test problems in optimization," *Mathematics*, vol. 2, no. 2, pp. 63–86, 2010.
- [31] J. G. Digenakis and K. G. Margaritis, "On benchmarking functions for genetic algorithms," *Int. J. Comput. Math.*, vol. 77, no. 4, pp. 481–506, Sep. 2001.
- [32] C. Xiang, Q. Zhou, J. Li, Q. Huang, H. Song, and Z. Zhang, "Comparison of dissolved gases in mineral and vegetable insulating oils under typical electrical and thermal faults," *Energies*, vol. 9, no. 5, p. 312, Apr. 2016.
- [33] H. Faria, Jr., J. G. S. Costa, and J. L. M. Olivas, "A review of monitoring methods for predictive maintenance of electric power transformers based on dissolved gas analysis," *Renew. Sustain Energy Rev.*, vol. 46, pp. 201–209, Jun. 2015.



fault diagnosis of smart grid.

**JIFANG LI** was born in Kaifeng, Henan, China, in 1971. He received the M.S. degree in control technology from 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



**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.



**ZHEN FENG** was born in Xianning, Hunan, China, in 1995. She is currently pursuing the master's degree with the North China University of Water Resources and Electric Power. Her current research interests include power electronics and power transmission.



**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.

• • •