Fault Diagnosis Method for Transformer Based on KPCA and SSA-SVM

Xiaohui Han[1], Shifeng Ma[1], Zhewen Shi[1], Guoqing An[1(✉)], Zhenbin Du[2], and Chunlin Zhao[3]

[1] College of Electric and Engineering, Hebei University of Science and Technology, Shijiazhuang 050018, China
anguoqing99@sina.com

[2] Hebei Provincial Key Laboratory of Electromagnetic and Structural Performance of Power Transmission and Transformation Equipment, Baoding Tianwei Baobian Electric Co., Ltd., Baoding 071056, China

[3] Hebei Weixun Electric Power Automation Equipment Co., Ltd., Hengshui 053000, China

Abstract. For the sake of integrate the ratio of dissolved gas in the oil and the advantages of artificial intelligence technology to enhance the accuracy rate of transformer fault diagnosis, the fault diagnosis method for transformer based on kernel principal component analysis (KPCA) and sparrow search algorithm (SSA)-support vector machine (SVM) was proposed. Firstly, based on the oil dissolved gas analysis (DGA), 24 fault features of the power transformer were extracted, Secondly, KPCA was used for dimensionality reduction to obtain a feature space with lower latitude. The fault diagnosis model of transformers based on SVM was designed with the 8 selected features as inputs, and the parameters in the model were simultaneously optimized by SSA. Then, SSA-SVM was adopted to diagnose the typical working conditions. To prove the superiority of the SSA-SVM diagnostic model combined with KPCA feature space, the comparative experiment of SSA-SVM classifier results in the origin feature space, the KPCA feature space was carried out, the comparison for accuracy of various methods in the KPCA feature space was proceed as assist.

Keywords: Dissolved gas analysis · Power transformers · Fault diagnosis · Kernel principal component analysis · Sparrow search algorithm

1 Introduction

The power transformer is one of the most significant electrical equipment in the electric power system. It undertakes the most important task of electric power transmission and voltage change. It will cause local or even large area power cut when it breaks down. And it must be sure to cause a huge financial losses [1]. Consequently, establishing the efficient and accurate transformer fault diagnosis model and mastering its operation state are of significant meaning to enhance the operational level of power transformer and guarantee the secure and dependable operation of power systems [2].

In the aspect of transformer fault diagnosis, Dissolved Gas Analysis (DGA) technology has become one of the most valid means for fault diagnosis of oil-immersed transformers [3]. At present, the transformer fault diagnosis methods based on DGA mainly include traditional methods and methods based on artificial intelligence. Traditional methods include main gas method [4], IEC ratio method [5], Rogers method [6] and David triangle method [7] and so on. All of them have advantages of simple principle, small amount of calculation and easy implementation. So, they play a crucial role in the transformer fault diagnosis. However, these methods are based on on-site experience, so there will be a large error in diagnosis [8].The methods based on artificial intelligence mainly include neural network [9], random forest [10], support vector machine [11, 12] and so on. Compared with traditional means, the above diagnosis methods have the advantages of continuous learning and updating. It can establish the complex curvilinear relations between the transformer fault and DGA gas content. Although the calculation accuracy of traditional methods are improved, there are also some limitations as follow. There are some problems in neural network, such as slow convergence and complex system structure [13]. Random forest has the disadvantages of long training time and over-fitting. Compare with the above two methods, support vector machine has strong generalization performance and can deal with local minimum well. Nonetheless, the diagnosis capability of SVM is influenced by penalty factor and kernel parameter. The diagnosis result will be a great error if the value is selected improperly.

When using artificial intelligence technology for transformer fault diagnosis, most studies are mainly founded on existing DGA ratios (such as IEC, Rogers) as the input characteristic parameter of artificial intelligence model to establish transformer fault diagnosis model. However, there is no unified standard of DGA gas ratio for diagnosis models. According to the literature [14], the effect of transformer fault diagnosis will be affected if use the partial DGA gas ratio. Therefore, in order to combine the advantages of DGA ratio and artificial intelligence methods [15], KPCA and SSA-SVM are proposed in this paper. In this method, 24 characteristic parameters of transformer are extracted on the basis of DGA. Dimensionality reduction and optimization of characteristic parameters by using KPCA. The data after dimensionality reduction were used as samples to train the SVM model. While SSA was used to optimize the model parameters. The transformer fault diagnosis means founded on KPCA and SSA-SVM is established. The real-time and accurate diagnosis of power transformer operation state will be realized.

2   Kernel Principle Component Analysis

Kernel Principle Component Analysis (KPCA) is an algorithm integrating principle component analysis and kernel method. It can make up for the defect that PCA can't extract the nonlinear structural characteristics. Finally, the nonlinear dimensionality reduction data samples will be realized [16].

The basic idea of KPCA is to transform a nonlinear problem in the input space into a linear problem in the mapping space by introducing a certain nonlinear mapping function [17, 18]. The nonlinear mapping is realized by using the kernel inner product arithmetic in the input space. RBF is used as kernel function in this paper. Suppose to input the spatial dataset $X = \{x_i\}$, $i = 1, 2, 3\ldots, N$, the x_i is a vector of m * 1. Each observation

has m characteristics. The RBF kernel matrix is calculated as:

$$K(x_i, x_j) = exp\frac{-\|x_i - x_j\|^2}{2\sigma^2} (i, j = 1, \ldots, N) \tag{1}$$

Centralize the kernel matrix, the centralization formula is:

$$\overline{K} = K - P_m K - K P_m + P_m^T K P_m \tag{2}$$

In the formula, P_m refers to the m × m matrix whose elements are all 1/m.

Next, solve the eigenvalues and eigenvectors of the central kernel matrix $\overline{K}$:

$$\overline{K}v = (AA^T)v = \lambda v \tag{3}$$

Make sure the s largest eigenvalues λ_j (j = 1,2…s)and its corresponding eigenvectors v_j (j = 1,2…s). The valve of s is defined judging by the cumulative variance contribution rate. The formula is as follow:

$$\frac{\lambda_1 + \lambda_2 + \lambda_3 \ldots \lambda_s}{\lambda_1 + \lambda_2 + \lambda_3 \ldots \lambda_m} \geq 85\% \tag{4}$$

The cumulative variance contribution rate, that is the eigenvalue corresponding to the eigenvector as the dimensionality reduction weight. The matrix after the dimensionality reduction is:

$$Z = \overline{K}^T v = \overline{K}^T \left[\frac{1}{\sqrt{\lambda_1}} v_1, \frac{1}{\sqrt{\lambda_2}} v_2 \cdots \frac{1}{\sqrt{\lambda_s}} v_s \right] \tag{5}$$

3 SVM Parameter Optimization Based on Sparrow Search Algorithm

3.1 Support Vector Machine

SVM is a model with strong generalization ability and is suitable for small specimen learning [19]. It has been extensive used in the domain of fault diagnosis.

SVM was originally proposed to study the linear separable binary classification problem. The most of the cases are nonlinear in fact. The samples of the original input space need to be mapped into the high dimensionality feature space through the nonlinear mapping function, after that the optimum classification super function will be constructed in the space. In this time, the optimization problem is transformed into:

$$\begin{cases} \begin{matrix} min \\ \omega, b, \xi \end{matrix} & \frac{1}{2}\omega^2 + C\sum_i^n \xi_i \\ s.t. & y_i(\omega \cdot \phi(x) + b) \geq 1 - \xi_i \\ & \xi_i \geq 0, i = 1, 2 \ldots n \end{cases} \tag{6}$$

In this formula,ω is the weight coefficient, C is the penalty parameter. ξ_i is the slack variable, which the classification error of training samples was characterized; b is the partial value constant.

The form of the dual optimization is as follows:

$$
\begin{cases}
\max\limits_{\alpha} \ \sum\limits_{i=1}^{n} \alpha_i - \frac{1}{2} \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} \alpha_i \alpha_j y_i y_j \left(K\left(x_i, x_j\right)\right) \\
s.t. \ \begin{array}{l} \sum\limits_{i=1}^{n} \alpha_i y_i = 0 \\ 0 \le \alpha_i \le C, i = 1, 2 \ldots n \end{array}
\end{cases}
\tag{7}
$$

In this formula, α_i is the Lagrange multiplier, $y_i \in \{-1, 1\}$ is the category label. Thus, the optimal classification function of the nonlinear problem is:

$$
f(x) = \text{sgn}(\sum_{i=1}^{n} \alpha_i^* y_i K(x, x_i) + b^*)
\tag{8}
$$

$K(x_i, x_j)$ is the kernel function, different kernel functions can be selected to construct different classification performances of SVM. The common kernel functions are: Linear kernel function, polynomial kernel function, RBF and Sigmoid. The RBF is easier than others, and the classification effect is better, so the RBF was used in the paper. The expression is as follows:

$$
K(x, x_i) = exp\left(-g\|x - x_i\|^2\right), g > 0
\tag{9}
$$

In this formula, the kernel parameter g is a positive real number.

3.2 SSA Algorithm

3.2.1 Algorithm Principle

SSA is a novel swarm intelligence majorization algorithm enlightened with foraging behavior and anti-predation behavior of sparrows in nature [20, 21].

In SSA, each sparrow position corresponds to a feasible solution. Supposes that in a D-dimensional search space, there is a population of N sparrows: $X = [x_{i1}, x_{i2}, \ldots, x_{id}]$, $i = 1, 2, \ldots n$, x_{id} indicates the location of sparrow i in d-dimension. When sparrows are foraging, there are 3 behavior.

Finders lead the population to search and forage, so the search area is wide. They use the memory to constantly update their location in order to obtain the food sources. Finders take up 10%–20% of the population. And its location is updated as follows:

$$
X_{i,j}^{t+1} =
\begin{cases}
X_{i,j}^t \cdot exp\left(-\dfrac{i}{\alpha \cdot iter_{max}}\right) & if \ R2 < ST \\
X_{i,j}^t + Q \cdot L & if \ R2 \ge ST
\end{cases}
\tag{10}
$$

In this formula, X_{ij} means the location of sparrow i in j-dimension. t means the present Iteration ordinal number. $iter_{max}$ means the maximum iteration ordinal number. α is a uniform random number between(0,1]. $R2 \in [0,1]$ represents the warning value. $ST \in [0.5,1]$ represents the safe value, Q is a stochastic number that follows a normal distribution. L indicates the array of $1 \times d$. The elements of this array are all 1.

In order to obtain the higher fitness, the new entrants follow to the finders to forge constantly. The location is updated as follows:

$$
X_{i,j}^{t+1} = \begin{cases} Q \cdot exp\left(\dfrac{X_{worst}-X_{i,j}^{t}}{i^2}\right) & if\ i > n/2 \\ X_p^{t+1} + \left|X_{i,j}^{t} - X_p^{t+1}\right| \cdot A^+ \cdot L & otherwise \end{cases} \tag{11}
$$

X_p expresses the optimal location occupied by the finders during the $t+1$th iteration. X_{worst} expresses the global worst location of the tth iteration. i expresses the population size. A represents the matrix of the $1 \times d$. Each element is randomly evaluated of 1 or -1. $A^+ = A^T(AA^T)^{-1}$.

Scouts watch for the predators' threats and warn the population to anti predation. The location is updated as follows:

$$
X_{i,j}^{t+1} = \begin{cases} X_{best}^{t} + \beta \cdot \left|X_{i,j}^{t} - X_{best}^{t}\right| & if\ f_i > f_g \\ X_{i,j}^{t} + K \cdot \left(\dfrac{\left|X_{i,j}^{t}-X_{worst}^{t}\right|}{(f_i-f_w)+\varepsilon}\right) & if\ f_i = f_g \end{cases} \tag{12}
$$

X_{best} is the global best position at present. β means the step length control parameters. It is a stochastic number that follows a normal distribution with a mean value of 0 and a variance of 1. K shows the orientation in which the sparrows move. It is a stochastic number between $[-1,1]$. f_i expresses the fitness value of the i sparrow. f_g represents the optimum fitness value of the sparrow population. f_w means the worst fitness value of the sparrow population. ε is a minimal constant, in order to avoid the situation of the denominator is 0.

3.2.2 Algorithm Performance Test

To verify the effectiveness of the SSA, test functions Rosenbrock and Ackley were selected to test its performance, and compared with PSO and GA for analysis.

Rosenbrock only has one global optimal value, it doesn't have the local value. Just like Fig. 1, the formula is as follows:

$$
F_1(x) = \sum_{i=1}^{n-1}\left[100\left(x_{i+1} - x_i^2\right)^2 + (x_i - 1)^2\right] \tag{13}
$$

x $\in[-30,30]$, the global minimum in the domain is 0.

There are a great quantity of local optimum values in the outer region of the Ackley. The center concave position is the global optimal value of the whole function. Just like Fig. 2, the formula is as follows:

$$
F_2(x) = -20exp(-0.2\sqrt{\frac{1}{2}\sum_{i=1}^{n}x_i^2}) - exp(\frac{1}{n}\sum_{i=1}^{n}cos(2\pi x_i)) + 20 + e \tag{14}
$$

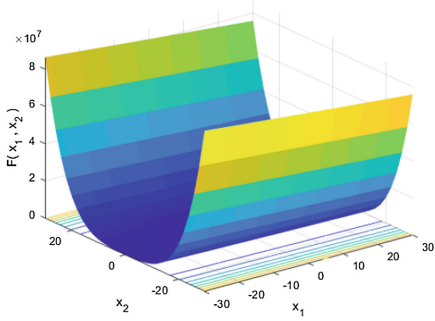x $\in[-32,32]$, the global minimum in the domain is 0.
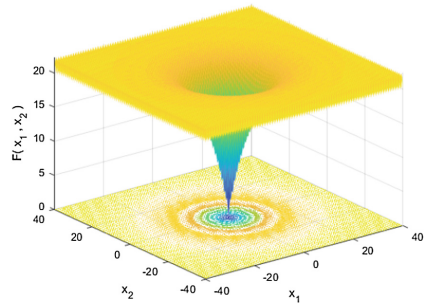
Fig. 1. TryItStatic F1

Fig. 2. TryItStatic F2

SSA,PSO,GA were used to conduct 30 independent simulation experiments on F1 and F2 test functions. The maximum number of the iterations of each algorithm is 1000. Population size is set to N = 100. The fitness change curves of the 3 algorithms during the optimization of F1 and F2 test functions are shown in the Fig. 3 and 4. Table 1 lists the comparison of test accuracy and stability results of the 3 algorithms for the 2 test functions.
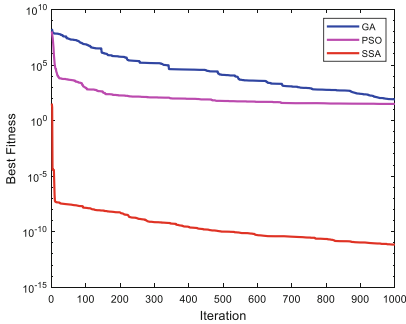

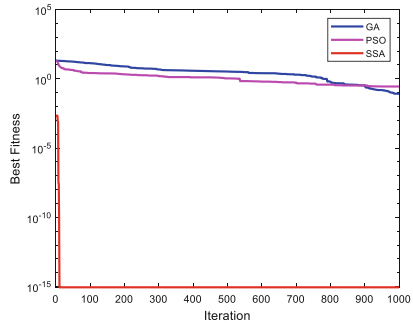
Fig. 3. F1(x) Function Fitness Change Curve

Fig. 4. F2(x) Function Fitness Change Curve

According to the F1(x) Function Fitness Change Curve in Fig. 3, although SSA did not find the optimal fitness value on F1(x), the fitness value is better than PSO, GA obviously.

According to the F2(x) Function Fitness Change Curve in Fig. 4, SSA had reached the optimal fitness value after 13 iterations on $F_2(x)$. However, PSO needs 922, GA needs more.

According to the results of the optimal value and average value in Table 1, we can know that the constriction accuracy of SSA is better than PSO and GA on test functions F1 and F2 obviously. SSA algorithm has better stability compared with PSO and GA. We can get the results from the worst difference, median and Std. Deviation.

Table 1. Comparison of test accuracy and stability

Function	Algorithm	Optimal value	Worst value	Median	Average value	Std
F1	PSO	31.6539	381.6988	81.1324	108.5045	87.156
	GA	80.6832	1405.9881	222.5685	286.2865	249.0589
	SSA	6.811e−12	3.0416e−07	7.0691e−09	5.2461e−08	9.4188e−08
F2	PSO	0.2752	2.4193	1.7369	1.7347	0.52752
	GA	0.082683	1.2485	0.16271	0.26456	0.25981
	SSA	8.8818e−16	8.8818e−16	8.8818e−16	8.8818e−16	0

In conclusion, we can get that SSA algorithm is better than GA and PSO on the part of convergence speed, convergence accuracy and stability through the test of Rosenbrock and Ackley.

3.3 Sparrow Search Algorithm Optimizes SVM Parameters

From the foregoing analysis, the punishment parameter C and kernel function parameter g of SVM are the main factors that affect its fault diagnosis results. In this paper, SSA is used to seek out the best C and g in the training stage of SVM, so as to enhance the fault diagnosis rate of SVM. The steps of SSA optimizes SVM parameters are as follows:

(1) SSA parameters are initialized. Including sparrow population Pop, search space dimension D, maximum iteration number $iter_{max}$, proportion of sparrows aware of danger SD, proportion of Discovers PD, warning value ST.
(2) Compute the fitness of every sparrow to find the position X_{best} corresponding to the best fitness value and the position X_{worst} corresponding to the worst fitness value at present.
(3) According to the formula (10), (11), and (12), update the positions of finders, entrants and sparrows aware of danger.
(4) Recalculate the fitness value of every sparrow after position renew and compare it with the fitness value in the last iteration. If it is higher than the original fitness value, the new location value is taken as the best fitness value X_{best}. Otherwise, the original fitness value remains unchanged.
(5) Determine whether the maximum number of iterations or solving accuracy is attained. If not, return to step (3). If so, the iterative process will be stopped and the sparrow position information of the best fitness will be returned. The optimal combination (C,g).

4 Transformer Fault Diagnosis Based on KPCA and SSA-SVM

4.1 Fault Type Classification and Feature Selection

According to IEC 60599 standard and some actual transformer operation data, the transformer fault mode will be divided into 6 types: medium low temperature overheating

T1($<$700 °C), high temperature overheating T2($>$700 °C), low energy discharge D1, partial discharge PD, high energy discharge D2 and normal working condition N.

When the transformer occurs the above thermal or electrical faults will lead to the decomposition of transformer oil and insulating paper. According to the "Oil-immersed transformer State Evaluation and Maintenance Guide" [22], the products include H2, CH4, C2H2, C2H4, C2H6 and other gases. Based on the above 5 gases, a total of 24 gas ratios are generated as the set of features to be selected, as shown in Table 2.

Table 2. Gas ratios dissolved in oil

Number	DGA ratio	Number	DGA ratio
1	H_2	13	H_2/TH
2	CH_4	14	$CH_4/(SUM)$
3	C_2H_6	15	$C_2H_4/(SUM)$
4	C_2H_4	16	$C_2H_2/(SUM)$
5	C_2H_2	17	C_2H_2/C_2H_4
6	CH_4/C_2H_6	16	C_2H_4/C_2H_6
7	CH_4/C_2H_4	19	$C_2H_2/(C1 + C2)/\%$
8	CH_4/TH	20	$H_2/(H_2 + C1 + C2)/\%$
9	C_2H_6/C_2H_4	21	$C_2H_4/(C1 + C2)/\%$
10	C_2H_6/TH	22	$CH_4/(C1 + C2)/\%$
11	C_2H_4/TH	23	$C_2H_6/(C1 + C2)/\%$
12	C_2H_2/TH	24	$(CH_4 + C_2H_4)/(C1 + C2)/\%$

(Note: SUM is the sum of CH_4, CH_2, and C_2H_4 contents, TH is the sum of 5 gases, and (C1 + C2) is the sum of the CH_4, C_2H_6, C_2H_4 and C_2H_2.)

4.2   Concrete Realization Process of Transformer Fault Diagnosis

Figure 5 shows the realization procedure of transformer fault diagnosis means founded on sparrow search algorithm to optimize SVM parameters. The specific steps are as follows:

(1) DGA data of various fault patterns during transformer operation are gathered to compose fault sample data;

(2) The numerical dispersion between different DGA ratios maybe large was considered, it may affect the diagnosis is accuracy if it is input directly as a sample. Normalized processing is carried out for all fault sample data according to formula (15).

$$x_{si} = \frac{x_i - x_{imin}}{x_{imax} - x_{imin}} \tag{15}$$

In this formula, x_{imin}, x_{imax} is the minimum and maximum value before normalization, x_i is the value before normalization.

(3) KPCA dimension reduction is performed on normalized sample data to eliminate redundant gas ratio that affects fault diagnosis accuracy. In this paper,KPCA non-linear dimension reduction technique based on RBF kernel is selected to lessen the dimension of dissolved gas ratio in transformer oil according to formula (1)-(5).

(4) Sample data are randomly separated into training set and test set in proportion. The data after dimensionality reduction is chose as the feature import of the model.

(5) Combined with SSA, a transformer fault diagnose model based on SSA to optimize SVM parameters was established.

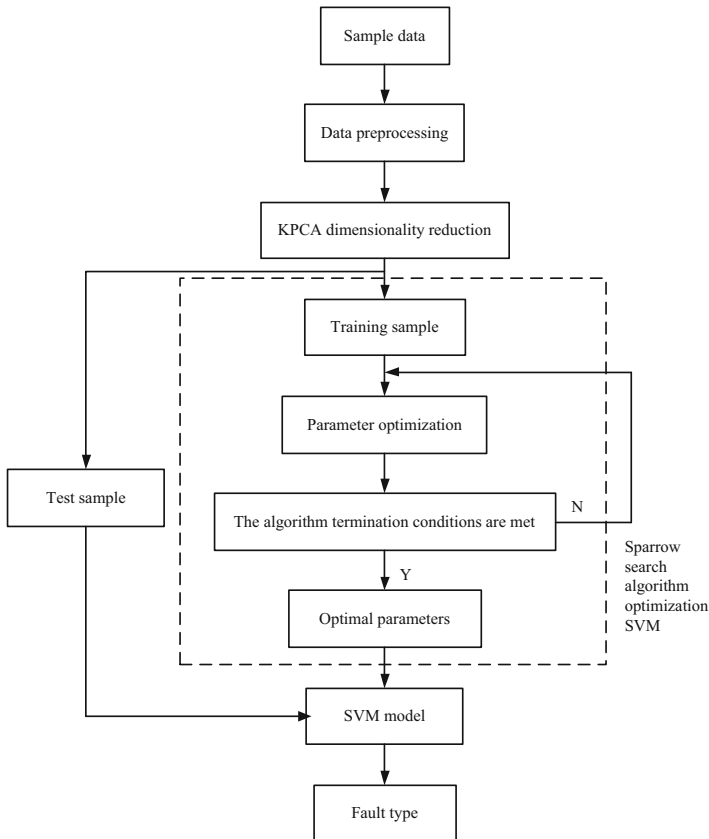(6) Test sets were used to test the diagnosis effect of SSA-SVM model and output fault diagnosis results.

Fig. 5. Fault diagnosis flowsheet of SSA optimized SVM

5   Example Analysis

5.1   Fault Sample and Data Pretreatment

In this paper, 240 transformer DGA data with determined state types are selected from reference [23] for fault diagnosis of SSA-SVM model. Firstly, 240 transformer DGA data are converted into the ratios in Table 2 to obtain a 240 * 24 dimensional DGA ratio set. Normalization pretreatment was performed on all ratios according to formula (15) to eliminate differences between different values.

5.2   Analysis of Feature Value Optimization Results Based on Kernel Principal Component Analysis

The normalized 240 * 24 dimensional DGA ratio was optimized for dimensionality reduction, the characteristic values and feature vector of its RBF kernel matrix were calculated, and the characteristic values whose cumulative variance contribution ratio was greater than 85% were selected for analysis to achieve dimensionality reduction. The eigenvalue and variance contribution ratio of RBF kernel matrix are shown in the table, and the variation trend of variance contribution rate with components is shown in Fig. 6.

Table 3. Eigenvalue and cumulative variance contribution ratio

Principal component	Eigenvalue	Variance contribution rate/%	Cumulative variance contribution ratio /%
1	0.0523	27.5003	27.5003
2	0.0376	19.7915	47.2918
3	0.0225	11.8084	59.1002
4	0.0194	10.2115	69.3117
5	0.0131	6.8765	76.1882
6	0.0084	4.4162	80.6044
7	0.0059	3.1211	83.7255
8	0.0051	2.6851	86.4106
9	0.0041	2.1733	88.5839
10	0.0036	1.8768	90.4607

As can be seen from Table 3 that the cumulative variance contribution of the eight components is 86.4106%. It means these eight components can represent the internal fault characteristics of the transformer. It can be seen from Fig. 6, the change trend of variance contribution ratio progressively decreases, and the variance contribution rate after 8 components has stabilized. Therefore, taking the first 8 components can reflect most of the variable information, which is transformed from 24 indicators to 8 new indicators, so as to reach the aim of dimension lessen.
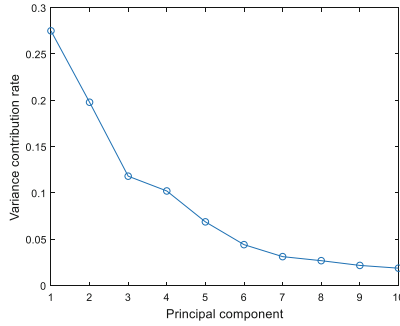
Fig. 6. Trend chart of contribution rate with the principal components changing

Further, the distribution of the first principal component in different fault types and normal working conditions in the original feature space and KPCA space is plotted as shown in Fig. 7. It can be displayed from Fig. 7a that the distribution of the first principal component in the original feature space is disordered and there is serious aliasing, which affects the diagnostic performance of the subsequent model; Comparing Fig. 7b with Fig. 7a, it can be seen that the distance between the first principal components becomes larger and the aliasing phenomenon is significantly reduced under the six working conditions after dimensionality reduction. It verifies the effectiveness of KPCA dimensionality reduction method.
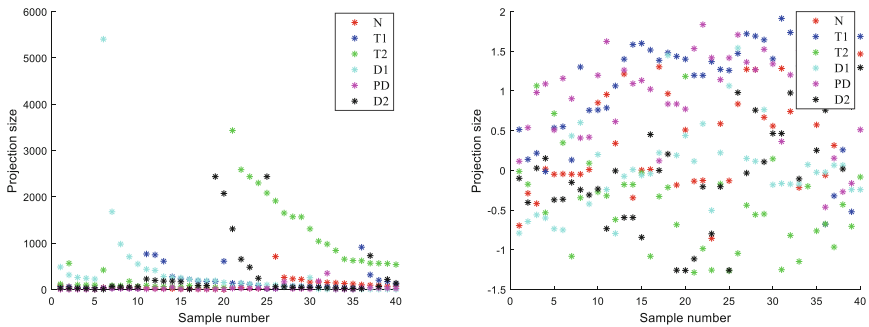
(a)   Distribution in original feature space   (b)   Distribution in KPCA feature space

Fig. 7. Distribution of six states in the origin and the KPCA feature space

5.3 Parameter Optimization

150 groups of failure samples were randomly selected as training samples (25 groups for each type of failure), and the rest of the 90 groups as test samples (15 groups for each type of failure). SSA was used to optimize parameters in the SVM model, and relevant parameters were set as follows: The search range of SVM kernel parameters

and penalty factor is [0.01,100], population size Pop = 50, maximum iteration times $iter_{max} = 100$, search space dimension $d = 2$, the sparrows aware of danger SD = 0.2, sparrows discovers PD = 0.7, warning value is 0.6. During the optimization of SVM parameters by sparrow search algorithm, the curve of fitness change is shown in Fig. 8.
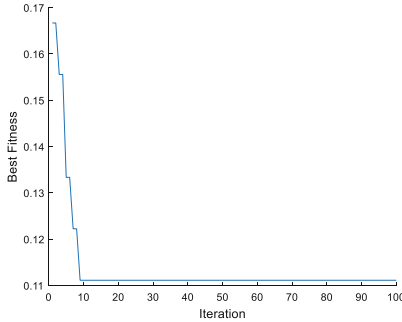

Fig. 8. Fitness change curve

As can be seen from Fig. 8, when the sparrow search algorithm is employed to optimize the SVM parameters, the optimal fitness curve gradually converges with the increase of the iterations, and converges to the minimum after 9 iterations. It shows that sparrow search algorithm can better resolve the optimization matter of SVM parameters.

5.4 Contrast of Diagnostic Results of Diverse Diagnostic Models in KPCA Dimensionality Reduction

To verify the superiority of the method in the paper, firstly, the diagnosis effects of SSA-SVM in original and KPCA feature spaces are compared, Table 4 shows the diagnosis outcomes of SSA-SVM model in two feature spaces.

Table 4. Eigenvalue and cumulative variance contribution rate

Type	Original feature space	KPCA
N	86.67	86.67
T1	93.33	93.33
T2	100	93.33
D1	60	93.33
PD	86.67	86.67
D2	93.33	80
Mean value of accuracy/%	86.67	88.89
Time/s	13.29	9.91

According to the data in Table 4, the mean diagnosis accuracy of SSA-SVM in KPCA feature space is 2.22% higher than that in original feature space, and the diagnosis time is reduced by 3.38 s. Through the comparison with that, the SSA-SVM diagnosis method in KPCA feature space is better than the original feature space in both effect and time.

5.5 Comparison of Diagnostic Results of Different Diagnostic Models in KPCA Dimensionality Reduction

In Sect. 4.4, the diagnostic effects of SSA-SVM in original and KPCA feature space are compared. In this section, SVM, GA-SVM, PSO-SVM are selected for comparing with SSA-SVM in KPCA feature space. Different fault diagnosis results of each diagnosis model are contrasted as shown in Table 5.

Seen from Table 5 that in KPCA feature space, the mean diagnostic accuracy of SVM, PSO-SVM, GA-SVM and SSA-SVM are 83.33%, 84.44%, 84.44% and 88.89% respectively; The diagnostic time of them are 0.14 s, 35.79 s, 19.13 s and 9.91 s respectively. It shows that the fault diagnosis outcomes of this means are more accurate than other diagnosis models in KPCA feature space.

5.6 Influence of the Number of Training Samples on Fault Diagnosis Outcomes

To further explain the availability of SSA-SVM, the number of training samples of each type was reduced to 4/5 and 3/5 respectively based on the original training samples. Then the fault diagnosis model was trained and tested. The outcomes are shown in Table 6.

Table 5. Comparison outcomes using various models in the KPCA feature space

Type	Accuracy/%			
	SVM	PSO-SVM	GA-SVM	SSA-SVM
N	86.67	86.67	86.67	86.67
T1	86.67	93.33	93.33	93.33
T2	100	86.67	86.67	93.33
D1	80	93.33	86.67	93.33
PD	93.33	86.67	86.67	86.67
D2	53.33	60	66.67	80
Mean value of accuracy/%	83.33	84.44	84.44	88.89
Time/s	0.14	35.79	19.13	9.91

From Table 6 we can see that the diagnosis accuracy of various types of faults is less affected by the reduction of the quantity of training samples. In the case of the test samples in Table 5, when the quantity of various training samples is reduced, the overall fault diagnosis accuracy of SSA-SVM model is still better than that of SVM,

Table 6. Fault diagnosis results after reducing the training samples

Type	Test set	4/5 Training sample		3/5 Training sample	
		Number of misjudgments	Accuracy/%	Number of misjudgments	Accuracy/%
N	15	2	86.67	3	80
T1	15	0	100	1	93.33
T2	15	0	100	0	100
D1	15	3	80	4	73.33
PD	15	0	100	0	100
D2	15	7	53.33	5	66.67
Total	90	12	86.67	13	85.56

GA-SVM and PSO-SVM models in Table 5. It shows that the transformer fault diagnosis model based on SSA-SVM has the merits of good robustness and strong generalization capability. It can still obtain high fault diagnosis accuracy with less training samples.

6   Conclusion

This paper proposes a transformer fault diagnosis method founded on the kernel principal component analysis and the sparrow search algorithm optimization support vector machine, and the main conclusions are as follows:

KPCA is used to reconstruct the feature space and reduce the dimension, and the effective features of transformer fault diagnosis are obtained. Compared with the original feature space, the diagnosis accuracy is improved by 2.22%, the diagnosis time is reduced by 3.38 s, and high-precision diagnosis results are obtained.

Using sparrow search algorithm to optimize SVM parameters can improve the accuracy of SVM fault diagnosis model. Compared with SVM, PSO-SVM and GA-SVM fault diagnosis models, the fault diagnosis model established in the paper has a greater test accuracy (the accuracy rates are 88.89%, 83.33%, 84.44%, 84.44%), which proves the validity of the method in this paper.

In case of training sample reduction, the optimization of SVM fault diagnosis model parameters by sparrow search algorithm can still obtain high fault diagnosis accuracy, which is better than the accuracy of SVM, PSO-SVM and GA-SVM models without training sample reduction. It shows that SSA-SVM model is less affected by the reduction of the quantity of training samples. It proves the generalization of the method in this paper.

References

1. Zhao, J., Wen, F., Xue, Y., et al.: Architecture of power CPS and its implementation technology and challenges. Autom. Electr. Power Syst. **34**(16), 1–7 (2010). (in Chinese)
2. Han, J., Kong, X., Li, P., et al.: A novel low voltage ride through strategy for cascaded power Electronic transformer. Protect. Control Mod. Power Syst. **4**(3), 227–238 (2019)
3. Dai, J., Song, H., Sheng, G., et al.: Dissolved gas analysis of insulating oil for power transformer fault diagnosis with deep belief network. IEEE Trans. Dielectr. Electr. Insul. **24**(5), 2828–2835 (2017)
4. Ward, S.A.: Evaluating transformer condition using DGA oil analysis. In: 2003 Annual Report Conference on Electrical Insulation and Dielectric Phenomena. Albuquerque, USA, pp. 463–468. IEEE (2003)
5. Faiz, J., Soleimani, M.: Dissolved gas analysis evaluation in electric power transformers using conventional methods a review. IEEE Trans. Dielect. Electr. Insulat. **24**(2), 1239–1248 (2017)
6. Rogers, R.R.: IEEE and IEC codes to interpret faults in transformers, using gas in oil analysis. IEEE Trans. Electr. Insul. **13**(5), 349–354 (2007)
7. Bakar, N., Abu-Siada, A., Islam, S.: A review of dissolved gas analysis measurement and interpretation techniques. IEEE Electr. Insul. Mag. **30**(3), 39–49 (2014)
8. Shi, X., Zhu, Y., Ning, X., et al.: Transformer fault diagnosis based on deep auto-encoder network. Elect. Power Autom. Equip. **36**(5), 122–126 (2016). (in Chinese)
9. Ruan, L., Xie, J., Gao, S., et al.: Application of artificial neural network and information fusion technology in power transformer condition assessment. High Volt. Eng. **40**(3), 822–828 (2014). (in Chinese)
10. Cheng, Y., Zhang, Z.: Multi-source partial discharge diagnosis of transformer based on random forest. In: Proceedings of the CSEE 2018, vol. 38, no. 17, pp. 5246–5256+5322. (in Chinese)
11. Zhang, Y., Jiao, J., Wang, K., et al.: Power transformer fault diagnosis model based on support vector machine optimized by imperialist competitive algorithm. Electr. Power Autom. Equip. **38**(1), 99–104 (2018). (in Chinese)
12. Fang, T., Qian, Y., Guo, C., et al.: Research on transformer fault diagnosis based on a beetle antennae search optimized support vector machine. Power Syst. Protect. Control **48**(20), 90–96 (2020). (in Chinese)
13. Wu, J., Ding, H., Ma, X., Yan, B., Wang, X.: Application of improved adaptive bee colony optimization algorithm in transformer fault diagnosis. Power Syst. Protect. Control **48**(09), 174–180 (2020). (in Chinese)
14. Kim, S.W., Kim, S.J., Seo, H.D., et al.: New methods of DGA diagnosis using IEC TC 10 and related databases part 1: application of gas-ratio combinations. IEEE Trans. Dielectr. Electr. Insul. **20**(2), 685–690 (2013)
15. Tian, F., et al.: Transformer fault diagnosis model based on feature selection and ICA-SVM. Power Syst. Protect. Control **47**(17), 163–170 (2019). (in Chinese)
16. Li, J., Chang, Y.: Short-term probabilistic forecasting based on KPCA-KMPMR for wind power. Electr. Power Autom. Equip. **37**(2), 22–28 (2017). (in Chinese)
17. Wang, Y., Wu, J., Ma, S., et al.: Mechanical fault diagnosis research of high voltage circuit breaker based on kernel principal component analysis and Soft Max. Trans. Chin. Electrotech. Soc. **35**(Supplement1), 267–276 (2020). (in Chinese)
18. Li, D., Yang, B., Zhang, Y., et al.: Dimension reduction and reconstruction of multi-dimension spatial wind power data based on optimal RBF kernel principal component analysis. Power Syst. Technol. **44**(12), 4539–4546 (2020). (in Chinese)
19. Samanta, B., Al-Balushi, K.R., Al-Araimi, S.A.: Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection. Eng. Appl. Artif. Intell. **16**(7–8), 657–665 (2003)

20. Xue, J.K., Shen, B.: A novel swarm intelligence optimization approach: sparrow search algorithm. Syst. Sci. Control Eng. **8**(1), 22–34 (2020)
21. Lv, X., Mu, X., Zhang, J., et al.: Multi-threshold image segmentation based on improved sparrow search algorithm. Syst. Eng. Electron. 43(2), 318–327 (2021). (in Chinese)
22. Guide for condition evaluation of oil-immersed power transformers (reactors): DL/T 1685-2017. National Energy Administration, Beijing, China (2017)
23. Yin, J.: Research on fault diagnosis method of oil-immersed power transformer based on correlation vector machine. North China Electric Power University (2013)