TRANSFORMER FAULT DIAGNOSIS METHOD BASED ON ReliefF AND HPO-SVM

Xiaohu ZHANG[1], Huanyu NING[2]

In order to improve the correct judgment rate of fault diagnosis of oil-immersed power transformers, this paper proposes a fault diagnosis method using the combination of the relief feature weight method, HPO-SVM model, and dissolved gas analysis in oil (DGA). First, the method introduces the feature weight algorithm to filter and reduce the dimensionality of the input quantity; second, the probabilistic neural network model is optimized using the predator optimization algorithm, and the SVM model is used to process the DGA ratio set to finally obtain the fault diagnosis results of the transformer. The experiments show that the model with dimensionality reduction using the reliefF feature weight algorithm has higher diagnostic accuracy. The average fault judgment accuracy of HPO-SVM, GWO-SVM, WOA-SVM, and PSO-SVM is 94%, 91.33%, 90%, and 83.33%, respectively, and the average number of iterations is 6.5, 8.9, 12.5, and 15.1, respectively, and the simulation results show that the preferred hybrid feature model has higher correct diagnosis rate and faster convergence seeking speed The simulation results show that the superiority of this scheme is confirmed.

Keywords: transformer fault diagnosis; dissolved gas analysis in oil; Feature weight; Hunter-Prey Optimizer-support vector machine.

1. Introduction

Transformer is the core electrical equipment in the power system, used to connect the power system with different voltage levels. Because the transformer internal lines have different degrees of aging and damage, the operation of the whole transformer will produce great risks. Once the failure occurs, it will seriously harm the whole power network. Therefore, accurate and fast fault diagnostic methods are needed to ensure life safety and reduce property losses. Due to the degeneration and aging of the transformer's internal wiring and structure, the gas generated will dissolve in the immersed oil, resulting in different degrees of discharge, overheating, and other abnormal transformer states. At present, the judgment method based on dissolved gas analysis in oil (DGA) has been widely used in China's electric power system.

Among the many DGA methods, the three-ratio method (IEC) and the no-code method are the more commonly used methods, but there are problems such

[1] Prof., Dept. of Electrical and Information Engineering, Hunan University of Technology
[2] College of Electrical and Information Engineering, Hunan University of Technology, Zhuzhou, e-mail: a2501547985@163.com
The Project Supported by National Key R&D Program of China Grant No.2022YFE0105200）

as missing codes, too vague boundary problems, and too absolute classification methods [1]. Meanwhile, scholars have added various optimization algorithms to machine learning models, and the literature [2] used the BP model to diagnose transformer fault types, although this method is more accurate and fault tolerant, it requires more parameters to be controlled, and it is easy to fall into local optimization, and it requires more parameters, and there is no effective method for parameter selection; the literature [3] combined RBF model to transformer fault diagnosis, although the radial basis neural network has a simple structure, fast convergence, and can approximate arbitrary nonlinear functions, the learning speed is slow and not suitable for occasions with high real-time requirements; and the literature [4] uses a recurrent neural network optimized by the bat algorithm to identify the fault type of transformers, but the RNN has the disadvantages of different input and output sequences and decreasing accuracy of processing long-term dependence. In contrast, the method of using SVM and fusing feature weight algorithm to screen the mixed feature set can effectively solve the regression and classification problems of high-dimensional features, and the kernel techniques with excellent research results are available to cope with linear indistinguishability.

Selecting suitable feature extraction methods and screening out the most relevant sets of samples can effectively improve the prediction accuracy of transformer fault diagnosis while selecting suitable parameters to optimize the SVM model will further enhance the accuracy and reliability of the model. To this end, some mainstream algorithms, such as Grey Wolf Optimization (GWO), Whale Optimization Algorithm (WOA), and Particle Swarm Optimization (PSO), are selected in this paper. However, GWO suffers from low solution accuracy, slow convergence, and easy to falls into local optimum [5].WOA suffers from low convergence accuracy and slow convergence [6].PSO converges too early, i.e., the algorithm is not a globally optimal solution [7].HPO optimization (Hunter-Prey Optimization, HPO) is a new algorithm inspired by HPO is a new heuristic algorithm inspired by predator hunting, and this method possesses a better optimization finding ability and convergence speed. Inspired by this, this paper adopts a prediction model combining the HPO optimization algorithm and SVM, and this model outperforms the three models GWO, WOA, and PSO in terms of prediction accuracy and convergence speed. Therefore, this paper incorporates the weighted dimensionality reduction algorithm and the predator optimization support vector machine approach for further improving the correct fault diagnosis rate of predicted immersion transformers.

2. ReliefF feature weight method

In 1992, the Relief algorithm was born. The Relief method is a feature weighting algorithm, which gives different differentiated weights to features based on the degree of association of each feature and kind, and features with weights less than a certain set value will be removed. The Relief method is widely applicable because of its low threshold of understanding, fast operation, and high accuracy of test results, but it cannot handle multidimensional data [8], so the ReliefF algorithm was born two years later. Compared with the Relief algorithm, the ReliefF algorithm increases the ability to handle multidimensional data. In the process of multidimensional solution, ReliefF extracts a training set (*R*) in the overall data and then extracts k similar sets (Near Hit) from this training set, and each feature set extracts *k* such sets (Near Miss), and finally the feature The weight values of the sets are re-iterated. The following equation is used.

$$W(A)=W(A)-\sum_{j=1}^{k} diff(A,R,H_j)/(mk)+\sum_{C\neq class(R)}[\frac{p(C)\sum_{j=1}^{k} diff(A,R,M_j(C))}{1-p(Class(R))}]/(mk) \quad (1)$$

Where, $W(A)$ is the weight value, sample R_1 and sample R_2 make the difference on the feature A to get $diff(A,R_1,R_2)$, and $M_j(C)$ represents the j th similarity set in class C. The equation is as follows:

$$(A,R_1,R_2)=\begin{cases}\frac{|R_1[A]-R_2[A]|}{\max(A)-\min(A)} & \text{If } A \text{ is continuous}\\ 0 & \text{If } A \text{ is discrete, and } R_1[A]=R_2[A]\\ 1 & \text{If } A \text{ is discrete, and } R_1[A]\neq R_2[A]\end{cases} \quad (2)$$

The algorithmic flow of reliefF algorithm is shown below:

Set the training set as D , the data set sampled m times, δ as the critical value, the number of most similar samples k , and output the characteristic weight value T of each set.

1) Set the weight value to zero and T is the empty set.

2) $for\ i=1\ to\ m\ do$

1. Randomly select sample R from D;
2. Find *k* of the most similar $H_j(j=1,2,\cdots,k)$ of *R* from the homogeneous set of *R*, and pick out *k* of the most similar $M_j(C)$ from different sets of the same set;
3. $for\ A=1\ to\ N$ (all characteristics)

$$W(A)=W(A)-\sum_{j=1}^{k} diff(A,R,H_j)/(mk)+\sum_{C\neq class(R)}[\frac{p(C)\sum_{j=1}^{k} diff(A,R,M_j(C))}{1-p(Class(R))}]/(mk)$$

end

3. HPO-SVM model

Combined with the above-mentioned reliefF algorithm for feature set screening session, this section proposes to use HPO-SVM to diagnose the filtered model inputs and finally select the optimal parameter C values and gamma values, so as to perform the optimal diagnosis on the multidimensional DGA data.

3.1 Support vector machine

Support vector machines are a class of generalized linear classifiers for batch processing of binary data. Since its decision boundary is solved in the maximum margin plane, its ability to quadratically solve plans is significant in complex nonlinear problems [9]. Let the range of the training set be $T=\{(x_1,y_1),(x_2,y_2)\cdots(x_n,y_n)\}$, x be the input features, and y be the output metrics. The optimal solution function is： $f(x)=\mathrm{sgn}(g(x))$ ， The linear classifier $g(x)=w^T x+b$, w is the weight and b is the bias, and the maximum interval between the two vector hyperplanes is $\max\limits_{w,b}\frac{2}{\|w\|}$, so the objective function can be transformed into $\min\limits_{w,b}\frac{\|w\|^2}{2}$, and the constraint expression is $s.t.\ y_i(w^T x_i+b)\geq 1$. Fig. 1 is the SVM decision boundary diagram.

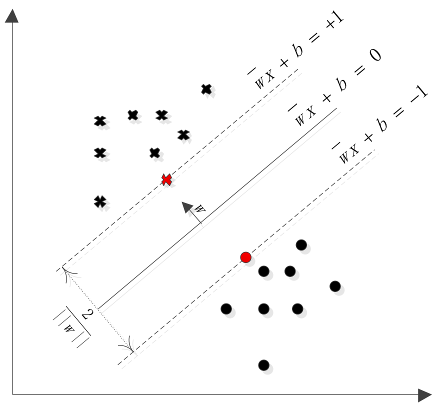

Fig. 1. SVM decision boundary diagram

When the data is non-linear and cannot accurately identify the type, ξ (relaxation constant) is introduced. When $\xi_i \geq 0$ is large enough, the preset training value can always be found. However, in order to avoid exceeding the value of ξ_i, C (penalty term) is introduced into the objective function, the objective function is updated to $\min_{w,b} \frac{\|w\|^2}{2} + C\sum_{i=1}^{n}\zeta_i$, and the constraint expression iteration is $s.t.\ y_i(w^T x_i + b) \geq 1 - \zeta_i$.

Quadratic optimization of its dual problem can be obtained:

$$\begin{cases} \max_{\alpha} \sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j x_i^T x_j \\ s.t. \sum_{i=1}^{n} a_i y_i = 0 \end{cases} \tag{3}$$

Where α is a Lagrange parameter ($0 \leq \alpha_i \leq C$)and i is a natural number.

Finally, A kernel function $k(x,z) = \Phi(x)\cdot\Phi(z)$ is needed to distinguish the original data of high-dimensional space, that is, the final optimal classification boundary decision expression is:

$$\begin{aligned} f(x) &= \operatorname{sgn}(w^T x + b) \\ &= \operatorname{sgn}(\sum_{i=1}^{n}\alpha_i y_i K(x_i, z) + b) \end{aligned} \tag{4}$$

3.2 Hunter-Prey Optimization

Hunter-Prey Optimization (HPO) is an innovative heuristic algorithm pioneered in 2022 [10], which was designed based on the idea of predator hunting (tiger, lion, cheetah) and predation scenarios (antelope, zebra) of the predated animals. The method used in this paper differs from the computational ideas of previous animal algorithms in that in this paper, the objects are divided into a predator and a prey population, and the predator attacks individuals that are separated from the prey population. The predator chooses the best place to initiate hunting based on distant prey, and the prey picks the safest habitat to escape from hunting. The location searched for is the optimal value of functional fitness. the mathematical model of HPO can be expressed as:

1) Within the search space, initial populations of prey or predators are generated using uniformly distributed random number theory within the upper and lower bounds of the respective populations.

$$hp_i = rand(1, n_v).*(u_{hp} - l_{hp}) + l_{hp} \tag{5}$$

In this equation, hp_i is the position of predator/prey, and u_{hp} and l_{hp} are the lower and upper limits of predator and prey; $rand$ is the uniformly distributed random number, and nv is the number of variables or the dimension of the problem. The fitness rating is performed in this step, and the population with the best fitness is taken as the global optimum for the current iteration and advanced to the stage before the iteration.

2) In the exploration phase, the search process must be repeated to seek the best place of refuge for the prey. Exploration and exploitation are the two phases of the search process. Exploration is the tendency of the algorithm to run randomly leading to the multiplicity of solutions. Therefore, in the search phase, the randomness needs to be reduced to enable the algorithm to find the global optimal solution. HPO in the detection phase of the k TH iteration update of predator position is defined as:

$$hp_{ij}(k+1) = hp_{ij}(k) + 0.5\{[2B_p A_p P_{pr(j)} - hp_{ij}(k)] + [2(1-B_p)A_p \gamma_j - hp_{ij}(k)]\} \quad (6)$$

Where, $hp_{ij}(k+1)$ is the modified predator orientation for the next iteration, $P_{pr(j)}$ is the prey orientation, γ_j is the average value of all prey orientation, A_p is the adaptive parameter, and B_p is a parameter to balance the exploration and utilization phase. The following equations are used to define A_p and B_p.

$$A_p = r_2 \otimes idx + r_3 \otimes (\sim idx) \quad (7)$$

$$B_p = 1 - \frac{0.98 \times k}{k_{\max}} \quad (8)$$

Where $\overline{r_1} = rand(1, n_v) < B_p$, $r_2 = rand()$, $\overline{r_3} = rand(1, n_v)$, $idx = (r_1 == 0)$,here $\overline{r_1}$ and $\overline{r_3}$ are random vectors in [0,1], their magnitudes are equal to the number of variables, respectively, r_2 is a random number, idx is the index number of $\overline{r_1}$ when $r_1 = 0$ is satisfied, $\otimes$ is the product between elements, k and $k_{\max}$ are the number of current update iterations and the maximum number of update iterations in turn.

Using the average of all prey locations γ_j and the Euclidean distance between each search variable to calculate the prey location P_{pr}. The equation is as follows:

$$\gamma_j = mean(hp_{ij}) \quad (9)$$

$$D_{euc(i)} = \left(\sum_{j=1}^{n_v} (hp_{ij} - \gamma_j)^2\right)^{\frac{1}{2}} \quad (10)$$

Consider the new random variables as follows:

$$P_{best} = round(B_p \times n_v) \quad (11)$$

n_v is the number of search agent assistants. The search agent assistant can be represented as:

$$\overline{P_{pr}} = \overline{hp_l} \mid i \ issorted \ D_{euc}(p_{best}) \quad (12)$$

3) In the process of hunting, the prey will escape and look for the next safe refuge point, assuming that the prey has the highest probability of survival in this place, that is, this then the difficult point is called the global best position.

$$hp_{ij}(k+1) = G_{pr(j)} + B_p A_p \cos(2\pi r_4) \times (G_{pr(j)} - hp_{ij}) \qquad (13)$$

Where $hp_{ij}(k+1)$ is the prey position of the next iteration, G_{pr} is the globally optimal prey position, and r_4 is a random number between 0 and 1.

4) The premise of using Eq. (7) and Eq. (14) is to identify the predator or prey. According to Eq. (14), we know that if $r<R$, hp_{ij} will be the predator, otherwise, it will be regarded as the prey.

$$hp_{ij}(k+1) = \begin{cases} Eq(6) \;\; if \;\; r_5 < R_p \\ Eq(13) \;\; else \end{cases} \qquad (14)$$

5) Where r_5 is the random number between [0,1], and R_p is the adjustment parameter for the selection process and is set to 0.1.

From the overall process link of the HPO algorithm, it is easy to see that it is a precise and powerful algorithm in practice because of its unique ability to solve single- and multi-peak problems while maintaining good stability between probing and exploitation.

3.3 HPO optimizes SVM

Based on the above conclusions, the HPO algorithm is used in this section to find the optimal solution of SVM in the input value. The process is described as follows:

1): Initialize the model randomly, input the maximum number of iterations, the number of optimization parameters, the upper and lower limit of optimization parameters, and the population number.

2): HPO-SVM optimization fitness was calculated and selected.

3): Evaluate the global optimal prey location.

4): According to Eq. (7) and Eq. (8), adaptive parameters and balanced exploration parameters were updated.

5): If , the predator successfully hunts, then Eq. (6) is used to calculate the modified position of the predator in the next iteration; otherwise, the new random variable  is considered to calculate.

6): Continue to evaluate the global best prey position, if the termination condition is satisfied, that is, output the optimal solution C value and gramma value as SVM model parameters, otherwise continues to step 4) iteration.

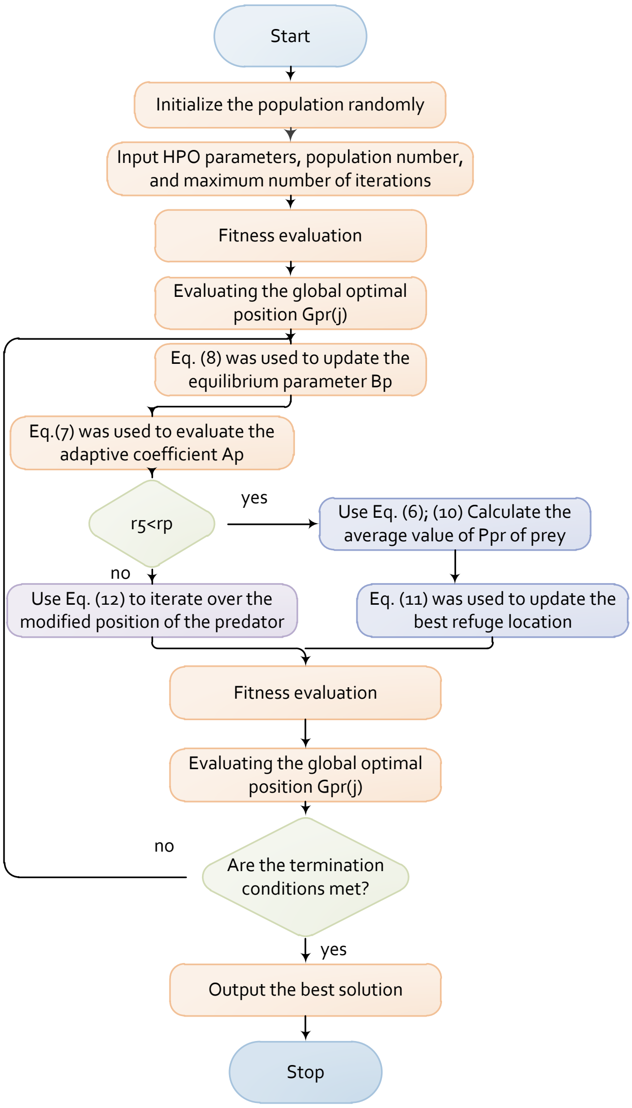

Fig. 2. Data optimization diagram based on HPO optimization algorithm

3.4 reliefF combined with HPO-SVM fault detection

To improve the correct rate of the diagnostic model, this paper specially fuses the reliefF feature weighting method and HPO-SVM model, and the total link is shown in Fig. 3, which includes the following four steps.

Step 1: Input the characteristic sample set of transformer fault dissolved gas, and obtain the ratio relationship between H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2 by the emissions of five kinds of gas. The ratios were imported into the reliefF feature weight model and ranked from high to low according to the weight values to obtain the correlation degree of different ratios.

Step 2: Construct the DGA sample set. 4/5 of the total transformer fault sample set is randomly selected as the training set and the remaining 1/5 as the test set. In this paper, the sets are selected as new sets in order of ranking of weight values until 14 sets of ratio sets are selected, and the correctness of these 14 sets are tested separately.

Step 3: Import the highest set of ratios into HPO-SVM, GWO-SVM, WOA-SVM and PSO-SVM respectively, compare their accuracy and the number of iterations of optimal solutions, and obtain the optimal parameters of SVM.

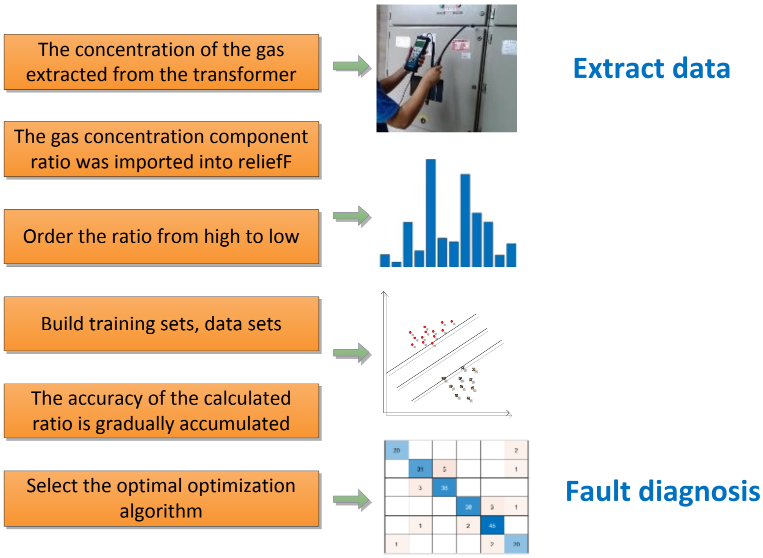

Fig. 3. Fault detection flow chart based on reliefF and HPO-SVM

4. Example fault detection model simulation

4.1 Experimental data analysis

In order to prove the superiority and universality of the fault data diagnosis method developed in this paper, 750 groups of DGA data of oil-immersed transformer faults were used for experimental data in this paper. The data were sourced from [11],[12]and[13], and the status types were divided into: Partial discharge, low temperature overheating, low energy discharge, high temperature overheating, high energy discharge, normal and label them in turn. 4/5 of the experimental samples are taken as the training set (600 sets of data), and the remaining 1/5 is taken as the test set (150 sets of data). The specific labels and quantities of the six fault types are shown in Table 1.

Table 1

Six transformer fault types and labels

Fault category	Sample number	Training set	Test set	label
Partial discharge	125	100	25	1

Low energy discharge	125	100	25	2
High energy discharge	125	100	25	3
Medium-low temperature superheating	125	100	25	4
High temperature superheating	125	100	25	5
Normal	125	100	25	6

4.2 reliefF feature weight method for data extraction

The DGA data is mainly composed of the concentrations of five gases: hydrogen H_2, methane CH_4, ethane C_2H_6, acetylene C_2H_2, and ethylene C_2H_4. C1(the sum of methane, acetylene and ethylene) and C2(the sum of all gases) were introduced to obtain the ratios of 14 groups of gases to subdivide the ratio relationship of the above gases. The correlation degree of these 14 groups of ratios was calculated, and the elements with obvious weights were selected to fully retain the characteristic of faults. The contribution rates of the weights of each group are shown in Table 2 and Fig. 4.

Table 2

Six transformer fault types and labels

14 groups of ratio labels	$W(A)$	Gas ratio
1	0.0040	CH_4/H_2
2	0.0030	C_2H_2/C_2H_4
3	0.0019	C_2H_4/C_2H_6
4	0.0039	C_2H_6/CH_4
5	0.0065	C_2H_2/CH_4
6	0.0052	C_2H_2/H_2
7	0.0919	$CH_4/C1$
8	0.0815	$C_2H_4/C1$
9	0.1076	$C_2H_2/C1$
10	0.0141	$H_2/C2$
11	0.0178	$CH_4/C2$
12	0.1103	$C_2H_6/C2$
13	0.0797	$C_2H_4/C2$
14	0.1272	$C_2H_2/C2$

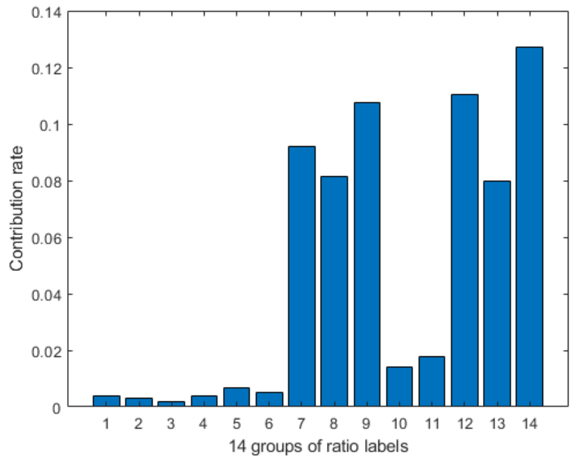

Fig. 4. Comparison of diagnostic accuracy under different ratios

In order to prove that reliefF feature weight method has higher detection accuracy, this paper chooses to compare with traditional gas detection models (IEC three-ratio method and non-coding nine-ratio method). The parameters of IEC three-ratio method are C_2H_2/C_2H_4, CH_4/H_2 and C_2H_4/C_2H_6. The parameters of the uncoded nine-ratio method are CH_4/H_2, C_2H_2/C_2H_4, C_2H_4/C_2H_6, $C_2H_2/(T1+T2)$, $H_2/(H_2+T1+T2)$, $C_2H_4/(T1+T2)$, $CH_4/(T1+T2)$, $C_2H_6/(T1+T2)$, $(CH_4+C_2H_4)/(T1+T2)$, ($T_1= CH_4$, $T2= C_2H_2 +C_2H_4+C_2H_6$). The three ratio sets were input into the HPO-SVM model.

4.3 Analysis of detection results of different SVM models

In this section, HPO algorithm is compared with some existing algorithms (such as Grey Wolf Optimization (GWO), Whale Optimization Algorithm (WOA) and Particle Swarm Optimization (PSO)) in terms of detection accuracy and convergence speed.

In the feature extraction stage, it can be seen from the reliefF model that the weight values of the 14 groups of data ranked from high to low (14, 12, 9, 7, 13, 11, 10, 5, 6, 1, 4, 2, 3). We obtained 14 sets of ratios in order from highest to lowest weight values, as shown in Table 4. Then they were substituted into HPO-SVM, GWO-SVM, WOA-SVM, and PSO-SVM with a uniform population size of 20, a dimension of 2, and a number of 20 iterations with upper and lower bounds of parameters [0.01, 5]. Finally, MATLAB simulation software is used to simulate diagnosis and prediction. The test results are shown in Table 3.

Table 3

Details of fault accuracy of different accumulative degrees

Groups	HPO-SVM/%	GWO-SVM/%	WOA-SVM/%	PSO-SVM/%
1group	45.33	45.33	45.33	43.33
2 groups	66.67	65.33	65.33	62

3 groups	73.33	73.33	73.33	68
4 groups	86.67	86.67	82.67	81.33
5 groups	88	86	83.33	80
6 groups	87.33	86.67	84	78
7 groups	89.33	88.67	85.33	82.67
8 groups	94	91.33	85.33	82.67
9 groups	92	90.67	86.67	81.33
10 groups	92.66	90	90	83.33
11 groups	92.66	88.67	88.67	82
12 groups	92	90.67	88.67	80
13 groups	93.33	91.33	88.67	82.67
14 groups	92	87.33	88.67	81.33
IEC	53.33	53.33	53.33	52
uncoded nine-ratio method	87.33	86.77	84.67	79.33

The experimental results showed that when the first 8 groups were selected, the accuracy of HPO was the highest (94%), which was higher than that of the complete 14 groups (92%), IEC ratio method (53.33%) and uncoded 9 ratio method (87.33%) by 2%, 40.67% and 6.67%.

Grey Wolf Optimization (GWO), Whale Optimization Algorithm (WOA) and Particle Swarm Optimization (PSO) reached the highest diagnostic accuracy rate (91.33%, 90%, 83.33%) in group 8, group 10 and group 10, respectively, which was significantly improved compared with the IEC three-ratio method and the uncoded nine-ratio method. It is proved that reliefF feature weight method can significantly improve the prediction accuracy.

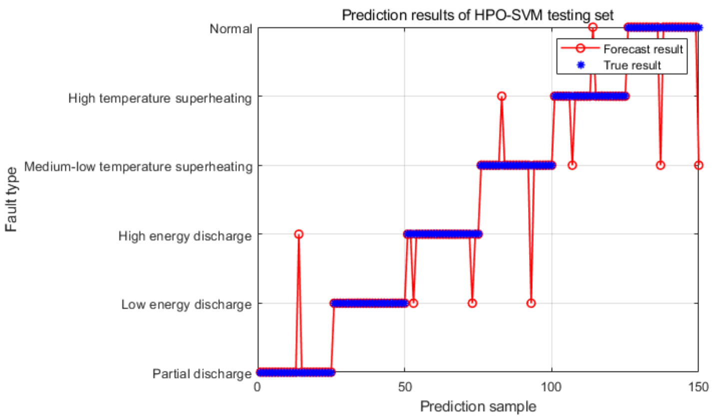

Fig. 5. HPO-SVM test results

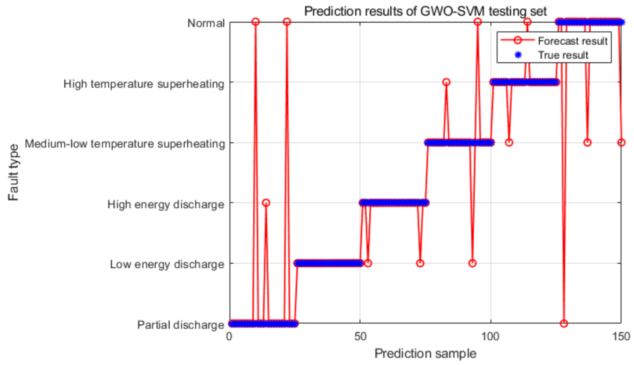

Fig. 6. GWO-SVM test results

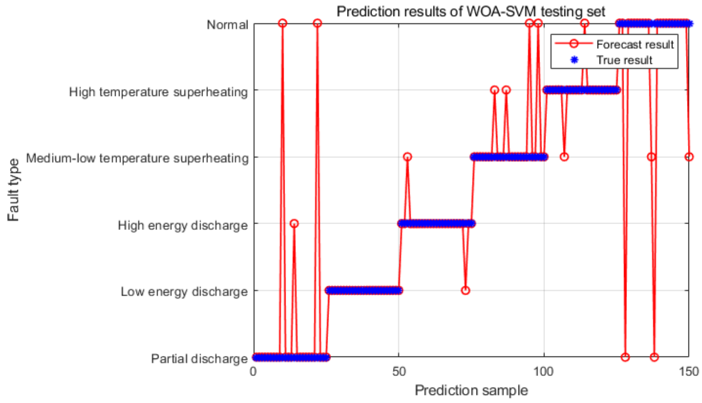

Fig. 7. WOA-SVM test results

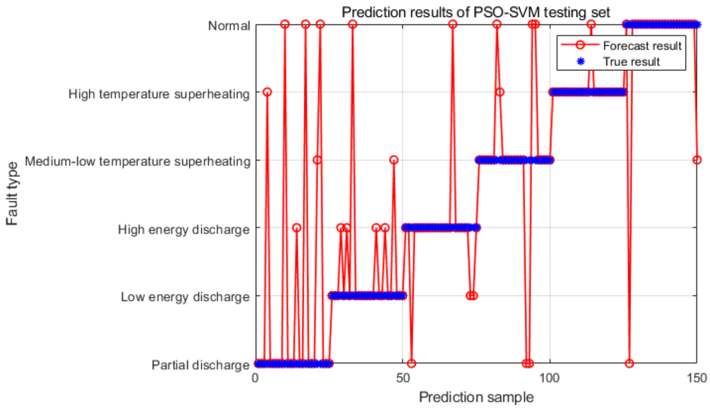

Fig. 8. PSO-SVM test results

Because the first 8 groups of HPO-SVM and GGO-SVM and the first 10 groups of WOA-SVM and PSO-SVM had the highest accuracy, the four groups of data were used to discuss the accuracy of different fault types, as shown in Table 4.

Table 4

Details of diagnosis accuracy of different fault types

Fault type	HPO-SVM/%	GWO-SVM/%	WOA-SVM/%	PSO-SVM/%
Partial discharge	96	88	92	76
Low energy discharge	100	100	100	76
High energy discharge	92	92	88	84
Medium-low temperature superheating	92	88	80	76
High temperature superheating	92	92	92	96
Normal	92	88	92	92

To explore the iterations of different SVM models in finding the optimal solution, this paper increases the search range, raises the maximum number of iterations to 50, sets the upper limit of the optimization parameter objective to 500, and increases the number of populations to 40. the red line in the figure is the HPO-SVM fitness curve, which completes its iteration in 7 generations, the orange line is the GWO-SVM fitness curve with 22 final iterations, and the green line is the PSO-SVM fitness curve with 11 final iterations. The number of iterations of WOA-SVM is 5, but its adaptation degree is larger than that of HPO-SVM model. Combined with the above experimental results, the HPO-SVM model has significant superiority in terms of adaptation degree and adaptation rate when compared with the other two optimization algorithms. Fig. 9 shows the adaptation curves of HPO-SVM, GWO-SVM, WOA-SVM, and PSO-SVM.

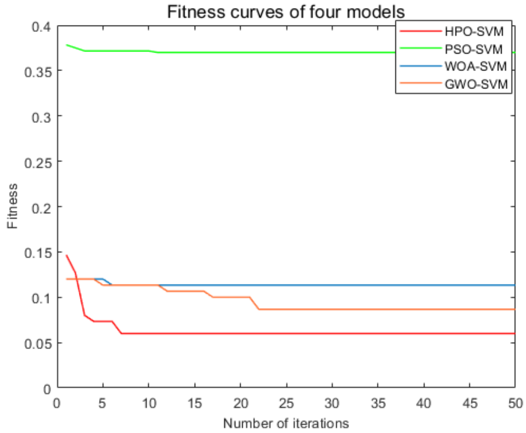

Fig. 9. Fitness results of the four models

The abscissa of fitness at a certain iteration time is the error rate of model prediction results. To further demonstrate the superior capability of the HPO-SVM models in terms of reliability, the three sets of models are run 10 times under the conditions above, respectively, and the correctness of the runs and the number of iterations of the locally optimal solutions are extracted. This is shown in Table 5. Comparing the accuracy, optimal number of iterations, average value and variance of the four optimization algorithms, it is easy to find that HPO-SVM is superior to the other algorithmic models in all data, and the stability rate is improved significantly.

Table 5

Details of diagnosis accuracy of different fault types

	HPO-SVM		GWO-SVM		WOA-SVM		PSO-SVM	
	Accuracy /%	Numbers	Accuracy /%	Numbers	Accuracy /%	Numbers	Accuracy /%	Numbers
1	94	6	92	9	89.33	15	82.67	8
2	94	11	92	12	90	2	82	20
3	94	8	88.67	3	90	3	84	13
4	94	4	90	3	90	20	84	13
5	94	3	92	15	90	13	87.33	19
6	94	7	89.33	13	89.33	18	83.33	18
7	94	6	91.33	10	90	10	84	14
8	94	5	92	4	90	13	82.67	13
9	94	4	90	9	90	11	81.33	16
10	94	10	92	11	89.33	20	82.67	17
Ave	94	6.4	90.93	8.9	89.80	12.5	83.4	15.1
Var	0	6.24	1.53	16.29	0.09	35.85	2.45	11.69

5. Conclusions

In this paper, a transformer fault identification method based on the reliefF and HPO-SVM model is described. Experiments prove that this method can effectively identify common oil-immersed transformer faults such as partial discharge, low-energy discharge, and high-energy discharge. The main research findings are as follows.

1) Compared with IEC three-ratio method and uncoded nine-ratio method, reliefF feature weight method can more effectively extract several groups of ratios with the highest feature weight from DGA fault samples. The ratio samples have good fault stripping performance, which is conducive to improving the recognition accuracy of the optimization model.
2) Compared with the support vector machine models based on GWO, WOA and PSO, the accuracy of the optimization strategy model based on HPO is higher. After the feature weight method of reliefF is integrated to simplify the

overall feature ratio set and reduce the feature dimension, the prediction accuracy is further improved.

3) Through experiments on the stability of several optimization algorithms, simulation experiments prove that the proposed HPO algorithm has better fault detection ability and has certain social practical value.

REFERENCES

[1] *T. F. Yang, P Liu, J I Li and Y Hu*. "FCM New Fault Diagnosis Method of Power Transformer by Combination of FCM and IEC Three-ratio Method". High Voltage Engineering, **vol .33**, no .8, pp. 66-70, Aug. 2007.

[2] *J. L. Yu and S. Bian*. "Fault Diagnosis Model of Transformer Based on BP Neural Network". Proceedings of the Csee, **Vol. 26,** no. 6, pp. 109-112, Jun. 2014.

[3] *X. L. Hu*. "Transformer Protection Based on RBF Neural Networks". Journal of System Simulation, **vol.** 24, no. 10, pp. 64-67.

[4] *J. F. Li*. "Fault Diagnosis of Transformer Based on Recurrent Neuron Network and Bat Algorithm". Electric Engineering, **vol.** 10, no. 20, pp. 38-41.

[5] *Y. Zhang and X. Z. Zhou*. "Modified grey wolf optimization algorithm for global optimization problems". Journal of University of Shanghai for Science and Technology, **vol.** 43, no. 1, pp. 73-82.2021.

[6] *D. G. Xu, Z. Q. Wang, Y. X. Guo and K. Z. Xing*. "Review of whale optimization algorithm". Application Research of Computers, **vol .40**, no .2, pp. 66-70, Feb . 2023.

[7] *A. G. Li, Z. Qin, F. M. Bao and S. P. He*. "Particle Swarm Optimization Algorithms". Computer Engineering and Applications, **vol .17**, pp. 1-3.

[8] *L. X. Zhang, J. X. Wang, ZHAO, Y. N and Z. H. YANG*. "Combination Feature Selection Based on Relief ". Journal of Fudan University, **vol .43**, no .5, pp. 893-898, Oct . 2004.

[9] *S. F. Ding, B. J. Qi and H. Y. Tan*. "An Overview on Theory and Algorithm of Support Vector Machines" . Journal of University of Electronic Science and Technology of China, **vol .40**, no .1, pp. 2-10, Jan. 2011.

[10] *Sai Ram Inkollu1, G.V. Prasanna Anjaneyulu, Kotaiah NC and Nagaraja Kumari.CH* "An Application of Hunter-Prey Optimization for Maximizing Photovoltaic Hosting Capacity Along with Multi-Objective Optimization in Radial Distribution Network." International Journal of Intelligent Engineering & Systems, **vol .15**, no .4, pp. 575-584, June 13 . 2022.

[11] *J. Y. Kang, S. X. Zhang, Q. P. Zhang, B. Gao, Z. H. Yan, H. Z. Cheng*. "Fault Diagnosis Method of Transformer Based on ANOVA and BO-SVM". High Voltage Engineering, **vol .48** Oct 28. 2022

[12] *T. Mei*. "Research on Power Transformer Fault Diagnosis Method Based on Improved ISOMAP and WkNN". Chinese Master's Theses Full-text Database,2020. **vol .01**, 2021.

[13] *C. Sun*. Transformer Fault Diagnosis Based on Machine Learning Algorithm. Chinese Master's Theses Full-text Database,2019. **vol .06**, 2020.