A Power Transformer Fault Diagnosis Method-Based Hybrid Improved Seagull Optimization Algorithm and Support Vector Machine

YUHAN WU, (Student Member, IEEE), XIANBO SUN, YI ZHANG, XIANJING ZHONG, AND LEI CHENG
School of Information Engineering, Hubei Minzu University, Enshi, Hubei 445000, China
Corresponding author: Xianbo Sun (sunxianbo@whut.edu.cn)
This work was supported in part by the National Natural Science Foundation of China under Grant 61661020.

ABSTRACT The method of Support Vector Machine (SVM) based on Dissolved Gas Analysis (DGA) has been studied in the field of power transformer fault diagnosis. However, there are still some shortcomings, such as the fuzzy boundaries of DGA data, and SVM parameters are difficult to determine. Therefore, this paper proposes a power transformer fault diagnosis method based on Kernel Principal Component Analysis (KPCA) and a hybrid improved Seagull Optimization Algorithm to optimize the SVM (TISOA-SVM). Firstly, KPCA is used to extract features from DGA feature quantities. In addition, TISOA is further proposed to optimize the SVM parameters to build the optimal diagnosis model based on SVM. For the SOA, three improvement methods are proposed. An improved tent map is used to replace the original population initialization to improve population diversity. In addition, the nonlinear inertia weight and random double helix formula are proposed to improve the optimization accuracy and efficiency of the SOA. Then, benchmarking functions are used to test the optimization performance of TISOA and six algorithms, and the results show that TISOA has the best optimization accuracy and convergence speed. Finally, the fault diagnosis method based on KPCA and TISOA-SVM is obtained, and it is noteworthy that three examples are tested to verify the diagnostic performance of the proposed method. These results show that the proposed method has higher diagnostic accuracy, shorter diagnosis time, stronger significance and validity than other methods. Therefore, a research idea is provided for solving practical engineering problems in the field of fault diagnosis.

INDEX TERMS Power transformer, fault diagnosis, kernel principal component analysis, support vector machine, hybrid improved seagull optimization algorithm.

I. INTRODUCTION
The power transformer is one of the most important pieces of equipment in the power grid, and the failure of a power transformer will affect the stable operation of a power grid, resulting in very large economic losses and even endangering people’s lives [1], [2]. On June 19, 2017, a transformer explosion accident, which caused great economic losses, one death and seven passer-by injuries, occurred in Yongzhou City, Hunan Province, China [3]. Thus, it is crucial to perform fault diagnosis of power transformers quickly and accurately to ensure safe and stable operation [4].

In the field of power transformer fault diagnosis, dissolved gas analysis (DGA) is the most versatile method [5]. There are several traditional DGA methods: the three-ratio method, Rogers ratio method, and Duval triangle method [6]–[9]. Nevertheless, the above methods mostly rely on manual experience and historical data. They encounter defects, such as incomplete coding and coding boundaries, so these methods cannot perform fault diagnosis of power transformers quickly and accurately. With the development of artificial intelligence, many new methods, such as artificial neural network (ANN) [10], [11], support vector machine (SVM) [12], [13],
and fuzzy theory [14], [15], have been proposed and applied in the field of power transformer fault diagnosis.

Zhou et al. proposed a transformer fault diagnosis model based on improved gray wolf optimization (IGWO) and probabilistic neural network (PNN). Compared with the simulation results of eleven models, the stability of the proposed model was effectively proven, and the accuracy of transformer fault diagnosis was improved significantly. However, the authors did not study the initialization of GWO. The poor initial position of the population will affect the optimization performance. Therefore, only improving the update strategy cannot stably improve the GWO optimization performance [16]. Kim et al. proposed a Semisupervised Autoencoder with an Auxiliary Task (SAAT) for power transformer fault diagnosis using DGA. The results show that the fault diagnosis method based on SAAT has good diagnostic performance. The proposed health feature space (HFS) contributes to the intuitive monitoring of the health state of transformers. The above methods only reflect the health of transformers and cannot identify specific fault types, so the proposed method cannot be applied to accurate fault identification. [17]. Tahir et al. produced an intelligent monitoring and classification algorithm for detecting transformer winding faults based on frequency response analysis (FRA). Then, four cases were used to analyze the proposed method. The model can accurately identify various winding states in the transformer. However, they did not study the disadvantage that the parameters of ANN are difficult to determine, and the diagnostic accuracy of the proposed model cannot be optimized only by using ANN [18]. Yuan et al. proposed a transformer fault diagnosis model based on chemical reaction optimization (CRO) and twin support vector machines (TWSVMs). The results show that the model not only ensures high simulation accuracy but also has fast diagnosis efficiency. However, CRO has the defects of precocity and poor optimization accuracy. They have not studied the above shortcomings, and the results have not been compared with recent research methods. Meanwhile, compared with the traditional diagnosis model, the improvement of diagnostic accuracy of the proposed model is not high [19].

By summarizing the above research methods from literature [10]–[20], ANN needs a large number of sample data, and the validity of the model is not strong, and the stability and robustness of fuzzy theory are poor, compared with ANN and fuzzy theory, SVM, which is suitable for nonlinear and small sample data, is not only more effective, and but also is more widely used in the field of fault diagnosis [20]. Thus, this paper selects SVM as the basic model. SVM is a classification algorithm with a perfect mathematical model based on statistical learning theory. SVM has strong generalization, but determining parameters is still difficult [21], [22]. Therefore, many optimization algorithms, such as whale optimization algorithm (WOA) [23], differential evolution (DE) [24], particle swarm optimization (PSO) [25], and GWO [26], have been successively applied to SVM transformer fault diagnosis models to find the optimal parameters.

Power transformer fault diagnosis based on SVM has been studied by many scholars at home and abroad. Zeng et al. proposed a transformer fault diagnosis model based on the IGWO-least squares SVM (LSSVM) model and used kernel principal component analysis (KPCA) to extract features from mixed DGA data. The results show that the model has high diagnostic performance. Nevertheless, DE has some defects, such as slow convergence and ease of falling into local optimization, and the authors did not research them. Thus, the performance improvement in the diagnosis model is not distinct [27]. Li et al. proposed a transformer fault diagnosis model based on an improved sparrow search algorithm (ISSA)-SVM, and the constructed model can accurately diagnose transformer faults. However, the population diversity of SSA is poor, and traditional DGA data have noise, and the authors have not further studied the above problems [28]. Benmahamed et al. proposed a transformer fault diagnosis system using Gaussian classification and bat algorithm (BA) to optimize SVM, and gas concentration and gas percentage were used as two inputs to diagnose faults. The results show that the diagnostic accuracy of BA-SVM reached 93.75%. However, the traditional BA, which easily falls into local optima, has poor population diversity, so it cannot find the optimal parameters of the SVM. If the authors can study the population behavior to improve BA, the diagnostic performance of proposed model will be better [29]. Zhang et al. used a genetic algorithm (GA) combined with an SVM to binary code the DGA feature quantity, and an improved krill optimization algorithm (IKH) was used to optimize the SVM parameters, then, the fault diagnosis model based on IKHSVM. The results show that the IKHSVM can improve the accuracy of transformer fault diagnosis. However, the authors only added adaptive random disturbance to the krill population, and they should also improve the initialization of the KH population, which can increase the optimization performance to improve the stability of the diagnostic model [30].

Three shortcomings of transformer fault diagnosis based on SVM are summarized: 1) a single fault diagnosis model cannot greatly improve the fault diagnosis performance; 2) the noise of transformer fault data will reduce the stability of the model; 3) research on optimization algorithms is not targeted and cannot significantly improve the optimization performance. Thus, a transformer fault diagnosis method based on the KPCA and TISOA-SVM method is proposed in this paper. It is noteworthy that the innovations and contributions of this paper are mainly divided into the following four improved methods. First, KPCA is used to extract the features of DGA data to reduce the influence of noise on the diagnosis results. In addition, SOA can be improved by the following three methods to obtain the TISOA. A modified tent map (MTent) is proposed to improve the initial diversity of the seagull population. Considering the influence of the inertia weight of SOA on population convergence and the optimization mode, a nonlinear inertia weight method, which can effectively improve the convergence speed of the SOA, is proposed, and a random double helix foraging formula is
proposed to improve the optimization accuracy of the SOA. Then, it is noteworthy that TISOA can be obtained from the above three improved methods, and the benchmark functions are used to test the optimization performance of TISOA and the other six algorithms. The results show that TISOA has the best optimization performance. Finally, three examples were used to test the diagnostic method of the proposed method. The results show that the proposed method has the highest diagnostic accuracy, the shortest diagnosis time, the best validity, and the strongest significance. These results can prove that the proposed method has the best diagnostic performance.

The rest of this paper is as follows. Section II introduces the basic theory of KPCA. Section III introduces the basic theory of SVM and SOA and the improvement methods of SOA. Meanwhile, the optimization performance of TISOA is tested. In section IV, the diagnosis model proposed in this paper is described. In section V, the diagnosis results are analyzed by three examples. In section VI, the research conclusions and shortcomings of this paper are summarized, and future research directions are proposed.

II. BASIC THEORY OF KPCA

KPCA is a feature extraction method based on principal component analysis (PCA) [31]. KPCA solves the disadvantage that PCA can extract only the features of linear data by introducing a kernel function. KPCA extracts the features of nonlinear data and increases the dimensionality of the data introducing a kernel function. KPCA extracts the features of that PCA can extract only the features of linear data by

The above derivation is the case where the mean value of the mapping data is zero. However, most cases are not consistent, and K will be effectively transformed.

$$\overline{K} = K - NK - KN + NKN$$ (6)

where N is an N-order matrix and N_{i,j} is 1/N. Eq. (6) ensures that the assumption that the mean value of the mapping data is 0 is true.

III. TISOA-SVM

A. SVM

SVM has a strong classification ability [34]. The principle is to construct the optimal hyperplane in space and classify the sample data. The kernel function is used to map x_i to the high-dimensional space \( \phi(x_i) \) to construct the optimal hyperplane. The model is as follows:

$$f(x) = sign(\omega^T \phi(x_i) + b)$$ (7)

where \( \omega \) is weight vector; \( b \) is deviation. The relaxation variable \( \xi_i \) is introduced to construct the maximum interval classifier.

$$\begin{align*}
\min & \quad \frac{1}{2} \| \omega \|^2 + C \sum_{i=1}^{N} \xi_i \\
\text{s.t.} & \quad y_i(\omega^T x_i + b) \geq 1 - \xi_i
\end{align*}$$ (8)

where \( \xi_i \geq 0 \); \( C \) is a penalty factor that can balance \( \xi_i \). The Lagrange multiplier \( \lambda_i \) is introduced to turn the original problem with constraints into a simple pair problem.

$$\begin{align*}
\min & \quad \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j y_i y_j \phi^T(x_i) \phi(x_j) - \sum_{i=1}^{N} \lambda_i \\
\text{s.t.} & \quad \lambda_i \geq 0
\end{align*}$$ (9)

Then, Eq. (10) can be obtained according to Karush-Kuhn-Tucker (KKT) conditions.

$$\begin{align*}
\frac{\partial f}{\partial \omega} = 0; \quad \frac{\partial f}{\partial b} = 0; \quad \frac{\partial f}{\partial \lambda} = 0 \\
\lambda_i[1 - y_i(\omega^T x_i + b) - \xi_i] = 0 \\
1 - y_i(\omega^T x_i + b) - \xi_i \leq 0
\end{align*}$$ (10)

where in the existing part, \( \lambda_i \neq 0 \), and solving Eq. (10) can obtain the best \( \omega \) and \( b \), which are brought into Eq. (7) to obtain the optimal decision model.

$$f(x) = sign(\sum_{i=1}^{N} \lambda_i y_i K(x_i, x_j) + b)$$ (11)

where \( K(x_i, x_j) = \phi^T(x_i) \phi(x_j) \) is the kernel function. In this paper, the radial basis function (RBF) is selected as the kernel function of SVM and KPCA. RBF has only one parameter, and the performance of processing nonlinear data is good. The expression is as follows:

$$K(x_i, x_j) = \exp \left\{ -\frac{\| x_i - x_j \|^2}{2\sigma^2} \right\}$$ (12)
where $\sigma$ is kernel function parameter. To establish the transformer fault diagnosis model, TISOA is used to find the parameters $C$ and $\sigma$, and KPCA is used to process the data.

**B. SOA**

SOA is a new metaheuristic optimization algorithm proposed by Gaurav Dhiman and Vijay Kumar in 2018 [35], [36]. The basic principle of SOA is to construct optimization rules by studying the migration and foraging behaviors of seagull populations. Migratory behavior refers to the migration of seagull populations from one place of residence to the next; foraging behavior refers to the behavior of seagull population predation in their residence. The optimization strategy of SOA is as follows and the optimization process is shown in step (a) of Figure 4.

1) **MIGRATION**

The migration behavior of seagulls should meet three conditions, namely, avoiding collisions, moving toward the direction of the best neighbor, and remaining close to the best search agent.

1) Avoiding the collisions. To avoid collisions among seagull populations, variable $A$ is used by SOA to adjust the position of seagulls.

$$\vec{C}_s = A \times \vec{P}_S(t)$$

(13)

where $\vec{C}_s$ represents a new location different from other seagulls; $\vec{P}_S$ represents the current location of the seagull; and $A$ represents the seagull’s movement behavior in a given search space.

$$A = f_c - f_c(t/T_{\text{max}})$$

(14)

where the value of $f_c$ is 2; $t$ is the current number of iterations; and $T_{\text{max}}$ is the maximum number of iterations.

2) Movement toward the best neighbor’s direction. After ensuring that there was no collision between seagulls, all seagulls moved toward the best position.

$$\vec{M}_s = B \times (\vec{P}_{\text{best}}(t) - \vec{P}_s(t))$$

(15)

where $\vec{M}_s$ indicates that the seagull population is moving toward $\vec{P}_{\text{best}}$, the best positioned seagull. $B$ is a random number and has the ability to explore balanced algorithm.

$$B = 2 \times A^2 \times \text{rand}$$

(16)

3) Remain close to the best search agent. After calculating the convergence direction of each seagull, the seagull begins to move.

$$\vec{D}_s = \left| \vec{C}_s + \vec{M}_s \right|$$

(17)

where $\vec{D}_s$ represents the new position of the seagull after moving.

2) **ATTACKING**

The seagull population engage in foraging behavior according to experience gained from migration behavior. The seagull population constantly changes its attack angles and flight speeds during migration. Seagulls spiral in the air to attack prey.

$$\begin{cases} x' = r \times \cos(k) \\ y' = r \times \sin(k) \\ z' = r \times k \\ r = u \times e^{kv} \end{cases}$$

(18)

where $r$ is the helix radius, which is controlled by $u$ and $v$. $u$ and $v$ are the correlation constants of the helix shape, and $k$ is the random angle between $[0, 2\pi]$. Combined with the new location of seagulls, the location update formula of the seagull population is obtained.

$$\vec{P}_s'(t) = (\vec{D}_s \times x \times y \times z) + \vec{P}_{\text{best}}(t)$$

(19)

where $\vec{P}_s'(t)$ is the location of updated seagull population.

**C. IMPROVING METHODS OF SOA**

SOA has strong optimization performance. However, it still has some defects. For example, it has poor location diversity of the initial population and easily becomes stuck in local optima. Aiming at the above shortcomings, this paper proposes the following three improvement methods to improve SOA: MTent, nonlinear inertia weight and the random double helix foraging formula.

1) **MTENT**

This paper uses a tent map to improve the defects of low ergodicity and weak randomness of the initialized population of SOA. Tent map is a traditional chaotic mapping method [37].

$$x_{n+1} = \begin{cases} 2x_n, & x_n \in [0, 0.5] \\ 2(1 - x_n), & x_n \in [0.5, 1] \end{cases}$$

(20)
However, the existence of small periodic points will make it jump out of the chaotic state [38]. For example, when the initial value $x_0 = \{0.2, 0.4, 0.6, 0.8\}$, there are small periodic points. Furthermore, there are unstable periodic points, such as $x_n = x_{n-i}, i = \{0, 1, 2, 3, 4\}$. Therefore, MTent is used to replace the original initialization method of SOA to enhance the population diversity. The formula is as follows:

$$x_{n+1} = \begin{cases} 2(x_n + \text{rand}()/20), & x_n \in [0, 0.5) \\ 2(1 - (x_n + \text{rand}()/20)), & x_n \in [0.5, 1] \end{cases} \quad (21)$$

In Eq. (21), rand() is a random number between $[0, 1]$. When the chaotic sequence is generated, $x_n$ will jump out of the periodic point and enter a chaotic state again because the random disturbance formula $\text{rand}()/20$. The bifurcation diagram of MTent is shown in Figure 2.

Figure 3 shows the population distribution in $(0,1)$ space generated by MTent and Tent. With MTent, seagulls can jump out of the boundary due to the random disturbance formula, and all seagulls that jump out of the boundary are counted as being on the boundary.

As shown in Figure 3, by comparing two methods, the results can be seen that most of the populations generated by the tent map are clustered at the boundary and have poor ergodicity and randomness. In contrast, the initial population generated by MTent has stronger ergodicity and randomness, and the generated population has better diversity.

### 2) NONLINEAR INERTIA WEIGHT

Migration behavior is an important part of the optimization strategy of SOA. Inertia weight $A$ plays a key role in the process of migration behavior. The size of $A$, which can avoid the collision of seagulls during migration, will affect the global search behavior of SOA.

In this section, a nonlinear formula is proposed to improve the inertia weight $A$ of the SOA to make the seagull population behavior update more flexible and to increase the optimization efficiency.

$$A = 2 \cdot \cos(\frac{\pi}{2} \cdot x + \frac{\pi}{2}) + 2 \quad (22)$$

The cosine function in Eq. (22), which can make the attenuation of the inertia weight smoother, is selected to replace Eq. (14). Figure 4 shows the function image of the inertia weight before and after improvement.

As shown in Figure 4, when the iteration ratio of the improved inertia weight reaches 0.33, the value of $A$ is equal to 1 and decays slowly in the subsequent iteration. The population starts a local search to ensure that the optimization accuracy and optimization efficiency of the SOA are effectively improved.

### 3) THE RANDOM DOUBLE HELIX FORAGING FORMULA

In this paper, the foraging behavior of SOA is studied and inspired by the bubble-net attacking method of WOA. A double helix formula with randomness is proposed to improve the foraging formula of SOA. The basic theory of WOA is as follows.

#### a: WOA

The WOA is a metaheuristic algorithm to simulate the predation behavior of humpback whales and was proposed in 2016 by Mirjalili [39]. WOA has the advantages of a simple structure and few adjustment parameters.

#### i) ENCIRCLING PREY

All whales in the population swim near the optimal whale in the population, causing the whale population to conduct local optimization.
ii) SEARCH FOR PREY
The whales randomly select a whale in the population and swim near it to conduct global optimization. The location update formula is as follows:

\[ B = a \cdot (2 \cdot \text{rand} - 1) \]
\[ C = 2 \cdot \text{rand} \]
\[ D = |C \cdot X_{\text{rand}} - X| \]
\[ X(t + 1) = X_{\text{rand}} - B \cdot D \]

where \( a \) is a linear function that decreases from 2 to 0 monotonically.

iii) BUBBLE-NET ATTACKING METHOD
Bubble-net predation is used to surround the prey by simulating the predation behavior of humpback whales. The location update formula is as follows:

\[ X(t + 1) = \begin{cases} 
  D \cdot e^{bl} \cdot \cos (2\pi t) + X, & p \geq 0.5 \\
  X - B \cdot D, & p < 0.5 
\end{cases} \]

where \( b \) is a constant that defines the shape of the spiral; \( l \) is a random number between \([-1, 1]\); and \( p \) is a random number between \([0, 1]\). When \( p < 0.5 \), WOA surrounds prey or searches prey. When \( p \geq 0.5 \), WOA carries out bubble-net attacking.

b: THE RANDOM DOUBLE HELIX FORAGING FORMULA
The variable parameter \( A \) in WOA is introduced to control the search state, the seagull position is randomly changed in the SOA iteration process, and the control parameter \( p \) is introduced to control the optimization method. When \( |A| > 1 \) and the control parameter is \( p < 0.5 \), SOA will conduct a random search.

\[ D = \left| C \cdot X_{\text{rand}} - \tilde{P}_s(t) \right| \]
\[ \tilde{P}_s(t + 1) = \overrightarrow{X_{\text{rand}}} - B \cdot \overrightarrow{D} \]

where \( \overrightarrow{D} \) represents the new position of the seagull after it has moved and \( \overrightarrow{X_{\text{rand}}} \) represents a random seagull in a seagull population. Finally, the double helix formula is constructed by combining Eq. (27)

\[ \tilde{P}_s(t + 1) = \begin{cases} 
  \overrightarrow{D} + e^{bl} \cdot \cos (2\pi t) \cdot \overrightarrow{P}_{\text{best}}(t), & p \geq 0.5 \\
  \overrightarrow{D} \times \overrightarrow{x} \times \overrightarrow{y} \times \overrightarrow{z} + \overrightarrow{P}_{\text{best}}(t), & p < 0.5 
\end{cases} \]

When \( |A| > 1 \), the formula of \( p < 0.5 \) is replaced with Eq. (29). In this paper, the double helix with randomness in Eqs. (29) and (30) is used to improve the foraging behavior of SOA in Eq. (19), the diversity of the seagull population is increased, and the optimization accuracy of SOA is improved.

D. TISOA
1) SUMMARY OF IMPROVEMENT METHODS
The improved method proposed in Part C of Section III is used to improve SOA from three aspects to obtain TISOA, the corresponding improvement objectives are shown in step (b) of Figure 4.

1) M谭澎 is used to initialize the seagull population, which can replace the original random initialization of SOA, and it can increase the diversity of the seagull population and greatly enhance ergodicity and randomness.

2) Under the condition of good population diversity, the inertia weight \( A \) is improved obtain the improve formula (22). The method is increasing its attenuation speed to 1, reduce the migration behavior of the seagull population, and increase the foraging behavior of the seagull population.

3) To greatly improve the optimization accuracy of SOA, based on the previous two improvements, a double helix formula with randomness is proposed to improve the foraging behavior of SOA to replace the original attacking. Significantly, this is an innovative improvement method, which can comprehensively improve the convergence speed and optimization accuracy of SOA.

To verify the validity and superiority of these three improved methods, the benchmark functions have been used to test the optimization performance of TISOA in Part E of Section III.

2) OPTIMIZATION PROCESS OF TISOA
In this section, the optimization process of TISOA is introduced in detail, and Figure 5 is the specific flow chart.

Step 1: parameter initialization. Initialize the algorithm parameters of the SOA and WOA, and set the population size \( M \), the maximum number of iterations \( T \), the dimension dim, \( F_c, A, a, l \) and \( B \);
Step 2: population initialization. MTent is used to initialize the seagull population and randomly generate the individual seagull population;

Step 3: fitness calculation. The initial fitness values of the seagull population and the individuals are calculated and compared to find the optimal individual;

Step 4: migration behavior. SOA performs the migration behavior using Eqs. (13), (15) and (17). Then, the linear inertia weight \( A \) (Eq. (14)) which controls the motion behavior, is replaced by Eq. (22);

Step 5: foraging behavior. The double helix with randomness in Eqs. (29) and (30) is used instead of the original foraging behavior in Eq. (19); SOA exhibits foraging behavior with strong randomness and high accuracy, and \( B \) is Eq. (23);

Step 6: population selection. After the SOA iteration, population selection is carried out, and the population positions and individual positions with the best fitness are taken as the next generation fitness values;

### TABLE 1. Parameters of algorithms.

| Algorithms     | Parameters | Values   |
|----------------|------------|----------|
| All Algorithms | Population | 50       |
|                | \( T_{\text{max}} \) | 1000     |
| SOA           | \( f_e \) | 2        |
| GWO           | \( A \)  | 2 \( \rightarrow \) 0 |
| PSO           | \( V_{\text{max}} \) | 6        |
|               | \( c_1, c_2 \) | 2        |
| SFO           | \( A \)  | 4        |
|               | \( e \)   | 0.001    |
|               | Percent of SF | 30%     |
| WOA           | \( a \)  | 2 \( \rightarrow \) 0 |
|               | \( b \)  | 1        |
|               | \( I \)  | \([-1, 1]\) |
|               | \( p \)  | \([0, 1]\) |
| ISOA and TISOA| \( f_e \) | 2        |
|               | \( A, a \) | 2 \( \rightarrow \) 0 |
|               | \( b \)  | 1        |
|               | \( p \)  | \([0, 1]\) |
Step 7: judgement. Judge whether the termination conditions are met. If yes, the algorithm ends; otherwise, go to step 4.

**E. OPTIMIZATION PERFORMANCE TEST OF TISOA**

In this paper, 24 benchmark functions are used to test the optimization performance of TISOA, GWO, PSO, sailed fish optimizer (SFO), WOA, SOA, and ISOA. The benchmark functions are unimodal benchmark test functions, multimodal benchmark test functions, fixed-dimension multimodal benchmark test functions, and CEC2015 benchmark functions. ISOA improves the nonlinear inertia weight and double helix formula with fault diagnosis.

The parameters of the above algorithms are shown in Table 1. All tests in this paper are realized in MATLAB 2020a. The operating environment is Intel (R) core (TM) i7–9700 CUP, 16.0 GB memory and 64-bit operating system. All algorithms are tested 20 times.
1) TRADITIONAL BENCHMARK FUNCTION

In this section, nine benchmark functions are selected to test the above algorithms, including three unimodal functions \( F_1, F_2 \) and \( F_3 \); three multimodal functions \( F_4, F_5 \) and \( F_6 \); and three fixed dimensional multimodal functions \( F_7, F_8 \) and \( F_9 \). The specific formula is shown in Appendix 1. Figure 6 is a three-dimensional graph of nine benchmark functions. Table 2 shows the simulation results of seven algorithms, in which the best value (Best), average value (Ave), and standard deviation (Std) of the fitness values are recorded as evaluation items.

In Table 1, the best value, average value and standard deviation of the function test results are in bold. For unimodal benchmark test functions \( F_1, F_2 \) and \( F_3 \), the test results of TISOA are better than the results of the other algorithms, in which the found optimal fitness value is 95 orders of magnitude higher than that of the traditional SOA and dozens of orders of magnitude higher than those of the other algorithms, and the performance of TISOA is obviously better than that of the other algorithms. From the results of multimodal benchmark test functions \( F_4, F_5 \) and \( F_6 \), it can be concluded that the results of ISOA and TISOA are the best compared with the results of the other algorithms, but TISOA, ISOA, GWO, WOA, and SOA can find the minimum fitness value in \( F_4 \) and \( F_6 \).

In the fixed-dimension multimodal benchmark test function \( F_7 \), the three evaluation terms obtained by TISOA are the best; in \( F_8 \) and \( F_9 \), GWO, WOA, SOA, and TISOA can find the best evaluation items. From the test results for the above benchmark functions, compared with other algorithms, the three evaluation items obtained from the TISOA test are optimal, which is enough to prove that TISOA has strong optimization performance and stability. However, there are also several benchmark functions in which TISOA and other algorithms can find optimal results. Therefore, the specific test results are shown in Figure 7 to further compare and analyze the optimization performance of these algorithms. For unimodal benchmark test functions \( F_1, F_2 \) and \( F_3 \), compared with the other algorithms, TISOA not only finds the optimal value of the function but also improves the convergence speed. For \( F_3 \), TISOA finds the optimal value only when it reaches 100 iterations. For multimodal benchmark test functions \( F_4, F_5 \) and \( F_6 \), TISOA and ISOA can stably find the minimum values of the functions, GWO and SOA also have the ability to find the minimum values of the functions, and the convergence speed of TISOA is significantly higher than the other algorithms.

For \( F_4 \) and \( F_6 \), TISOA converges quickly at the beginning of iteration and finds the minimum fitness values. For the fixed-dimension multimodal test functions \( F_7 \) and \( F_9 \), TISOA has the fastest convergence speed, and the minimum values of

### Table 2. The benchmark functions test results.

| Function | \( F_1 \) | \( F_2 \) | \( F_3 \) | \( F_4 \) | \( F_5 \) | \( F_6 \) | \( F_7 \) | \( F_8 \) | \( F_9 \) |
|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| **Best** | 7.92e-43 | 2.711e+1 | 5.12e-4  | 0        | 1.51e+14 | 0        | 3.07e-4  | 3        | -10.54   |
| GWO      | 9.271e-41| 3.143e+1 | 1.013e-3 | 1.645e-14| 1.688e-14| 8.8e-2   | 1.791e-3 | 3        | -10.54   |
| **Std**  | 8.319e-41| 3.718    | 9.834e-4 | 2.649e-14| 3.287e-15| 1.16e-2  | 8.377e-4 | 0        | 0        |
| **Best** | 1.283e-6 | 2.586e+1 | 1.603e-2 | 3.681e+1 | 1.115e-5 | 7.396e-3 | 3.379e-4 | 3        | -10.54   |
| PSO      | 3.469e-5 | 2.962e+1 | 8.672e-2 | 3.881e+1 | 1.304e-5 | 1.6e-2   | 7.685e-4 | 3        | -8.2389  |
| **Std**  | 5.662e-5 | 1.965    | 5.713e-2 | 1.614    | 2.849e-6 | 1.39e-2  | 3.853e-4 | 0        | 8.237    |
| **Best** | 5.878e-6 | 3.016e-3 | 4.362e-4 | 4.212e-11| 2.021e-6 | 8.725e-13| 3.162e-4 | 3        | 0.002    |
| SFO      | 4.725e-5 | 6.465e-3 | 6.525e-4 | 5.795e-10| 2.521e-6 | 2.112e-12| 3.651e-4 | 3        | -10.54   |
| **Std**  | 3.053e-5 | 7.667e-3 | 3.523e-4 | 4.485e-10| 3.674e-7 | 2.057e-12| 4.357e-5 | 1.210    | 2.332    |
| **Best** | 2.186e-36| 7.923    | 2.127e-3 | 0        | 4.441e-15| 0        | 3.075e-4 | 3        | -10.54   |
| WOA      | 8.164e-36| 1.137e+1 | 8.801e-3 | 9.173e-15| 7.639e-15| 9.1e-3   | 5.149e-4 | 3        | -10.54   |
| **Std**  | 9.992e-36| 3.112    | 7.512e-3 | 1.282e-14| 2.623e-15| 1.83e-2  | 1.817e-4 | 0        | 0        |
| **Best** | 2.079e-3 | 1.693e+3 | 2.786e-2 | 1.421e-4 | 9.5e-7   | 9.602e-7 | 8.335e-3 | 3        | -10.54   |
| SOA      | 4.331e-2 | 2.347e+3 | 9.477e-2 | 4.987e-3 | 6.965e-5 | 1.1176   | 1.23e-2  | 3.172    | -9.4341  |
| **Std**  | 4.860e-2 | 5.456e+2 | 9.231e-2 | 9.503e-3 | 1.313e-4 | 1.356e-1 | 9.6e-3   | 2.659e-1 | 8.9e-1   |
| **Best** | 2.506e-79| 1.634e-4 | 1.612e-4 | 0        | 8.882e-16| 0        | 3.075e-4 | 3        | -10.54   |
| ISOA     | 7.314e-78| 6.861e-3 | 7.514e-4 | 0        | 8.882e-16| 0        | 7.213e-4 | 3        | -10.54   |
| **Std**  | 6.295e-78| 6.559e-3 | 5.686e-4 | 0        | 0        | 8.232e-5 | 0        | 0        |
| **Best** | 4.624e-98| 1.129e-6 | 3.493e-6 | 0        | 8.882e-16| 0        | 3.075e-4 | 3        | -10.54   |
| TISOA    | 3.229e-97| 2.354e-6 | 5.479e-6 | 0        | 8.882e-16| 0        | 3.141e-4 | 3        | -10.54   |
| **Std**  | 2.136e-97| 2.198e-6 | 2.009e-6 | 0        | 0        | 1.334e-5 | 0        | 0        |

VOLUME 10, 2022
| CEC | TISOA Ave | TISOA Std | ISOA Ave | ISOA Std | SOA Ave | SOA Std | PSO Ave | PSO Std | WOA Ave | WOA Std | GWO Ave | GWO Std | SFO Ave | SFO Std |
|-----|-----------|-----------|---------|---------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 1   | 2.81e+3   | 3.14e-1   | 3.43e+3 | 4.71e-1 | 1.03e+6| 5.17e+7| 1.29e+6| 7.47e+5| 2.92e+6| 1.38e+6| 4.27e+6| 2.03e+6| 9.17e+5| 4.81e+5|
| 2   | 2.20e+2   | 1.93e+1   | 2.36e+2 | 6.13e+1 | 6.13e+5| 2.78e+5| 1.57e+4| 1.18e+4| 9.07e+5| 4.27e+5| 7.13e+5| 3.14e+5| 1.84e+5| 7.39e-4|
| 3   | 3.20e+2   | 6.14e-2   | 3.37e+2 | 8.91e-1 | 3.21e+2| 1.96e-1| 3.20e+2| 1.10e-1| 3.21e+3| 1.01e-1| 3.20e+3| 6.37e-1| 3.21e+2| 8.412  |
| 4   | 4.07e+2   | 1.13e-1   | 4.20e+2 | 5.34e-1 | 4.83e+2| 1.68e+1| 4.15e+2| 5.135  | 5.28e+3| 8.1647 | 4.71e+3| 3.1475 | 1.9e+3  | 3.17e-2|
| 5   | 5.10e+2   | 7.012     | 6.31e+2 | 3.91e+1 | 2.34e+3| 2.78e+2| 1.28e+3| 3.70e+2| 2.96e+3| 3.02e+2| 6.27e+3| 7.11e+1| 2.17e+3| 6.31e+2|
| 6   | 6.00e+2   | 5.13e-1   | 6.01e+2 | 5.13e-1 | 9.37e+5| 6.46e+5| 4.26e+4| 2.49e+4| 2.70e+5| 4.17e+4| 8.66e+5| 1.4e+4  | 3.39e+4| 4.21e+3|
| 7   | 7.00e+2   | 2.71e-2   | 7.01e+2 | 7.871   | 7.33e+2| 1.79e+1| 7.02e+2| 7.61e-1| 7.87e+2| 5.19e+1| 7.41e+2| 2.23e+1| 7.11e+2| 6.21e+1|
| 8   | 8.01e+2   | 1.813     | 1.21e+3 | 3.293   | 9.74e+3| 9.51e+3| 4.98e+3| 3.95e+3| 4.41e+6| 5.48e+6| 1.21e+6| 4.65e+2| 8.11e+3| 1.96e-3|
| 9   | 1.00e+3   | 3.76e-2   | 1.00e+3 | 3.76e+1| 1.04e+3| 1.76e+1| 1.00e+3| 7.21e-2| 1.07e+3| 2.32e+2| 1.05e+3| 2.61e+1| 1.00e+3| 4.44e-1|
| 10  | 1.31e+3   | 7.41e-1   | 1.97e+3 | 8.663   | 2.16e+4| 1.09e+4| 8.23e+3| 5.49e+3| 4.31e+5| 1.23e+5| 4.89e+5| 3.12e+4| 8.23e+3| 9.18e+2|
| 11  | 1.10e+3   | 8.37e-1   | 1.27e+3 | 7.87e-1| 1.69e+3| 8.68e+1| 1.42e+3| 3.60e+1| 1.65e+3| 1.56e+2| 1.67e+3| 7.71e+1| 1.40e+3| 6.41e-1|
| 12  | 1.30e+3   | 2.28e-3   | 1.30e+3 | 9.34e-1| 1.34e+3| 1.37e+1| 1.30e+3| 1.18e-1| 1.36e+3| 1.25e+1| 1.31e+3| 1.17e+1| 1.40e+3| 1.03e+1|
| 13  | 1.30e+3   | 7.03e-3   | 1.30e+3 | 3.64e+2| 1.30e+3| 1.4184 | 1.30e+3| 9.18e-1| 1.30e+3| 2.4378 | 1.30e+3| 1.9511 | 1.30e+3| 2.39e-1|
| 14  | 1.50e+3   | 4.11e-1   | 1.57e+3 | 3.82e+1| 1.80e+4| 4.95e+3| 7.42e+3| 2.72e+3| 1.84e+4| 2.63e+3| 1.84e+4| 4.68e+3| 7.42e+3| 2.72e-3|
| 15  | 1.60e+3   | 0         | 1.60e+3 | 3.71e-3| 2.56e+3| 1.03e+3| 1.61e+3| 2.25e-1| 5.23e+3| 4.51e+3| 3.81e+3| 1.67e+3| 2.31e+3| 8.73e-1|
the functions can also be found. For $F_8$, TISOA has significantly improved the convergence speed compared with SOA and ISOA.

2) CEC2015
The CEC2015 benchmark functions are used to further verify the optimization performance of TISOA. All algorithms are tested 20 times, and the average ($Ave$) and standard deviation ($Std$) were taken as the evaluation items. The specific function formula is shown in Appendix 2, and the results are shown in Table 3.

The boldface text in Table 3 is the best item of the test results. It can be clearly seen that the results of TISOA for these functions are the best. Among the CEC-9, CEC-12 and CEC-13 benchmark functions, the average values of several optimization algorithms are close, but the TISOA standard deviation is smaller and the optimization stability is stronger. For other functions, the average value and standard deviations of TISOA are the minimum, and TISOA has better optimization performance than other traditional algorithms, which can effectively prove the remarkable optimization performance of TISOA.

IV. POWER TRANSFORMER TROUBLESHOOTING BASED ON KPCA AND TISOA-SVM
In this section, a power transformer fault diagnosis method based on KPCA and TISOA-SVM is described in detail as follows.

A. DATA PROCESSING BASED ON KPCA
The DGA data used in this paper are 300 sets of DGA data obtained from the Hubei Power Grid in China, and some data are given in Table 4.

| $H_2$ | CH$_4$ | C$_2$H$_6$ | C$_2$H$_4$ | C$_2$H$_2$ | Type |
|-------|--------|-----------|-----------|-----------|------|
| 19.2  | 110    | 181.1     | 3.4       | 0.3       | $T_1$ |
| 20.9  | 212.3  | 324.1     | 73.1      | 0.9       | $T_1$ |
| 64.4  | 77.5   | 20.6      | 11.6      | 0         | $T_1$ |
| 45    | 168    | 82        | 330       | 3.1       | $T_2$ |
| 8.4   | 28.7   | 13        | 105       | 2.1       | $T_2$ |
| 27.4  | 63.2   | 32.7      | 137.8     | 0.5       | $T_2$ |
| 77.9  | 6.3    | 5.5       | 13.1      | 0         | PD   |
| 84.3  | 8      | 0.4       | 7.2       | 0         | PD   |
| 87.3  | 6.5    | 1.1       | 5.2       | 0         | PD   |
| 43    | 36     | 41        | 47        | 13        | $D_1$ |
| 46.7  | 8.2    | 1.8       | 5.1       | 4.1       | $D_1$ |
| 49.1  | 12.2   | 0.3       | 3.9       | 4.8       | $D_1$ |
| 90.8  | 24     | 5.2       | 8.9       | 9.1       | $D_2$ |
| 193   | 29     | 6         | 13        | 14        | $D_2$ |
| 176   | 40     | 11        | 72        | 79        | $D_2$ |
| 8.8   | 7.5    | 2.5       | 6.5       | 1         | NC   |
| 3.5   | 1.7    | 1.4       | 8.8       | 3.1       | NC   |
| 2     | 1.5    | 0.6       | 0.7       | 0.5       | NC   |

There are five characteristic quantities in the existing DGA data, namely, $H_2$, CH$_4$, C$_2$H$_6$, C$_2$H$_4$ and C$_2$H$_2$. There are six fault types: medium- and low-temperature overheating ($T_1$), high-temperature overheating ($T_2$), partial discharge (PD), low-energy discharge ($D_1$), high-energy discharge ($D_2$) and normal state (NC). The fault types are numbered, as shown in Table 5.

A. DATA PROCESSING BASED ON KPCA
The DGA data used in this paper are 300 sets of DGA data obtained from the Hubei Power Grid in China, and some data are given in Table 4.
TABLE 7. Partial extraction results.

| No.  | Feature extraction results |
|------|----------------------------|
| 1    | -0.70074 -0.71065 0.03988 -0.20045 |
| 2    | -0.00439 -0.63443 0.58448 -1.52634 |
| 3    | -0.69642 -0.68890 0.01870 -0.22586 |
| 4    | -0.96111 -0.64679 -0.27909 0.44839 |
| 5    | -0.14902 -0.68582 0.65029 -1.33334 |
| 6    | -0.78932 -0.69757 -0.04944 0.01254 |
| 7    | 0.00630 -0.79221 0.75303 -1.60750 |
| 8    | -0.27292 -0.79617 0.43222 -1.02230 |
| 9    | -0.83026 -0.65757 -0.13513 0.09306 |
| 10   | -0.88647 -0.69852 -0.23489 0.35086 |
| 11   | -0.82688 -0.63493 -0.13452 0.09348 |
| 12   | -0.82884 -0.72200 -0.15560 0.21791 |
| 13   | -0.96467 -0.64719 -0.36554 0.54597 |
| 14   | -0.90015 -0.68006 -0.27369 0.37358 |
| 15   | -0.89923 -0.68040 -0.27437 0.37293 |
| 16   | -0.89798 -0.68070 -0.27322 0.36994 |
| 17   | -0.85931 -0.68036 -0.30745 0.31630 |
| 18   | -0.76593 -0.74231 -0.25449 0.20537 |

TABLE 8. Data grouping.

| Type | 1 | 2 | 3 | 4 | 5 | 6 |
|------|---|---|---|---|---|---|
| Train| 20| 20| 20| 20| 20| 20|
| Test | 10| 10| 10| 10| 10| 10|
| Total| 30| 30| 30| 30| 30| 30|

TABLE 9. Diagnosis result of new models.

| Methods         | Diagnostic accuracy | Time(s) |
|-----------------|---------------------|---------|
| TISOA-SVM       | 91.67%              | 8.762   |
| KPCA-TISOA-SVM  | 96.67%              | 7.037   |

for each fault type, and a total of 180 groups of data are selected for fault diagnosis. The groupings of the training set and test set data are shown in Table 8 and have a ratio of 2:1.

B. DATA NORMALIZATION

Due to the different sizes and high complexity of the existing fault data features, it is necessary to normalize the data after KPCA feature extraction and to map all data into (0, 1).

\[ x_i^* = \frac{x_i - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} \]  \hspace{1cm} (31)

where \( i = 1, 2, 3, 4 \); \( x_i^* \) represents normalized data; \( x_{\text{min}} \) represents the minimum value in data; \( x_{\text{max}} \) represents the maximum value in data; \( x_i \) represents raw data.

C. THE PROCESS OF FAULT DIAGNOSIS

This section elaborates the specific flow of power transformer fault diagnosis based on KPCA and TISOA-SVM, and Figure 9 is the specific flow chart.

1) Feature extraction. KPCA is used to extract the features of DGA data, and the principal components are extracted as fault data according to the total contribution rate of the principal components, which is 95%;
2) Data preprocessing. The extracted fault data are loaded; the training sets includes 120 groups, and the test sets includes 60 groups. The data are normalized to (0,1);
3) Parameter optimization. TISOA is used to optimize the parameters \( C \) and \( \sigma \) of the SVM. The specific optimization process is described in Part C of Section III;
4) Judgement of whether the termination conditions are met. If the optimization algorithm reaches the maximum number of iterations, then the optimization is stopped and the optimal parameters are output; otherwise, the optimization process continues;
5) Result output. After obtaining the optimal parameters \( C \) and \( \sigma \), SVM starts fault diagnosis and outputs the results, including the algorithm running time, fault classification, test set accuracy, training set accuracy, and number of misclassifications.

V. EXAMPLE DIAGNOSIS RESULTS

The method based on TISOA-SVM to diagnose the faults with the KPCA-processed DGA data in this paper. Furthermore, three examples are compared and analyzed to verify the superiority of the proposed diagnosis method. The parameter
settings of the algorithms used in all methods are shown in Table 1, and $T_{\text{max}} = 100$; $\text{Population} = 30$; $ub = 10^3$; $lb = 10^{-3}$.

A. EXAMPLE 1
Firstly, TISOA-SVM is used to diagnose the data before and after KPCA processing to verify the validity of KPCA in data processing, Table 9 and figure 10 show the diagnostic results.

The results show that the diagnostic accuracy of proposed method is 96.67%, with 2 misjudgments, and the diagnostic accuracy of TISOA-SVM is 91.67%, with 5 misjudgments; The diagnostic time of proposed method is
7.037s, and diagnostic time of TISOA-SVM is 8.762s. It can be seen that TISOA-SVM is used to diagnosis the DGA data which are processed by KPCA can obtain the higher accuracy and shorter time than diagnosis original DGA data. The above results prove that the DGA data processed by KPCA is more effective in fault diagnosis.

Significantly, in order to carry out the diagnosis test more effectively and intuitively, all the following fault diagnosis tests are based on the DGA data processed by KPCA. Then, the performance of TISOA-SVM is preliminary tested, the fault diagnostic results of GWO-SVM, WOA-SVM, SOA-SVM and SFO-SVM are compared with diagnostic result of KPCA-TISOA-SVM. And the diagnostic results include diagnostic accuracy, diagnostic time and fitness curve. The specific diagnostic results are shown in Figure 11 and Table 10.

By analyzing the results in Figure 11 and Table 10, the fault diagnosis accuracy of GWO-SVM is 90%, with 6 misjudgments; the fault diagnosis accuracy of WOA-SVM is 90%, with 6 misjudgments; the fault diagnosis accuracy of SOA-SVM is 88.33%, with 7 misjudgments; and the fault diagnosis accuracy of SFO-SVM is 88.33%, with 7 misjudgments. It can be seen that diagnostic accuracy of proposed method is approximately 5% higher than other methods, and it has 4-5 fewer misjudgments. The proposed method is significantly better than four models and the diagnostic accuracy has great improvement.

Then, the diagnostic time of five methods is given for further analysis. It can be seen that the result of proposed method is 7.037s, which is fastest compared with above methods. Thus, it can be concluded that proposed method has higher diagnostic efficiency than above methods.
Finally, as shown in Figure 12 and Table 11, fitness curves are given to compare the convergence performance of the proposed method.

As shown in Figure 12 and Table 11, the optimal fitness curve of TISOA-SVM converges fastest, which converges when iterating to the 3th generation, and the fitness value found is the highest. In addition, the average fitness curve converges of proposed method when it iterated to 50 generations, with the increase in the number of iterations, the average fitness also approaches the optimal fitness. A higher fitness value can be found with GWO-SVM, and it starts to converge during the 22th generation. The optimal fitness curve of WOA-SVM converges when it is iterated to 17th generations. A faster convergence speed can be found with SOA-SVM, it starts to converge during the 5th generation, but the fitness value is not high. And the optimal fitness curve of SFO-SVM converges when it is iterated to 3th generations and it has the lowest fitness value in five methods. It is worth noting that except for the proposed method, the average fitness curves of other methods all do not converge to the optimal fitness values.

In summary, the proposed method has the best performance than other methods in diagnostic accuracy, diagnostic time, and fitness curves.

Meanwhile, in order to prove the diagnostic performance of the proposed method more convincingly, five methods from the others in the literature are given to diagnostic the fault of KPCA data in Example 2, and the results are compared with the proposed method.

B. EXAMPLE 2

In the section, the diagnostic results of five methods will compared with the proposed method. These methods are IGWO-SVM [8], IGWO-PNN [16], HGWO-LSSVM [27], BA-SVM [29], and IKH-SVM [30].

Firstly, Figure 13 and Table 12 show the diagnostic accuracy and time of these methods.

The results show that the fault diagnosis accuracy of IGWO-SVM is 93.33%, with 4 misjudgments; the fault diagnosis accuracy of IGWO-PNN is 93.33%, with 4 misjudgments; the fault diagnosis accuracy of HGWO-LSSVM is 91.67%, with 5 misjudgments; the fault diagnosis accuracy of BA-SVM is 90%, with 6 misjudgments; and the fault diagnosis accuracy of IKH-SVM is 95%, with 3 misjudgments. Among them, IKH-SVM has the highest diagnostic accuracy, but it also lower than the proposed method.

Then, diagnostic time of six methods is analyzed. Because the model of PNN has high complexity, so the diagnostic time of IGWO-PNN is the slowest of the other methods, with 21.013s. The same improve methods were used in IGWO-SVM and HGWO-SVM, but these methods reduce the complexity of diagnostic model, because the diagnostic time of IGWO-SVM is not reduce. However, LSSVM has lower complexity than SVM, so the diagnostic time of HGWO-LSSVM has a litter lower than IGWO-SVM. The


### TABLE 12. Diagnosis result of six methods.

| Fault type       | Diagnostic accuracy and the number of misjudgments of each method | Time(s) |
|------------------|---------------------------------------------------------------------|---------|
|                  | 1               | 2               | 3               | 4               | 5               | 6               | Total            |
| IGWO-SVM         | 90% (9/10)      | 100% (10/10)    | 100% (10/10)    | 90% (9/10)      | 80% (8/10)      | 100% (10/10)    | 93.33% (56/60)   | 14.316           |
| IGWO-PNN         | 90% (9/10)      | 100% (10/10)    | 100% (10/10)    | 90% (9/10)      | 80% (8/10)      | 100% (10/10)    | 93.33% (56/60)   | 21.013           |
| HGWO-LSSVM       | 90% (9/10)      | 100% (10/10)    | 100% (10/10)    | 80% (8/10)      | 80% (8/10)      | 100% (10/10)    | 91.67% (55/60)   | 12.591           |
| BA-SVM           | 80% (8/10)      | 90% (9/10)      | 100% (10/10)    | 100% (10/10)    | 70% (7/10)      | 100% (10/10)    | 90% (54/60)      | 11.467           |
| IKH-SVM          | 90% (9/10)      | 100% (10/10)    | 100% (10/10)    | 9% (10/10)      | 90% (9/10)      | 100% (10/10)    | 95% (57/60)      | 8.182            |
| TISOA-SVM        | 90% (9/10)      | 100% (10/10)    | 100% (10/10)    | 100% (10/10)    | 90% (9/10)      | 100% (10/10)    | 96.67% (58/60)   | 7.037            |

Diagnosis time of BA-SVM and IKH-SVM is 11.467s and 8.182s respectively. It can be concluded that the proposed method is the fastest than other methods.

### C. EXAMPLE 3

In this section, statistical test is used to judge about the superiority of the proposed method. 20 times results of six methods are tested by Wilcoxon rank sum test to obtain p-value. That is, whether the proposed method is significantly different from other methods under the standard of \( p = 0.05 \). When \( p > 0.05 \), it can be considered that the assumption of H0 is accepted, there is no significant difference between the two methods, and the diagnostic ability of the methods is the same; When \( p < 0.05 \), H0 is rejected, that is, there is a significant difference between the two methods. And Table 13 shows the results.

The results show that the tested p-value are all far less than 0.05, so the conclusion can be obtained that the proposed method has stronger significance than other methods.

Then, A single test set does not prove the validity of the method. So, 3-fold cross-Validation is used to verify the validity of the proposed method. Table 14 shows results of all methods.

It can be seen that the result of IGWO-PNN is 4% less than that of single test set because the model of PNN has low stability. And compared with other methods, the proposed method has the highest accuracy.

### TABLE 13. p-value of wilcoxon rank sum test.

| Methods       | p-value |
|---------------|---------|
| GWO-SVM       | 1.48e-4 |
| WOA-SVM       | 1.41e-4 |
| SOA-SVM       | 1.54e-4 |
| SFO-SVM       | 1.61e-4 |
| IGWO-SVM      | 1.1e-4  |
| IGWO-PNN      | 1.2e-4  |
| HGWO-LSSVM    | 1.0e-4  |
| BA-SVM        | 1.62e-4 |
| IKH-SVM       | 2.27e-4 |

In this section, the significance and validity of proposed method are proved. Meanwhile, summarizing the results of section A, B, and C, it can be concluded that the proposed method has the best diagnostic performance compared with other methods.

### VI. CONCLUSION

This paper advances a new power transformer fault diagnosis method based on the SVM. First, KPCA is used to extract the features of DGA data. Then, aiming at the defects of SOA, three improved methods are proposed: improved tent chaos initialization, nonlinear inertia weight, and the stochastic double helix formula. Finally, a transformer fault diagnosis method based on KPCA and TISOA-SVM is constructed, and three examples are used to test its diagnostic performance. Significantly, the following conclusions are obtained:

1) The traditional benchmark functions and CEC2015 benchmark functions are used to simulate TISOA and six optimization algorithms. The results show that TISOA has the highest optimization accuracy and fastest convergence speed. The results of these functions effectively prove that TISOA has excellent optimization performance, which can optimize SVM parameters quickly and accurately.
TABLE 15. Traditional benchmark function.

| Function | Dim | Range      | \( f_{\text{min}} \) |
|----------|-----|------------|----------------------|
| \( F_1(x) = \sum_{i=1}^{n} |x_i| + \prod_{i=1}^{n} |x_i| \) | 30  | [-10, 10] | 0 |
| \( F_2(x) = \sum_{i=1}^{n-1} \left[ 100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2 \right] \) | 30  | [-30, 30] | 0 |
| \( F_3(x) = \sum_{i=1}^{n} ix_i^4 + \text{rand}(0,1) \) | 30  | [-1.28, 1.28] | 0 |
| \( F_4(x) = \sum_{i=1}^{n} \left[ x_i^2 -10 \cos(2\pi x_i)+10 \right] \) | 30  | [-5.12, 5.12] | 0 |
| \( F_5(x) = -20 \exp \left(-0.2 \left[ \sum_{i=1}^{n} x_i^2 \right] - \exp \left( \frac{1}{n} \sum_{i=1}^{n} \cos \left( 2\pi x_i \right) \right) + 20 + e \) | 30  | [-32, 32] | 0 |
| \( F_6(x) = \frac{1}{4000} \sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n} \cos \left( \frac{x_i}{\sqrt{i}} \right) + 1 \) | 30  | [-600, 600] | 0 |
| \( F_7(x) = \sum_{i=1}^{11} \left[ \frac{x_i (x_i^2 - 10x_i^5 + 5x_i^3 + 5x_i)}{x_i^2 + 100} \right] -1 \) | 4   | [-5, 5] | 0.0003 |
| \( F_8(x) = \left[ 1 + \left( x_1 + x_2 + 1 \right)^2 \left( 9 - 14 x_1 + 6 x_2^2 \right) \times \left[ 1 + \left( 2 x_1 - 3 x_2 \right)^2 + \left( 2 x_2 - 1 \right)^2 + \left( 9 x_1 - 14 x_2 \right)^2 + \left( 27 x_1^2 + 7 x_2^2 \right) \right] \) | 2   | [-2, 2] | 3 |
| \( F_9(x) = -\sum_{i=1}^{n} \left[ (X - a_i)^2 \left( X - a_i \right)^2 + c_i \right]^{-1} \) | 4   | [0, 10] | -10.536 |

TABLE 16. Function of CEC2015.

| Type                  | No.  | Functions                                      | \( f_{\text{min}} \) |
|-----------------------|------|------------------------------------------------|----------------------|
| Unimodal Functions    | CEC-1| Rotated High Conditioned Elliptic Function     | 100                  |
|                       | CEC-2| Rotated Cigar Function                         | 200                  |
|                       | CEC-3| Shifted and Rotated Ackley's Function           | 300                  |
| Simple Multimodal Functions | CEC-4| Shifted and Rotated Rastrigin's Function       | 400                  |
|                       | CEC-5| Shifted and Rotated Schwefel's Function         | 500                  |
|                       | CEC-6| Hybrid Function 1 (N=3)                        | 600                  |
| Hybrid Functions      | CEC-7| Hybrid Function 2 (N=4)                        | 700                  |
|                       | CEC-8| Hybrid Function 3(N=5)                         | 800                  |
|                       | CEC-9| Composition Function 1(N=3)                    | 900                  |
|                       | CEC-10|Composition Function 2 (N=3)                   | 1000                 |
|                       | CEC-11|Composition Function 3(N=5)                    | 1100                 |
| Composition Functions | CEC-12|Composition Function 4(N=5)                    | 1200                 |
|                       | CEC-13|Composition Function 5(N=5)                    | 1300                 |
|                       | CEC-14|Composition Function 6 (N=7)                   | 1400                 |
|                       | CEC-15|Composition Function 7 (N=10)                  | 1500                 |

Search Range: [-100, 100]
2) The DGA data based on KPCA are used for fault diagnosis, and three examples are used to verify the diagnostic performance of proposed method. Compared with other method, the diagnostic accuracy and diagnostic time of proposed method have been greatly improved, and the significance and validity of the method is also the strongest.

In conclusion, the fault diagnosis method proposed in this paper has excellent diagnostic performance, can diagnose transformer faults quickly and accurately, and has high reference values. However, the research on DGA data in this paper is not comprehensive. In the future, DGA data will be deeply analyzed and studied to further improve the accuracy and stability of transformer fault diagnosis.

APPENDIX A
See Table 15.

APPENDIX B
See Table 16.

REFERENCES
[1] L. Tao, X. Yang, Y. Zhou, and L. Yang, “A novel transformers fault diagnosis method based on probabilistic neural network and bio-inspired optimizer,” Sensors, vol. 21, no. 11, p. 3623, May 2021.
[2] S. Lu, W. Gao, C. Hong, and Y. Sun, “A newly-designed fault diagnostic method for transformers via improved empirical wavelet transform and kernel extreme learning machine,” Adv. Eng. Informat., vol. 49, Aug. 2021, Art. no. 101320.
[3] J. Wu, A Transformer in Yongzhou, Hunan Suddenly Exploded and Caught Fire, Killing 1 and Injuring 7. [Online]. Available: https://www.sohu.com/a/150193454_652081
[4] 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, Mar. 2017.
[5] D. R. Morais and J. G. Rolim, “A hybrid tool for detection of incipient faults in transformers based on the dissolved gas analysis of insulating oil,” IEEE Trans. Power Del., vol. 21, no. 2, pp. 673–680, Apr. 2006.
[6] J. Fan, F. Wang, Q. Sun, F. Bin, F. Liang, and X. Xiao, “Hybrid RVM–ANFIS algorithm for transformer fault diagnosis,” IET Gener., Transmiss. Distrib., vol. 11, no. 14, pp. 3637–3643, Aug. 2017.
[7] H. A. Illias, X. R. Chai, and A. H. A. Bakar, “Hybrid modified evolutionary particle swarm optimisation-time varying acceleration coefficient-artificial neural network for power transformer fault diagnosis,” Measurement, vol. 90, pp. 94–102, Aug. 2016.
[8] X. Huang, X. Huang, B. Wang, and Z. Xie, “Fault diagnosis of transformer based on modified grey wolf optimization algorithm and support vector machine,” IEEE Trans. Electr. Electron. Eng., vol. 15, no. 3, pp. 409–417, Mar. 2020.
[9] W. Teng, S. Fan, Z. Gong, W. Jiang, and M. Gong, “Fault diagnosis of transformer based on fuzzy clustering and the optimized wavelet neural network,” Syst. Sci. Control Eng., vol. 6, no. 3, pp. 359–363, Sep. 2018.
[10] A. Li, X. Yang, Z. Xie, and C. Yang, “An optimized GRNN-enabled approach for power transformer fault diagnosis,” IEEE Trans. Electr. Electron. Eng., vol. 14, no. 8, pp. 1181–1188, Aug. 2019.
[11] H.-H. Yi, J. Wang, and G.-G. Wang, “Improved probabilistic neural networks with self-adaptive strategies for transformer fault diagnosis problem,” Adv. Mech. Eng., vol. 8, no. 1, pp. 1–13, Jan. 2016.
[12] 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, Nov. 2018.
[13] G. Xie and L. Niu, “Transformer fault diagnosis based on an artificial bee colony-support vector machine optimization algorithm,” Power Syst. Protection Control, vol. 48, no. 15, pp. 156–163, Aug. 2020.
[14] A. Hoballah, D.-E.-A. Mansour, and I. B. M. Taha, “Hybrid grey wolf optimizer for transformer fault diagnosis using dissolved gases considering uncertainty in measurements,” IEEE Access, vol. 8, pp. 139176–139187, 2020, doi: 10.1109/ACCESS.2020.3012633.
[15] G. Mani and J. Jerome, “Intuitionistic fuzzy expert system based fault diagnosis using dissolved gas analysis for power transformers,” J. Electr. Eng. Technol., vol. 9, no. 6, pp. 2058–2064, Nov. 2014.
[16] Y. Zhou, X. Yang, L. Tao, and L. Yang, “Transformer fault diagnosis model based on improved grey wolf optimizer and probabilistic neural network,” Energies, vol. 14, no. 11, p. 3029, May 2021.
[17] S. Kim, S.-H. Jo, W. Kim, J. Park, J. Jeong, Y. Han, D. Kim, and B. D. Youn, “A semi-supervised autoencoder with an auxiliary task (SAAT) for power transformer fault diagnosis using dissolved gas analysis,” IEEE Access, vol. 8, pp. 179295–179302, 2020.
[18] T. Mehran and T. Stefan, “Transformer winding condition assessment using feedforward artificial neural network and frequency response measurements,” Energies, vol. 14, no. 11, pp. 1–22, May 2021.
[19] F. Yuan, J. Guo, Z. Xiao, B. Zeng, W. Zhu, and S. Huang, “A transformer fault diagnosis model based on chemical reaction optimization and twin support vector machine,” Energies, vol. 12, no. 5, p. 960, Mar. 2019.
[20] Y. Zhou, J. Guo, L. Qiu, W. Zhu, and S. Huang, “A transformer fault diagnosis model optimized by GWO to the prediction of soil liquefaction,” Environ. Earth Sci., vol. 80, no. 9, pp. 1–9, Apr. 2021.
[21] H. Li, J. Tao, D. Zhang, L. Chang, Y. Lu, X. Zhao, and Y. Guo, “Analog circuit fault diagnosis based on support vector machine classifier and fuzzy feature selection,” Electronics, vol. 10, no. 12, p. 1496, Jun. 2021.
[22] W. Yan and H. Shao, “Application of support vector machines and least squares support vector machines to heart disease diagnoses,” Control Decis., vol. 18, no. 3, pp. 360–358, May 2003.
[23] 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, Mar. 2017.
[24] Y. Zhou, X. Yang, L. Tao, and L. Yang, “Transformer fault diagnosis model based on improved grey wolf optimizer and probabilistic neural network,” Energies, vol. 14, no. 11, p. 3029, May 2021.
[34] Y. Xu, A. Fenerci, O. Øiseth, and T. Moan, “Efficient prediction of wind and wave induced long-term extreme load effects of floating suspension bridges using artificial neural networks and support vector machines,” Ocean Eng., vol. 217, Dec. 2020, Art. no. 107888.

[35] G. Dhiman and V. Kumar, “Seagull optimization algorithm: Theory and its applications for large-scale industrial engineering problems,” Knowl.-Based Syst., vol. 165, pp. 169–196, Feb. 2019.

[36] L.-L. Li, S.-J. Zheng, M.-L. Tseng, and Y.-W. Liu, “Performance assessment of combined cooling, heating and power system operation strategy based on multi-objective seagull optimization algorithm,” Energy Convers. Manage., vol. 244, Sep. 2021, Art. no. 114443.

[37] Y. Li, M. Han, and Q. Guo, “Modified whale optimization algorithm based on tent chaotic mapping and its application in structural optimization,” KSCE J. Civil Eng., vol. 24, no. 12, pp. 3703–3713, Oct. 2020.

[38] S. Y. D. Nezhad, N. Safdarian, and S. A. H. Zadeh, “New method for fingerprint images encryption using DNA sequence and chaotic tent map,” Optik, vol. 224, Dec. 2020, Art. no. 165661.

[39] S. Mirjalili and A. Lewis, “The whale optimization algorithm,” Adv. Eng. Softw., vol. 95, pp. 51–67, May 2016.

YUHAN WU (Student Member, IEEE) received the B.E. degree from the Department of Electrical Engineering and Automation, School of Science and Technology, China Three Gorges University, in 2019. He is currently pursuing the master’s degree with the School of Information Engineering, Hubei Minzu University. His research interests include power transformer fault diagnosis and artificial intelligence algorithm improvement.

XIANBO SUN received the B.S. degree in electrical engineering and automation from Hubei Minzu University, in 2000, the M.S. degree in control theory and control engineering and the Ph.D. degree in traffic information engineering and control from the Wuhan University of Technology, in 2008 and 2019, respectively. He is currently a Professor and a Master Tutor with Hubei Minzu University. Doing multiple research work in the fields of control theory and control engineering, industrial automation. With more than ten articles published in domestic and foreign academic journals, he has chaired Natural Science Foundation projects, and two lateral projects. Besides, he is also the Core Member of Hubei Province Innovation Team. He has received three national utility model patents, two national invention patents, and won two awards for Enshi Scientific and Technological Progress, one award for Hubei Provincial Scientific and Technological Progress, and the Enshi Second Youth Science and Technology Award.

XIANJING ZHONG received the B.E. degree from the Department of Electrical Engineering and Automation, School of Science and Technology, Xi’an University of Science and Technology, in 2021. She is currently pursuing the master’s degree with the School of Information Engineering, Hubei Minzu University.

YI ZHANG received the B.E. degree from the Department of Electrical Engineering and Automation, School of Science and Technology, North China Electric Power University, in 2020. He is currently pursuing the master’s degree with the School of Information Engineering, Hubei Minzu University. His research interest includes power quality, especially three-phase unbalance of distribution networks.

LEI CHENG received the B.E. degree from the Department of Electrical Engineering and Automation, School of North China Electric Power University, in 2019. He is currently pursuing the master’s degree with the School of Information Engineering, Hubei Minzu University.

* * *