Hindawi
Mathematical Problems in Engineering
Volume 2019, Article ID 8784154, 19 pages
https://doi.org/10.1155/2019/8784154

Research Article

Fault Diagnosis of Bearings Based on KJADE and VNWOA-LSSVM Algorithm

Tao Wu,[1] Chang Chun Liu,[1] and Cheng He[2]

[1]*Graduate Student, School of Environmental and Material Engineering, Shanghai Polytechnic University, Shanghai, Pudong 201209, China*
[2]*Associate Professor, School of Intelligent Manufacturing and Control Engineering, Shanghai Polytechnic University, Shanghai, Pudong 201209, China*

Correspondence should be addressed to Cheng He; 20181510007@stu.sspu.edu.cn

Received 6 July 2019; Revised 29 August 2019; Accepted 6 September 2019; Published 4 December 2019

Academic Editor: Qiuye Sun

In order to accurately diagnose the faulty parts of the rolling bearing under different operating conditions, the KJADE (Kernel Function Joint Approximate Diagonalization of Eigenmatrices) algorithm is proposed to reduce the dimensionality of the high-dimensional feature data. Then, the VNWOA (Von Neumann Topology Whale Optimization Algorithm) is used to optimize the LSSVM (Least Squares Support Vector Machine) method to diagnose the fault type of the rolling bearing. The VNWOA algorithm is used to optimize the regularization parameters and kernel parameters of LSSVM. The low-dimensional nonlinear features contained in the multidomain feature set are extracted by KJADE and compared with the results of PCA, LDA, KPCA, and JADE methods. Finally, VNWOA-LSSVM is used to identify bearing faults and compare them with LSSVM, GA-LSSVM, PSO-LSSVM, and WOA-LSSVM. The results show that the recognition rate of fault diagnosis is up to 98.67% by using VNWOA-LSSVM. The method based on KJADE and VNWOA-LSSVM can diagnose and identify fault signals more effectively and has higher classification accuracy.

1. Introduction

Rolling bearings are widely used in modern machinery and are one of the most vulnerable mechanical parts in rotating machinery [1]. Sometimes it is possible to determine the safety of the entire machine due to the state of motion of the bearing. Serious accidents may occur, directly threatening the safety of life and property, so researchers pay close attention to these accidents [2]. In order to reduce the damage caused by bearing failure, fault detection is very necessary and has great practical value [3]. At this stage, with the continuous development of artificial intelligence algorithms, optimization problems, and improved algorithm convergence, as well as improving system reliability and stability [4, 5], scholars in various fields have carried out in-depth research [6]. However, mechanical failure is a complex nonlinear time-varying system that is difficult to understand and estimate [7, 8], and its failure process is more complicated. How to perform rapid fault detection and diagnosis through optimization algorithms for specific mechanical faults has always been a challenge. Therefore, in-depth and scientific research on the rotating mechanical bearing parts will have great value and significance in industrial production or daily life [9–11].

1.1. Literature Review. The traditional bearing fault diagnosis process based on feature extraction mainly includes four steps: vibration signal acquisition, signal preprocessing, feature extraction, and fault identification. In the traditional bearing fault diagnosis method, the three steps of signal preprocessing, feature extraction, and fault identification are closely connected [12]. If any of the steps are not satisfactory, it will affect the final recognition effect. Since the 21st century, with the tremendous improvement of computer computing power, hardware support for deep neural

network training has been provided. At present, deep learning has become the research object of many scholars and has been successfully applied in the fields of image recognition [13], target detection [14], and computer vision [15]. Deep learning has the advantage of automatically extracting features, eliminating the need for cumbersome artifacts. Many scholars at home and abroad have used deep learning in bearing fault diagnosis research [16, 17] to achieve better results. For instance, Manjurul and Jong [18] used a discrete wavelet packet transform combined with an adaptive learning method to train CNN bearing fault diagnosis methods to improve the performance of the network and achieve high-precision diagnosis under multiple fault types. Du and Yang [19] proposed a method for demodulating the fault signal of a bearing based on the wavelet transform and Hilbert transform. The phase compensation method is used to refine the envelope spectrum to improve the frequency of the envelope spectrum. The resolution makes the recognition accuracy of bearing fault diagnosis improved. Shao et al. [20] proposed a bearing fault diagnosis method based on a double-tree complex wavelet packet transform and DBN. The dual-tree complex wavelet packet transform is used to refine the feature information, and the diagnostic effect is improved on the basis of the original DBN.

In the aspect of mechanical fault diagnosis based on empirical mode decomposition, Ali et al. [21] proposed a new fault diagnosis method based on EMD decomposition and energy entropy feature extraction. Firstly, multiple inherent modal components of the bearing-fault signal were obtained by EMD. The energy entropy characteristics of each component are obtained, and the fault type diagnosis is realized by the artificial neural network. Liu et al. [22] provided new ideas for the detection and fault diagnosis of rolling bearing-vibration signals. They use the EMD method to decompose the noise-reduced bearing-fault signal and calculate the AR spectrum characteristics of each intrinsic modal component to achieve bearing fault type diagnosis. Yang et al. [23] obtained the envelope spectrum of multiple scale components of the bearing-fault signal through empirical mode decomposition. The amplitude ratio characteristics are calculated for the envelope spectrum and used as feature vectors and are classified into classifiers for classification. Finally, an effective diagnosis of the inner ring and outer ring of the rolling bearing is realized. The SVM (Support Vector Machine) method [24] is a very important identification method in mechanical fault diagnosis. It is proposed for machine learning problems under finite samples and can effectively solve problems such as small samples and nonlinearities. Chen [25] uses JADE combined with EMD (Empirical Mode Decomposition) spectral correlation to extract the characteristic parameters reflecting the type of bearing fault. Then, the combination of the processed data and the strong generalization can realize the recognition of different fault types for the support vector machine (SVM) and finally achieved good results. He [26] extracted the time domain characteristics of the bearing-vibration signal, the frequency domain characteristics, and the energy characteristics of the wavelet packet to form a multidomain feature set. Then, used KJADE to effectively extract the low-dimensional nonlinear features contained in the multidomain feature set and finally combine the SVM for fault diagnosis and achieve good results. Shao et al. [27] used the original vibration signal to train the automatic encoder, which significantly improved the diagnostic accuracy compared with traditional fault identification methods such as BPNN, SVM, and Boosting. Yan and Jia [28] proposed a support vector machine (SVM) fault classification algorithm based on multisource features. This method introduces the Laplacian fractional algorithm to filter redundant information, select sensitive features, and finally use PSO-SVM to achieve identification of multiple fault conditions of the bearing. Saimurugan et al. [29] used the decision tree to select the optimal statistical characteristics of the vibration signal and then used the support vector machine to realize fault classification. In the fault type discrimination, a better diagnosis classification effect was obtained.

Past research has made significant progress, but there are some limitations. The fault diagnosis is essentially a pattern recognition problem, and deep learning can be used for pattern recognition by extracting the features layer by layer, which can be largely rid of relying on various advanced signal processing techniques and heavy artificial feature extraction. Therefore, the fault diagnosis method improved by the conventional method has achieved a large room for improvement and the effect is also good. Therefore, the study of reasonable methods and optimization algorithms is particularly important for the realization of intelligent and efficient diagnosis of rolling bearings.

1.2. Motivation and Contributions. The main purpose of this study is to extract the nonlinear characteristic parameters of normal bearing, inner ring fault, outer ring fault, rolling element fault, and three fault levels (a total of ten bearing states) under various fault modes. Then, the KJADE algorithm is used to feature the extracted original high-dimensional feature matrix to extract the effective features that are more sensitive to the bearing state. Finally, a fault diagnosis method for rolling bearing based on the improved whale algorithm is proposed to effectively identify bearing faults. The diagnostic results of LSSVM, GA-LSSVM, PSO-LSSVM, WOA-LSSVM, and VNWOA-LSSVM are compared, respectively, to verify the effect of the improved whale algorithm on fault diagnosis accuracy and diagnostic recognition rate of rolling bearings. Three contributions have been made.

1.2.1. Experimental Study on Rolling Bearing Fault. This study used Case Western Reserve University's bearing test data as the verification object. The experimental data of the model SKF6205 deep groove ball bearing were used as the simulation data. The experimental data are processed into different points of failure at different positions of the bearing by electric spark. The vibration signal of the bearing is recorded by the acceleration sensor at a sampling frequency of 48 kHz. The torque sensor is mainly used to monitor the speed of the bearing at a speed of 1750 rpm. The load motor

provides a load to the bearing. The bearing failure damage diameters are 0.007 inches (0.1778 mm), 0.014 inches (0.3556 mm), and 0.021 inches (0.5334 mm), respectively. By taking the bearing test data as the original input signal, the bearing normal, inner ring fault, outer ring fault, and rolling element fault are extracted. Finally, the fault diagnosis research and analysis of the nonlinear characteristic parameters of three fault levels (a total of ten bearing states) under different fault modes are carried out.

1.2.2. Modeling. Vibration signal feature extraction is an important step in the research of rolling bearing fault diagnosis. It is essential to extract the most useful features from the vibration signal. The extracted feature is a high-dimensional feature matrix, but there will still be some useless signals in it. Therefore, how to eliminate the redundancy between features and reduce the complexity of decision making is very important. In terms of fault diagnosis, due to some problems such as the convergence and accuracy of the optimization algorithm, the final fault diagnosis accuracy and fault recognition rate are not very high. This paper attempts to extract features from the time domain, frequency domain, and time-frequency domain of the vibration signal. Then, the algorithm is used to reduce the dimension of the feature. Finally, the Von Neumann topology is used to improve the WOA and optimize the LSSVM regularization parameters and the parameters of the nuclear parameters to diagnose the rolling bearing fault.

1.2.3. Model-Based Fault Diagnosis. A method based on KJADE and VNWOA-LSSVM is proposed to diagnose the rolling bearing-fault signal. Firstly, the multidomain feature set of nonlinear feature parameters is obtained by extracting time-frequency domain features from time domain, frequency domain, and VMD envelope spectrum. The feature fusion of the high-dimensional feature matrix in the feature set is then performed by KJADE. Finally, the Von Neumann topology is used to improve the WOA and optimize the LSSVM regularization parameters and kernel parameters, which effectively improves the accuracy and convergence speed of the LSSVM. At the same time, it maintains a good overall performance, which improves the ability to diagnose rolling bearing faults.

1.3. Organization of This Paper. The remainder of this paper is organized as follows. In Section 2, the theoretical methods of the time-frequency domain feature extraction, feature fusion, and two types of models are mainly described. In Section 3, the description of the proposed fault diagnosis research method based on the VNWOA-optimized LSSVM algorithm is introduced. In Section 4, several experimental cases are used to verify the effectiveness of the fault diagnosis method used in this paper compared with other traditional algorithms and the results and analysis of the method are given. Finally, the conclusion is presented in Section 5.

Figure 1 shows the framework of the specific idea of this article.

2. Time-Frequency Domain Feature Extraction and Feature Fusion

2.1. VMD Feature Extraction. The original signal is decomposed by VMD (Variational Mode Decomposition) to obtain corresponding components, wherein the ith component is represented as $c_i(t)$. Then, each component is separately Hilbert transformed to obtain the transformed amplitude function $a_i(t)$. Finally, the Fourier transform is performed on each amplitude function signal to obtain the envelope spectrum [30, 31].

The Hilbert transform is performed on the ith component obtained by the decomposition to obtain

$$H\left[c_i(t)\right] = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{c_i(\tau)}{t-\tau}\mathrm{d}\tau. \tag{1}$$

Construct the parsing signal:

$$z_i(t) = c_i(t) + jH\left[c_i(t)\right] = a_i(t)\exp\left(j\phi_i(t)\right). \tag{2}$$

Corresponding amplitude function

$$\begin{aligned} a_i(t) &= \sqrt{c_i^2(t) + H^2\left[c_i(t)\right]}, \\ x(t) &= \mathrm{Re}\sum_{i=1}^{n} a_i(t)\exp\left[j\phi_i(t)\right]. \end{aligned} \tag{3}$$

The resulting Hilbert spectral expression is

$$H(\omega, t) = \mathrm{Re}\sum_{i=1}^{n} a_i(t)\exp\left[j\int\omega_i(t)\mathrm{d}t\right]. \tag{4}$$

From the entire process of HHT analysis, the frequency and amplitude of each IMF (Intrinsic Mode Function) are a function of time, so $H(\omega, t)$ describes the distribution of the amplitude of the signal over time and frequency. If we consider $|x(t)|^2$ as the energy density of the signal, $H^2(\omega, t)$ also has the physical meaning of energy density after the Hilbert transform. $H^2(\omega, t)$ is called the Hilbert energy spectrum, which is the time-frequency representation of the signal energy.

2.2. KJADE Dimensionality Reduction Method. KJADE is a new feature fusion method based on the JADE (Joint Approximative Diagonalization of Eigenmatrix) algorithm, and it has greatly improved the handling of nonlinear problems and the improvement of the robustness of nonlinear results. The core idea of this algorithm is to perform high-dimensional mapping of the observed data $X \in R^{n\times m}$ (and $X = \{x_1, x_2, \ldots, x_m\}$) in the sample space through the nonlinear function $\varnothing$ and obtain the corresponding high-dimensional feature space $F = \{\phi(x_1), \phi(x_2), \ldots, \phi(x_m)\}$. Then, the JADE algorithm is applied to the high-dimensional feature space F, which can finally transform the nonlinearity into a linearly separable problem [32–34]. The mapping process is shown in Figure 2.

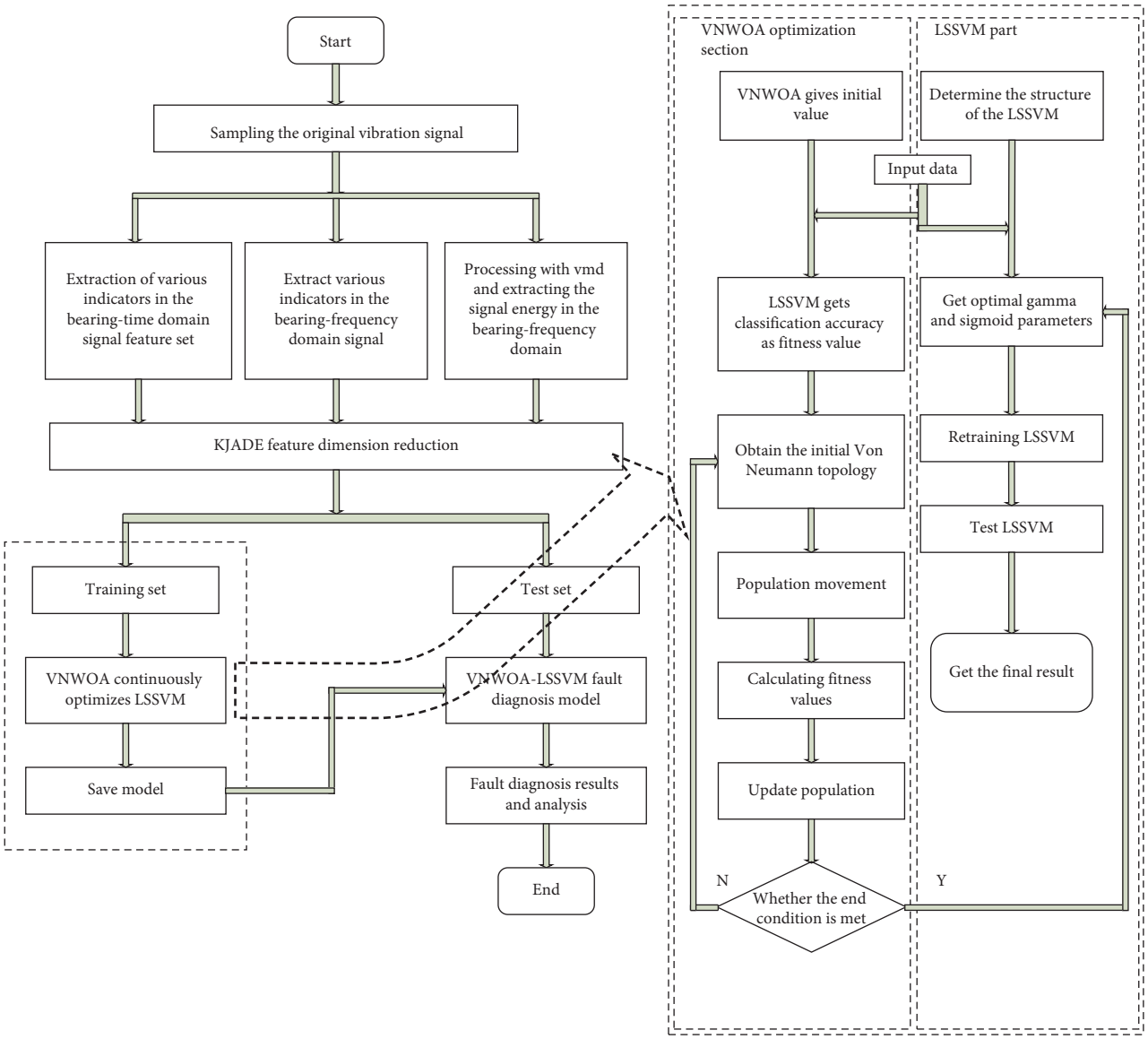

Figure 1: Schematic framework diagram.

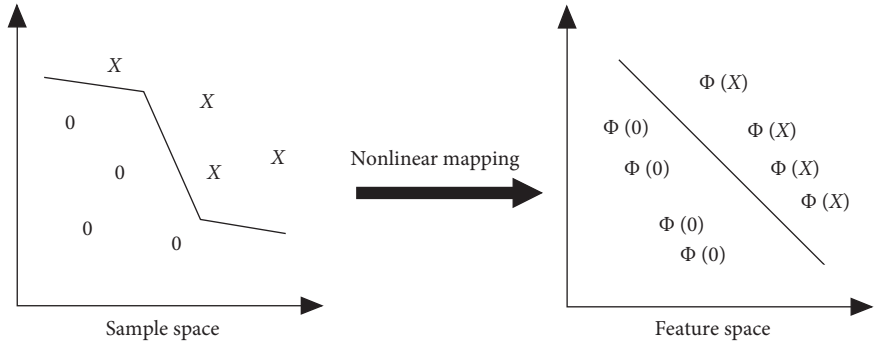

Figure 2: Nonlinear mapping.

The specific steps of the KJADE algorithm are summarized as follows:

(1) The sample point data in the observed data $X \in R^{n\times m}$ are mapped by high dimension or infinite dimension through the nonlinear mapping function $\varnothing$ and the feature space is $F = \{\phi(x_1), \phi(x_2), \ldots, \phi(x_m)\}$.

(2) Calculating the covariance matrix in the feature space F

$$R_F = \frac{1}{N}\varnothing(x_i)\varnothing(x_i)^T = \frac{1}{N}FF^T. \tag{5}$$

Introducing the concept of kernel function, convert the complex and time-consuming inner product calculation into kernel function and get a kernel matrix [33]:

$$K_{ij} = \langle \phi(x_i) \cdot \phi(x_j) \rangle = k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right). \tag{6}$$

The form is $N \times N$, where σ represents the width parameter of the function and K_{ij} is required to satisfy the Mercer condition, i.e., $K = FF^T$.

(3) For the abovementioned kernel matrix K whitening process, $z(t) = Wx(t)$, respectively, obtain the whitening matrix W and the whitened data $z(t)$.

(4) Selecting the matrix group $M = [M_1, M_2, \ldots, M_P]$ and finding the fourth-order cumulant matrix corresponding to each M_i, respectively:

$$Q_Z(M_i) = \sum_{i=1}^{P} K \cdot M_i, \quad i = 1, 2, \ldots, P. \tag{7}$$

(5) Calculating the rotation matrix

$$U = \arg_{\min} \sum_i \operatorname{off}\left(U^{\#} Q_Z(M_i) U\right), \tag{8}$$

where $\arg_{\min}$ is the argument of the complex number, off is the square of the nondiagonal element, and $U^{\#}$ is the pseudoinverse of U so that the cumulant matrix obtained in the previous step can be a matrix of diagonalized form.

(6) Finally, the mixed matrix B is calculated to obtain the optimal approximation matrix of the source signal:

$$y = B^F \phi(x) = U^T W \phi(x). \tag{9}$$

2.3. The Two Classes of Model. When the bearing fails, the extracted characteristic distribution of each bearing has a good class separability. Therefore, the difference between the vibration signals of each bearing and each part can be calculated by constructing two classes of model, which can be used as the basis for evaluating the performance degradation of the bearing. In the classification measurement of the sample, the distance between between-class and within-class has been successfully applied to the class separability metric [35] and the extraction of the bearing performance degradation curve [36].

The two classes of model are shown in Figure 3. It is assumed that the feature set extracted by the bearing in the healthy state is X_0 and the feature set extracted by the bearing at time t is Y_t. The two classes of model is $Z_t = [X_0, Y_t]$, where $X_0, Y_t = (x_1, x_2, \ldots, x_i, \ldots, x_n)$, n is the number of samples, and $x_i \in R^D$ (D is the extracted feature dimension).

Then, the between-class scatter matrix is

$$S_b = \sum_{i=1}^{C} P_i \|m_i - m\|^2. \tag{10}$$

The within-class scatter matrix is

$$S_w = \sum_{i=1}^{C} P_i \frac{1}{n_i} \sum_{k=1}^{n_i} \|x_k^i - m_i\|^2, \tag{11}$$

where

$$\begin{cases} P_i = \dfrac{n_i}{\sum_{j=1}^{C} n_j}, \\ m_i = \dfrac{1}{n_i \sum_{k=1}^{n_i} x_k^j}, \quad 1 \le i,\ j \le C, \\ m = \sum_{i=1}^{C} P_i m_i. \end{cases} \tag{12}$$

C is the number of categories, and $C = 2$ in this model; m_i is the feature mean in category i; and m is the feature mean in the entire sample. The between-class scatter matrix S_b represents the degree of convergence between different classes, and the within-class scatter matrix S_w represents the degree of aggregation of the same class. In order to comprehensively describe the feature part, the evaluation factor SS composed of between-class and within-class scatters is employed to depict the clustering performance of the proposed new features quantitatively. Then, the SS is defined as the following equation:

$$\mathrm{SS} = \operatorname{trace}\left(\frac{S_b}{S_w}\right). \tag{13}$$

3. Research on Fault Classification Based on the VNWOA-Optimized LSSVM Algorithm

3.1. LSSVM Algorithm. SVM is a relatively new machine learning method based on statistical theory as basic research. In the case, where the sample is very limited, it can be used to find the optimal solution. LSSVM introduces the least squares based on the traditional SVM. Based on Vapnik's support vector machine, the optimization goal is defined by a quadratic loss function with a square term. Finally, the

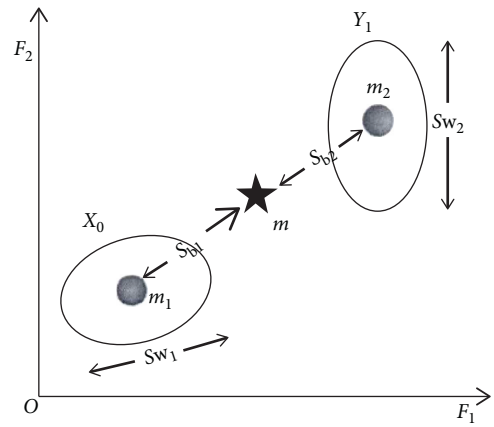

Figure 3: The diagrammatic sketch of the two classes of model.

inequality constraint in the original quadratic optimization problem becomes the equality constraint problem for solving linear equations, which can reduce the complexity of the overall calculation and increase the operation speed [37–39].

The optimization objective function using the LSSVM algorithm is

$$f(x) = \operatorname{sgn}\{\omega \cdot \phi(x) + b\}. \tag{14}$$

The final optimization problem becomes

$$\begin{cases} \min J(\omega, \xi) = \dfrac{1}{2}\|\omega\|^2 + \gamma \sum_{i=1}^{N} \xi_i^2, \\ \text{s.t. } y_i\left[\omega^T \phi(x_i) + b\right] = 1 - \xi_i, \\ i = 1, 2, \ldots, N. \end{cases} \tag{15}$$

In order to solve the problem of optimization well, the Lagrange multiplier α_i is introduced and the function of Lagrange is constructed. The formula is as follows:

$$L(\omega, b, \xi, \alpha_i) = J(\omega, \xi_i) - \sum_{i=1}^{N} \alpha_i \left\{ y_i \left[\omega^T \phi(x_i) + b\right] - 1 + \xi_i \right\}. \tag{16}$$

Then, using the constraints on the KKT condition. The relevant parameters of the Lagrange function are separately subjected to a partial derivative operation at the extreme points sought, and the result is zero. The resulting linear matrix expression is as follows:

$$\begin{pmatrix} 0 & y^T \\ y & ZZ^T + C^{-1}I \end{pmatrix} \begin{pmatrix} b \\ \alpha \end{pmatrix} = \begin{pmatrix} o \\ 1_N \end{pmatrix}. \tag{17}$$

There are some relationships in the above expression: $y = [y_1, y_2, \ldots, y_N]^T$; $1_N = [1, 1, \ldots, 1]^T$; $I = R \times R$; $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_N]^T$; and $Z = [\phi(x_1), \phi(x_2), \ldots, \phi(x_N)]^T$.

In the LSSVM algorithm, the resulting optimal classification function expression is

$$f(x) = \operatorname{sgn}\left[\sum_{i=1}^{N} \alpha_i y_i \mathrm{K}(x, x_i) + b\right]. \tag{18}$$

In the above formula, $K(x, x_i)$ is defined as a kernel function satisfying the Mercer condition and the expression is the same as equation (6).

3.2. Whale Optimization Algorithm

3.2.1. Principle of Whale Optimization Algorithm. Mirjalili and Lewis studied the predatory behavior of humpback whales [40] and designed a new heuristic search optimization algorithm with simple operation, less adjustment parameters, and strong local optimal ability. The algorithm is called the WOA (Whale Optimization Algorithm), which has some natural inspiration, mainly to imitate the predation behavior of humpback whales, including modeling the three behaviors of the humpback whales surrounding prey, hunting prey, and hunting prey, as shown in Figure 4.

(1) In the model surrounding the prey, assuming that the target prey position or the position closest to the target prey is equated with the current best candidate solution, after the defined best search agent is subsequently defined, the other agents will update their position and gradually nearest to the best search agent. Its location update expressions are as follows:

$$D = \left|CX^*(t) - X(t)\right|, \tag{19}$$

$$X(t+1) = X^*(t) - AD, \tag{20}$$

$$A = 2ar_1 - a, \tag{21}$$

$$C = 2 \cdot r_2, \tag{22}$$

$$a = 2\left(1 - \frac{t}{T_{\max}}\right). \tag{23}$$

In the above equations (19)~(23), t represents the number of current iterations, $X^*(t)$ is the vector of the whale in the best position at the current stage, $X(t)$ represents the vector of the current whale position, a is represented as a constant and the range is changed from 2 to 0 linear decreasing law, r_1 and r_2 belongs to any number in [0, 1], and $T_{\max}$ represents the maximum number of iterations.

(2) In the model of hunting prey, the model consists of two parts.

The first part is narrowing the range of prey, that is, the value of a in equation (21) is decreasing. At the same time, the value of A varies by $[-a, a]$ and the value of a will also change with changes. If $A \in [-1, 1]$, it means that the target position of the whale hunting prey is at any coordinate between the

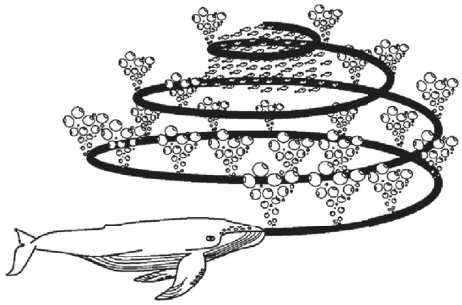

Figure 4: Bubble-net prey behavior of whales.

current position and the prey, thus showing that the humpback whale has a good local search ability.

The second part is position update with spiral features:

$$X(t+1) = D' e^{bl} \cos(2\pi l) + X^*(t), \tag{24}$$

$$D' = \left| X_t^* - X_t \right|. \tag{25}$$

In equations (24) and (25), D' represents the distance between the whale and the prey at the current stage; b represents the logarithmic spiral shape constant; l represents a random number; and the value ranges from $[-1, 1]$.

Humpback whales contain both the first and second parts described above when hunting prey, and the system defines a 50% probability as a threshold to apply to the position of the humpback whale. The specific expression is as follows:

$$X(t+1) = \begin{cases} X(t) - A \cdot D, & p < 0.5, \\ D' e^{bl} \cos(2\pi l) + X(j), & p \geq 0.5. \end{cases} \tag{26}$$

In the above equation, p is any number belonging to [0, 1].

(3) In the model of searching for prey, for the individual whales, they are randomly supplemented with prey within a certain range. When $A \in (-\infty, -1) \cup (1, +\infty)$, the distance D will also change randomly and the whale will change accordingly in the search for the target of feeding. Therefore, the method can improve its original feeding ability and has a global search ability. The expressions are as follows:

$$\begin{aligned} D &= \left| C \cdot X_{\text{rand}} - X(t) \right|, \\ X(t+1) &= X_{\text{rand}} - A \cdot D. \end{aligned} \tag{27}$$

In the above formulas, X_{rand} is expressed as the position of any individual in the current whale population.

3.2.2. Improvement of Whale Algorithm. In view of the shortcomings of the WOA in optimizing some complex problems, the accuracy is not high and the convergence speed is slow. Therefore, the WOA is improved by using the Von Neumann topology. As shown in Figure 5, the Von Neumann topology [41], each humpback whale is surrounded by up, down, left, and right directions to form a grid, while the humpback whale is in contact with the surrounding whales in the center. The head whales have the best solution to affect other whales around them, allowing the fish to communicate frequently and maintain the multidirectionality and diversity of the group. The premature occurrence has been improved in terms of maintaining globality, convergence speed, and adjustment accuracy.

Studying the abovementioned whale position update formula, it shows that the change has a great relationship with the global optimal solution and will change with the update of the global optimal solution. The midpoints of the local and global optimal positions are selected and adopted so that the whales are affected by the global and local optimal solutions while performing position updating. It can exchange sufficient information with other local whales to enhance the local search ability of the algorithm. As the number of iterations increases, the local optimum and the global optimal will appear approximately coincident and the position of the whale will be updated to return to the original formula [42]. The improved expression is as follows:

$$X_i^* = \frac{P_{l\text{best}_i} + G}{2}. \tag{28}$$

In the above formula, $P_{l\text{best}_i}$ indicates the optimal position of the ith search agent of the humpback whale in the Von Neumann topology field $L(i)$, which can also be understood as being in this field $L(i)$. Each whale search agent's own fitness function obtains the historical optimal position of the particle corresponding to the minimum value; G represents the optimal position of the whale in the global update.

3.2.3. Analysis of the Advantages of the Improved Whale Optimization Algorithm

(1) The VNWOA uses the Von Neumann topology. Within each neighborhood, the optimal solution found by the humpback whale at each central location affects only the other four whales in the same neighborhood. The full exchange of whales in the neighborhood can maintain the multidirectionality and diversity of the population.

(2) While multiple neighborhoods maintain the diversity of fish stocks, avoiding the fact that a whale finds a local optimal solution prematurely, the whole population falls into local optimum, and it is improved in terms of maintaining globality and convergence speed.

(3) By improving the whale position update formula in the algorithm, the midpoint of the local optimal

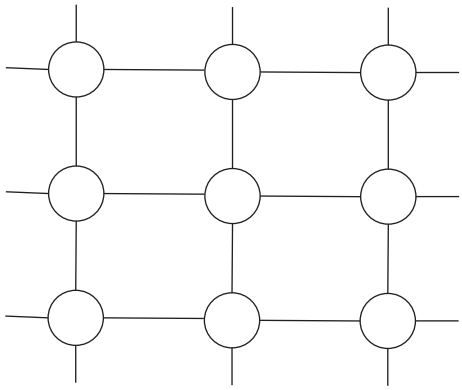

Figure 5: Von Neumann topology.

position and the global optimal position is selected. In this way, the whale can not only be guided by the global optimum when the location is updated, but also enhance the whale communication ability in the local area. The update of its location is also affected by a part of the local optimum, thus enhancing the local search ability of the algorithm.

(4) As the iteration progresses, the local optimal value gradually coincides with the global optimal value and the whale's position update eventually returns to the original formula. The ability of the algorithm to achieve convergence and precision is enhanced.

3.3. VNWOA-Optimized LSSVM Algorithm Steps

Step 1 (data preprocessing): The time domain, frequency domain, and time-frequency domain characteristics are obtained for the original vibration signals and normalized. The feature reduction is performed by KJADE, and the training set and test set are divided by 7 : 3.

Step 2: Initializing the whale position and setting the population to $N = 10$, and the maximum number of iterations is $T_{\max} = 100$.

Step 3: Determining the network topology of the VNWOA and the range of values for initializing the LSSVM, that is, the range of values of σ^2 and gamma [0 1000].

Step 4: Calculating the corresponding fitness of each whale, and determining the initial optimal individual and initial optimal fitness according to the order of fitness.

Step 5: Using the Von Neumann topology to perform neighborhood search, exchanging information in the neighborhood, finding the best whale in the neighborhood, and then following the formulas (19), (20), (26), and (28) to perform a location update.

Step 6: The whale swims in a spiral shape to the prey while shrinking the encirclement. The position of the other whales is updated according to the selected whale position, forcing the whale to deviate from the prey thus obtaining the best training accuracy in the LSSVM.

Step 7: Repeating steps 5 and 6. Until the maximum number of iterations $T_{\max} = 100$ is reached, the whale position with the best fitness is trained as a parameter of the LSSVM, and then the test set is diagnosed and classified.

4. Experimental Results and Analysis

4.1. Basic Steps for Diagnosis

(1) Vibration signal preprocessing: by selecting the normal state of the bearing and three types of faults (corresponding to three fault levels for each of the three fault types), a total of ten bearing states are used as verification objects. Then, the time domain and frequency domain of ten bearing states and the method of using VMD envelope spectrum are used to extract the time-frequency domain features of the rolling bearing so as to obtain the multidomain feature set of nonlinear characteristic parameters.

(2) Feature fusion: KJADE is used for feature fusion in the high-dimensional feature matrix of the feature set, and the evaluation factor SS composed of between-class and within-class scatters is employed to depict the clustering performance of the proposed new features quantitatively. It is then compared to the PCA, LDA, KPCA, and JADE clustering effects.

(3) Selection of fault signal for training and testing: a total of 1000 fault signal samples are randomly selected from ten types of bearing states. Among them, 700 samples were used as training samples and 300 samples were used as test samples and the data length of each sample was 15.

(4) Fault diagnosis: the method used for fault diagnosis mainly uses the Von Neumann topology to improve the WOA and optimize the LSSVM regularization parameters and the parameters of the nuclear parameters. In the experiment, the maximum number of iterations is set to 100, and then the 700 sets of samples in step 3 are input to the whale position with the best fitness as the parameters of the LSSVM for training. Finally, the remaining 300 sets of test data are diagnosed.

4.2. Introduction to Experimental Equipment. The rolling bearing fault data used in this experiment were obtained from the Case Western Reserve University Bearing Data Center [43], which provides a large number of rolling bearing fault data. The corresponding experimental device is shown in Figure 6. The experimental support consists of a 2 hp motor (left), a torque sensor/decoder (middle), a dynamometer (right), and an electrical control unit (not shown). The treated bearing will be mounted on the drive end of the motor, while the inner ring of the bearing will follow the rolling axis but will remain stationary. The acceleration signal is used to record the vibration signal of the

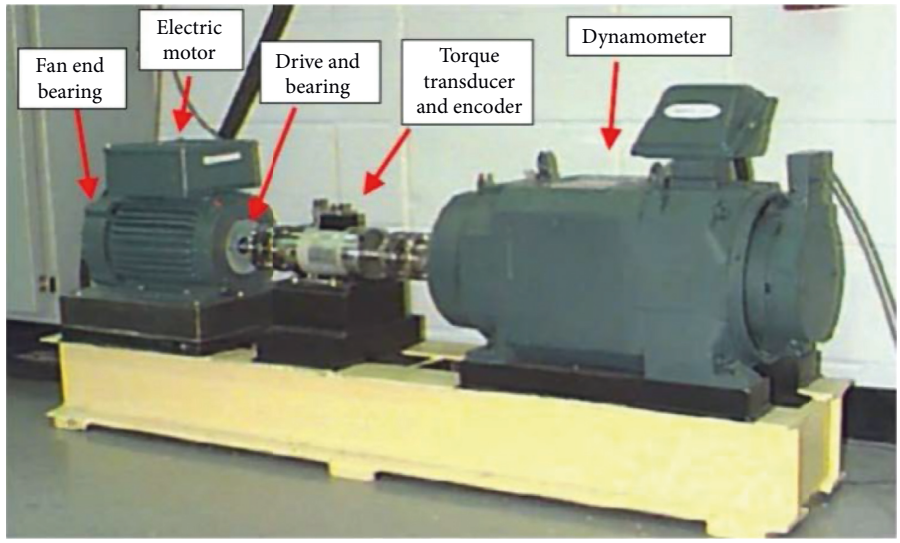

Figure 6: Bearing fault diagnosis test equipment.

bearing during the whole process. The sampling frequency is 48 kHz. The torque sensor is mainly used to monitor the rotation speed of the bearing. The rotation speed is 1750 rpm, and the load motor provides the load for the bearing. In the test, the model data of SKF6205 deep groove ball bearing is used as the simulation data to test the bearing support motor shaft and three types of test bearings are prepared.

This experiment uses an acceleration sensor adsorbed on a magnetic body to obtain the vibration data. At the drive and blast ends of the motor frame, the sensor is placed at 12 o'clock. The vibration signal was recorded with a 16-channel digital recording recorder and saved as a Matlab (*mat) format in the form of a file. The speed and power data are manually recorded by a torque sensor/decoder. By using the bearing test data as the original input signal, the nonlinear characteristic parameters of the bearing normal, inner ring fault, outer ring fault, rolling element fault, and three fault levels (a total of ten bearing states) in various fault modes are extracted. The EDM technology is used to test the inner ring, the outer ring, and the rolling element of the bearing, respectively. Single point damage failures of 0.007 inches (0.1778 mm), 0.014 inches (0.3556 mm), and 0.021 inches (0.5334 mm) with a depth of 0.011 inches (0.2794 mm) were set [44, 45].

"Normal" indicates bearing under normal conditions; "IR" indicates inner ring failure; "OR" indicates outer ring failure; and "B" indicates rolling element failure. The ten bearing data obtained are numbered and indicated by I-X. The specific bearing data are shown in Table 1.

4.3. Feature Extraction. When the motor speed is 1797 rpm (the frequency conversion is also considered to be $f_i = 29.95$), the drive end fault bearing is selected as the analysis object. In the case of bearing failure damage diameters of 0.007 inches (0.1778 mm), 0.014 inches (0.3556 mm), and 0.021 inches (0.5334 mm), respectively, the bearing-vibration signals of the outer ring, the inner ring, and the rolling element position are selected as the object. In the case of bearing failure damage diameters of 0.007 inches (0.1778 mm), 0.014 inches (0.3556 mm), and 0.021 inches (0.5334 mm), the bearing-vibration signals of the three failure positions were selected as objects and compared with healthy bearings. The bearing signals of the various bearing states collected therein are shown in Figures 7(a)–7(j) correspond to the bearing health status, outer ring failure status, inner ring fault status, and rolling element fault status.

A time-domain waveform diagram of the bearing in a healthy state is shown in Figure 7(a), and it can be found that its vibration and noise are relatively small. The vibration waveform is disorderly and irregular, and the amplitude is small. Figures 7(b)~7(d) show time-domain waveforms for the case, where the bearing inner ring failure damage diameter is 0.007 inches (0.1778 mm), 0.014 inches (0.3556 mm), and 0.021 inches (0.5334 mm). Similarly, Figures 7(e)~7(j) are the three impaired waveforms corresponding to the rolling element and the outer ring, respectively. From the waveform in the figure, it can be concluded that when the bearing is partially damaged, an impact signal is generated, so a short-time low-frequency pulse occurs in the vibration signal of the faulty bearing. At the same time, this shock will arouse the high-frequency natural vibration and the amplitude will be modulated by this excitation.

Four characteristic indexes of the center frequency, frequency standard deviation, root mean square frequency, and frequency concentration of the bearing signal are extracted as frequency domain features, and the time-frequency domain characteristic signals are extracted by using the VMD envelope spectrum method to form a multidimensional feature set. Since the corresponding spectrum and the envelope spectrum have the same range of values on the abscissa, image blending is employed and normalized.

The two characteristics of the time-frequency spectrum and the envelope spectrum when the bearing is in normal

Table 1: Bearing data.

Bearing status	Degree of failure (inches)	Abbreviation	Category labeling
Normal	0	Normal	I
Inner ring failure	0.007	IR07	II
	0.014	IR14	III
	0.027	IR21	IV
Outer ring fault	0.007	OR07	V
	0.014	OR14	VI
	0.027	OR21	VII
Rolling element failure	0.007	B07	VIII
	0.014	B14	IX
	0.027	B21	X

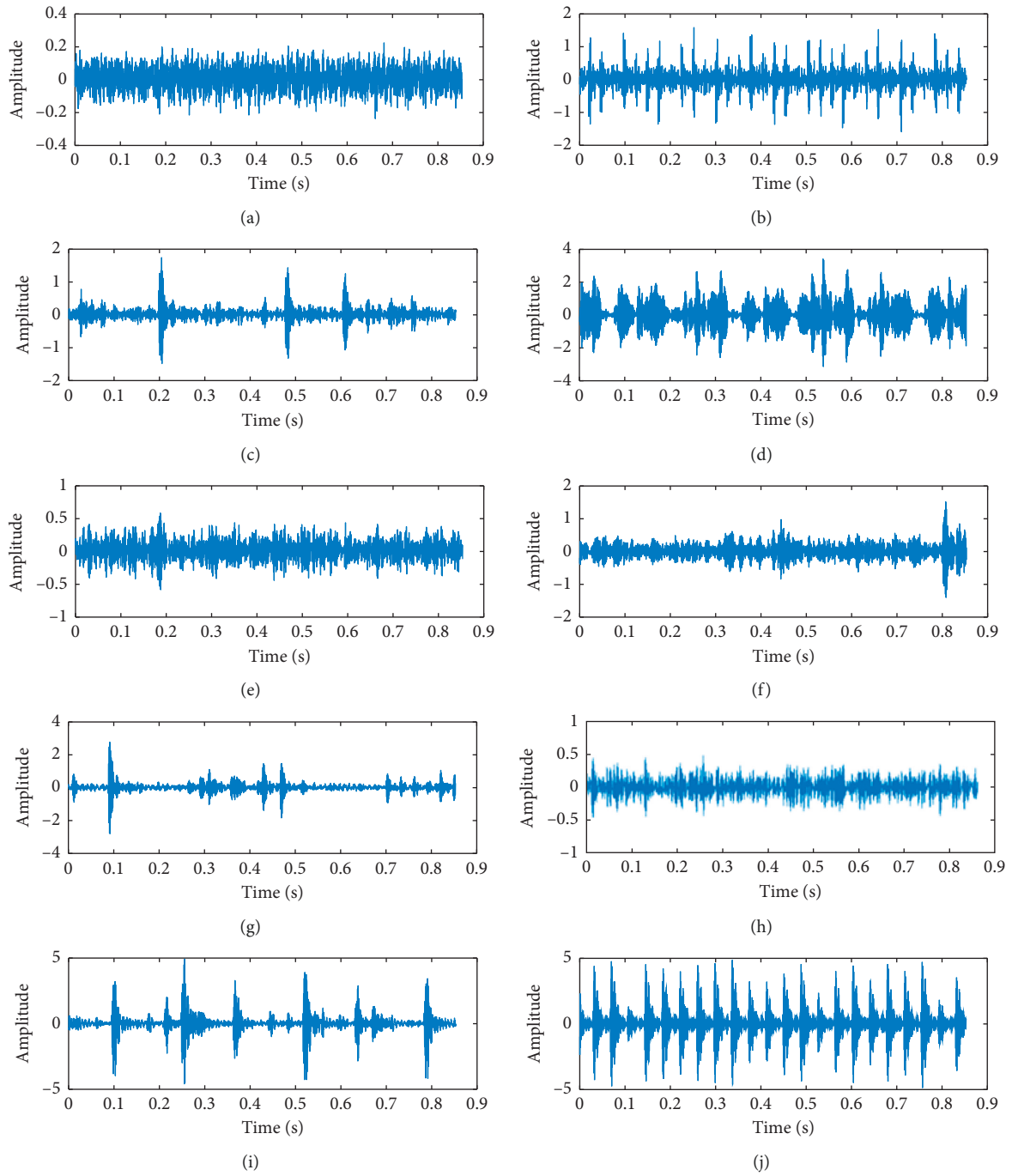

Figure 7: Time domain waveform of 10 kinds of bearings.

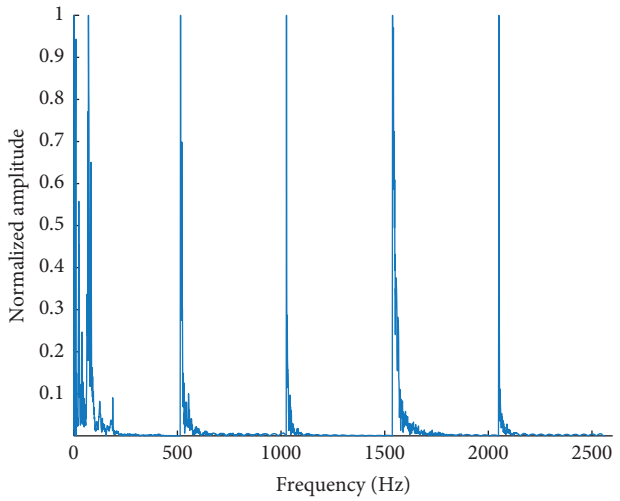

Figure 8: Fault-free normalized waveform.

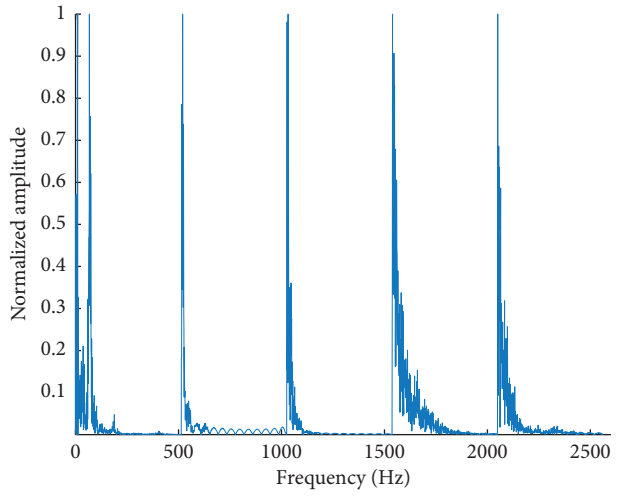

Figure 9: B07 normalized waveform.

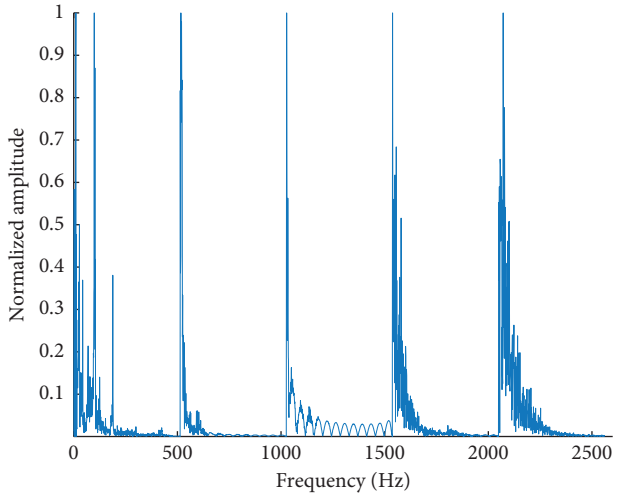

Figure 10: B14 normalized waveform.

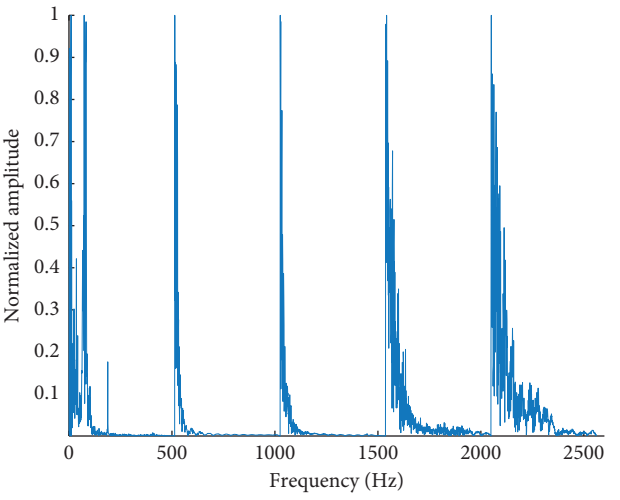

Figure 11: B21 normalized waveform.

operation is shown in Figure 8. It can be seen that the bearing has the most prominent frequency conversion, the spectrum is relatively simple, and the energy is concentrated in the low-frequency band. Figures 9–11 show the corresponding image of the rolling element. The signal is rather messy. This is because when the roller fails, the failed roller not only rotates with the inner ring, but also rotates itself, so the defect information is often submerged in the noise. In the signal, it is difficult to be as obvious as the outer ring and the inner ring, so we will see that there is a strong energy in both the low- and high-frequency bands. Figures 12–17 show the image of the inner and outer rings of the bearing. We will find that the energy is mainly concentrated in the middle frequency band because the local position of the damage will cause the natural vibration of each component. On the other hand, for sensors with a resonant frequency less than 20 kHz, the resonance information caused by the impulse signal is also reflected in this frequency band. In addition, the high-frequency natural vibration of the bearing is modulated by the shock caused by these damages, so the fault information can be separated from the chaotic modulation signal by the demodulation technique. By comparing these processes, the results show that when the bearing develops from a normal state to a fault, the main energy in the spectrum will gradually shift from the low-frequency band to the middle-frequency band. At the same time, for the same part of the bearing, as the degree of damage increases, the amplitude corresponding to the same frequency will also become larger.

4.4. Feature Fusion. From the original bearing vibration information, the number of samples is 1000 and the multidomain feature set is extracted from the time domain, the frequency domain, and the time-frequency domain, thereby obtaining a feature matrix of *2560. On this basis, the parameters of PCA, LDA, KPCA, and JADE are reasonably set to ensure the fairness of feature fusion. The PCA selects a k value of 3 and a guaranteed error of less than 0.05 (95% of the information is retained). The dimension reduction

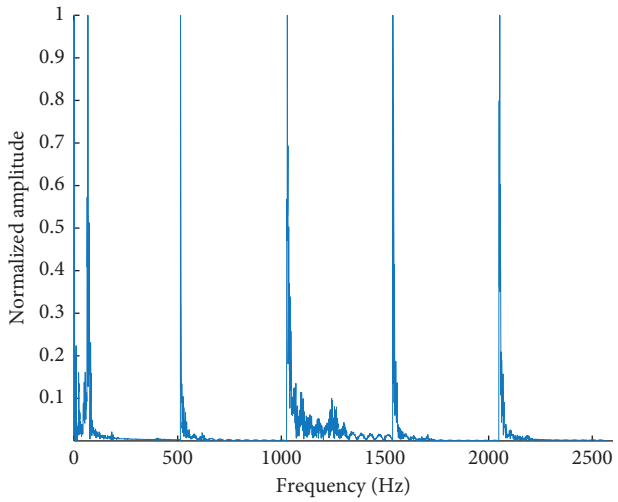

Figure 12: IR07 normalized waveform.

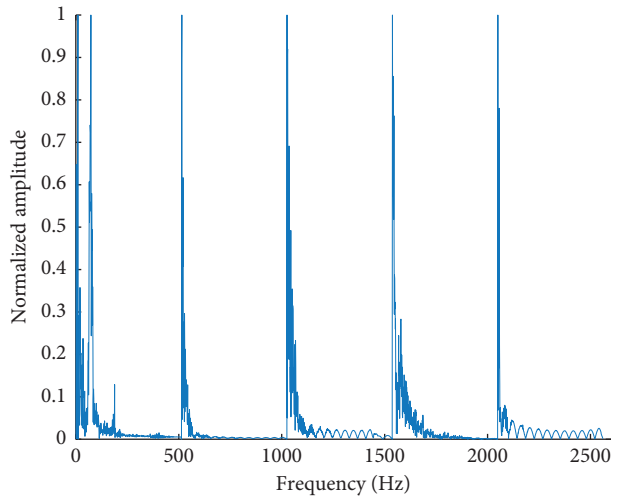

Figure 13: IR14 normalized waveform.

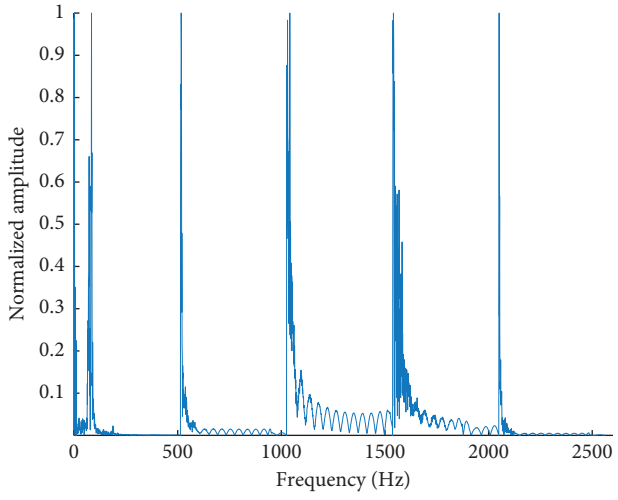

Figure 14: IR21 normalized waveform.

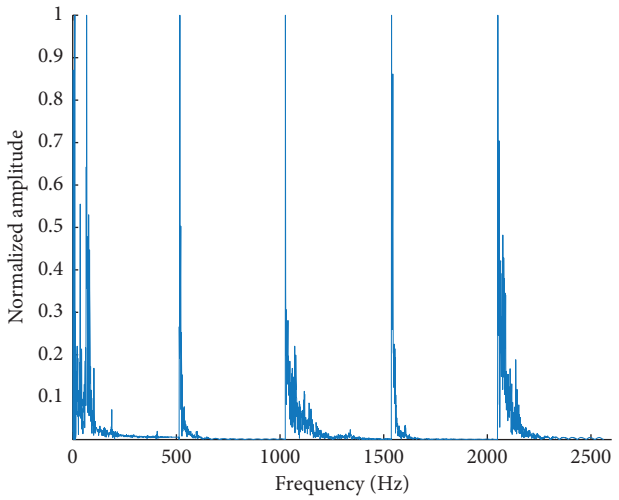

Figure 15: OR07 normalized waveform.

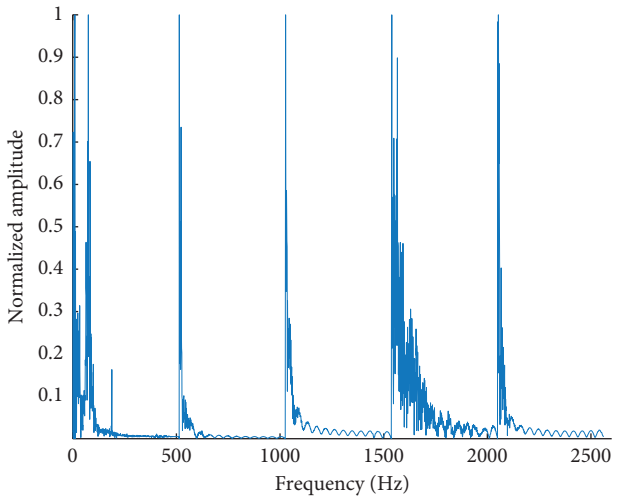

Figure 16: OR14 normalized waveform.

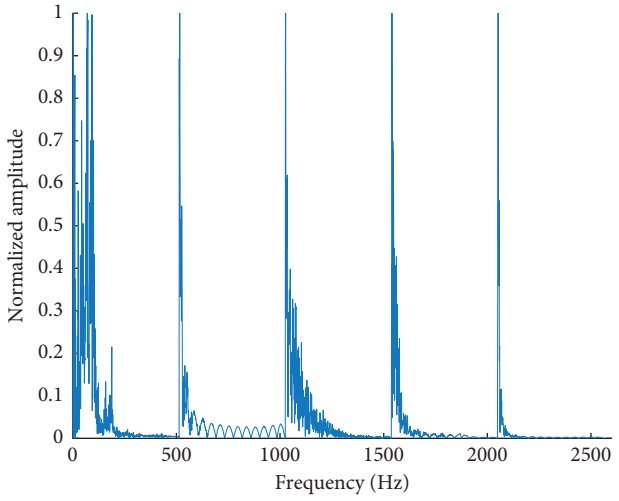

Figure 17: OR21 normalized waveform.

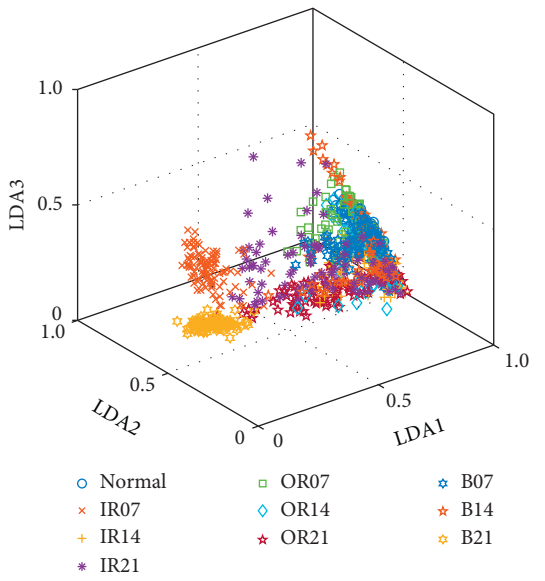

Figure 18: LDA feature distribution map.

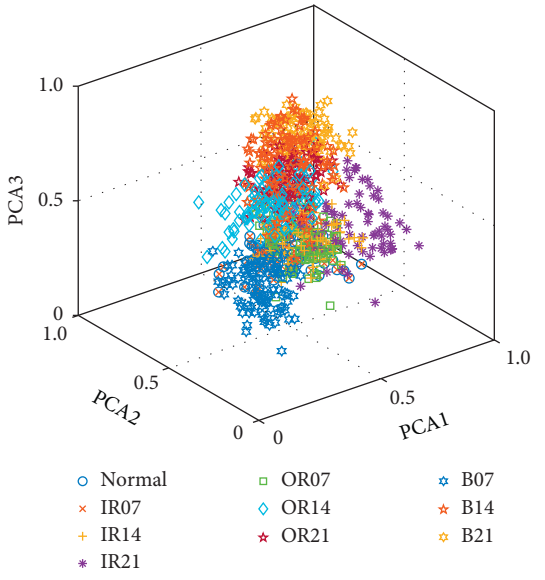

Figure 19: PCA feature distribution map.

selected by LDA is 3. The kernel function used by KPCA is a radial basis function with a parameter value of 4. The dimension of JADE is taken as 3 for feature fusion and sensitive feature extraction. Through the above parameters and dimension settings, on the one hand, it is ensured that enough sensitive feature information is retained and the feature set can be displayed more conveniently and facilitated by us. The distribution of the feature fusions of the 10 fault states using PCA, LDA, KPCA, JADE, and KJADE is shown in Figures 18–22.

It can be clearly seen from Figures 18–22 that the low-dimensional features extracted by the KJADE method compared with PCA, LDA, KPCA, and JADE have very good clustering effect in the feature space. The KJADE-based feature fusion method is more compact, and the

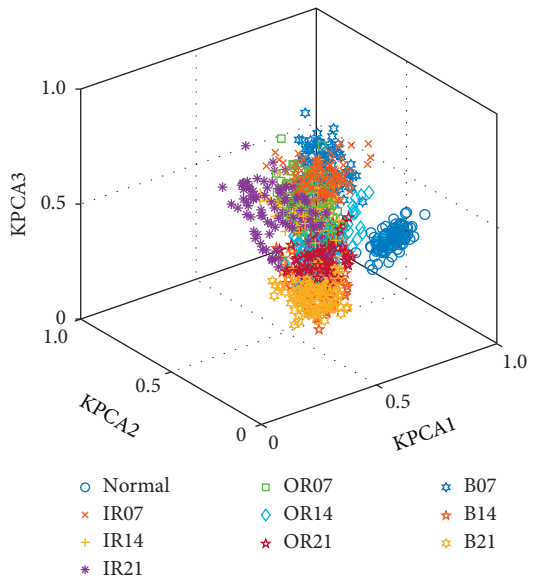

Figure 20: KPCA feature distribution map.

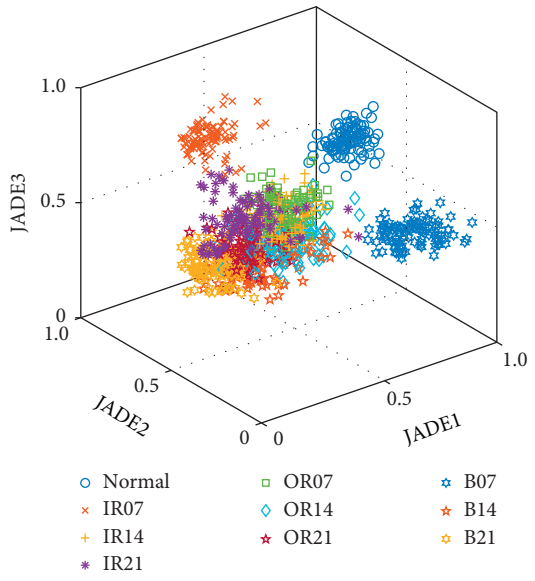

Figure 21: JADE feature distribution map.

characteristics of each bearing state are basically concentrated in a small area. Moreover, there are few overlapping parts of various sample features and these distribution characteristics will facilitate the identification of different fault types. This indicates that the extracted features obtained by KJADE are more suitable as classifiers than those obtained by conventional methods. In order to better quantitatively describe the clustering performance of the extracted new features, the evaluation factor SS composed of between-class and within-class scatters is employed to depict the clustering performance of the proposed new features quantitatively. It can be seen from Table 2 that the clustering evaluation SS of KJADE features is higher than the traditional method, which further shows that the KJADE algorithm is suitable for extracting more stable and effective low-dimensional feature nonlinear low-dimensional feature components embedded in high-dimensional data. It also shows that KJADE can eliminate the redundancy between features and reduce the dimension of features to obtain effective low-dimensional features. It can reduce the complexity of decision making under the premise of retaining the classifier classification performance in fault diagnosis degree.

4.5. Fault Diagnosis Results. From the total sample of bearing vibration data, 700 samples were randomly selected as training samples and 300 samples were used as test samples, and the data length of each sample was 15. The characteristics of the training set are identified by the

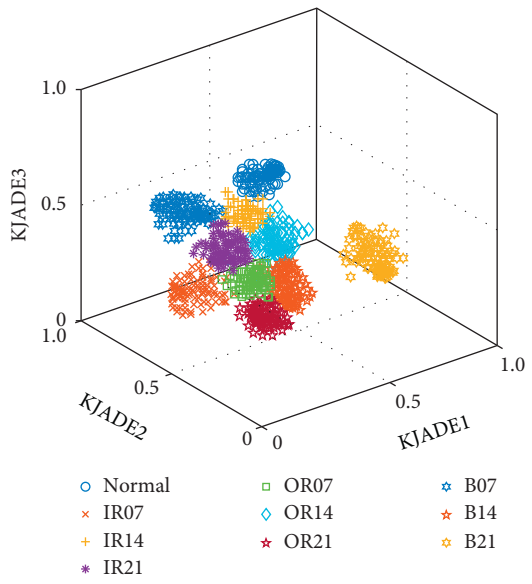

Figure 22: KJADE feature distribution map.

Table 2: SS of different methods.

Method	LDA	PCA	KPCA	JADE	KJADE
ss	16.7	9.5	27.6	48.2	202.7

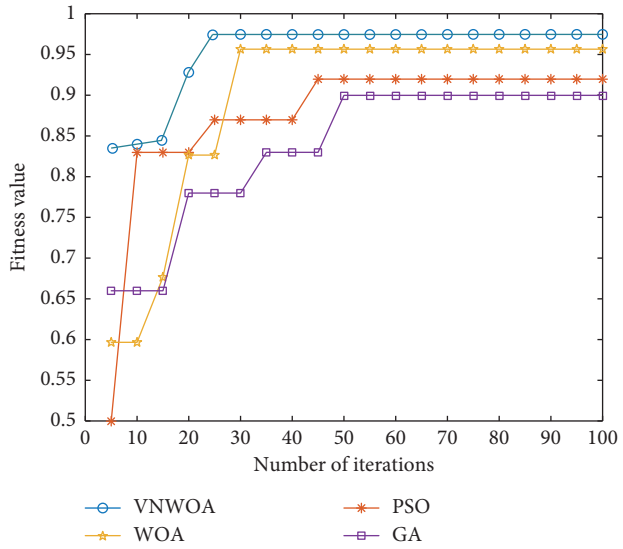

Figure 23: LSSVM optimal fitness value.

category and then input to the VNWOA-LSSVM training and compared with the LSSVM, GA-LSSVM, PSO-LSSVM, and WOA-LSSVM.

It can be seen from the fitness curve of Figure 23 that the fitness value of the VNWOA optimization algorithm proposed in this paper can reach 0.9784 and stabilize after 25 iterations. Its characteristics show that the convergence rate of the population is relatively slow at the beginning, and then the convergence speed is accelerated so that the algorithm quickly converges to the optimal fitness. Compared with several other algorithms, several other optimization algorithms will eventually have higher fitness values but slower convergence. The number of iterations will be 30 or more from the graph. This shows that the VNWOA algorithm is superior to other algorithms, indicating that this method can be used as an effective means of bearing fault pattern recognition.

In Figures 24–28, the ordinates 1–10 represent the category labeling symbol I-X in Table 1 , indicating the type of diagnosis in the case of the algorithm. Comparing Figures 24~28 with Table 3, it can be seen that the result of

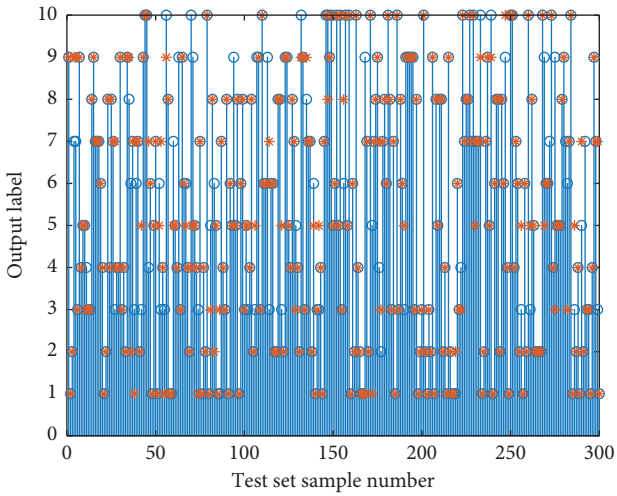

Figure 24: LSSVM troubleshooting results.

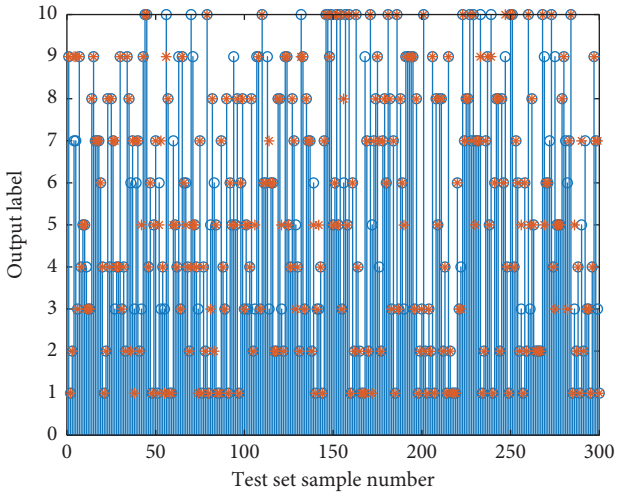

Figure 25: GA-LSSVM troubleshooting results.

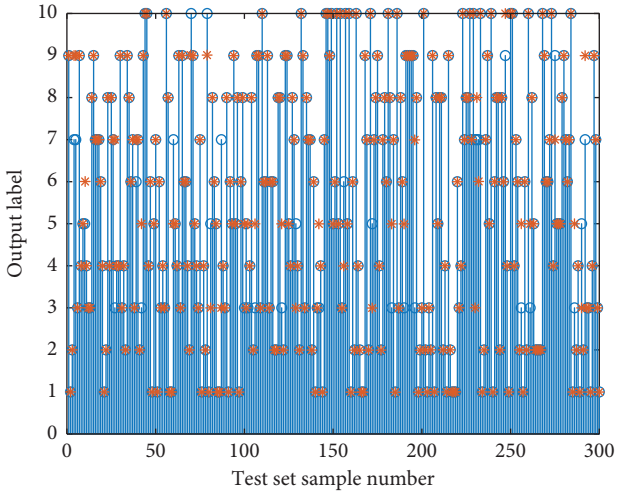

Figure 26: PSO-LSSVM troubleshooting results.

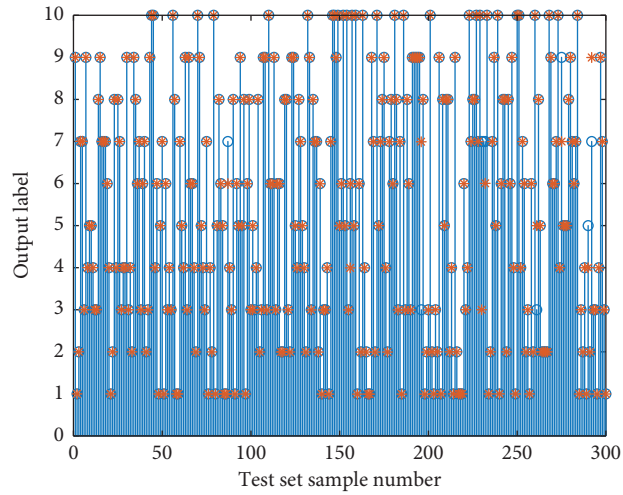

Figure 27: WOA-LSSVM troubleshooting results.

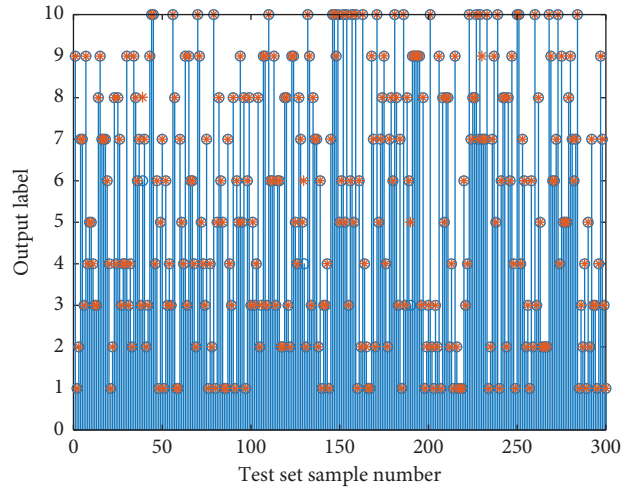

Figure 28: VNWOA-LSSVM troubleshooting results.

Table 3: Algorithm performance comparison.

Algorithm type	Fitness value	Training accuracy (%)	Diagnostic accuracy (%)	Training mean square error (%)
LSSVM	—	87.46	85.67	4.325×10^{-5}
GA-LSSVM	0.8996	89.38	88.00	3.268×10^{-6}
PSO-LSSVM	0.9187	90.96	90.67	3.059×10^{-6}
WOA-LSSVM	0.9583	94.32	94.00	3.156×10^{-6}
VNWOA-LSSVM	0.9784	98.79	98.67	2.792×10^{-6}

fault diagnosis of VNWOA-LSSVM can reach 98.67%, which is greatly improved compared to the result of fault diagnosis without optimized LSSVM of 85.67%. The

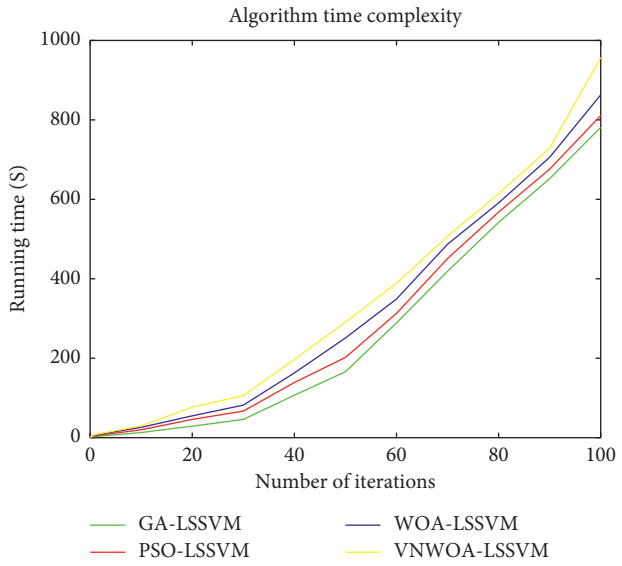

Figure 29: Algorithm time complexity comparison chart.

diagnostic results of GA, PSO, WOA, and VNWOA-optimized LSSVM algorithm can clearly show that the diagnostic accuracy obtained by VNWOA-optimized LSSVM is the highest, indicating that VNWOA-LSSVM can accurately diagnose faults.

Figure 29 shows the time complexity of GA-LSSVM, PSO-LSSVM, WOA-LSSVM, and VNWOA-LSSVM algorithms. It can be seen from the figure that the time complexity of the algorithm for fault diagnosis of LSSVM optimization by different algorithms is also different. The population set in this study is 10, the ordinate is the running time of the algorithm, and the abscissa is the number of iterations. As the number of iterations increases, the time complexity of the four optimization algorithms is increasing. Moreover, it can be clearly found in combination with Table 3 that VNWOA-LSSVM is much more accurate in terms of diagnosis results, but it is also much more complicated in terms of algorithm complexity. On the other hand, it can be seen from Table 3 that after optimizing the LSSVM with the VNWOA algorithm, the training accuracy, test accuracy, and training variance are obviously superior to the other four algorithms, and the mean square error is relatively small, indicating that the algorithm has good stability and generalization ability.

5. Conclusion

Through the normal, inner ring fault, outer ring fault, rolling element fault and three fault levels in various fault modes (a total of ten bearing states), the time domain, the frequency domain, and the method using the VMD envelope spectrum, the time-frequency domain features are extracted to obtain the multidomain feature set of nonlinear feature parameters. Then, the KJADE is used to feature the high-dimensional feature matrix in the feature set. The extracted low-dimensional features have very good clustering in the feature space. The effect and the effective characteristics that are more sensitive to the bearing state can be obtained. At the same time, in the fault diagnosis, the Von Neumann topology is used to improve the WOA and optimize the LSSVM regularization parameters and the parameters of the nuclear parameters. The accuracy and convergence speed of LSSVM are improved and a good overall globality is maintained, thus improving the ability of fault diagnosis of rolling bearings. It shows that the proposed method for fault diagnosis of rolling bearings has a good diagnostic effect and also provides a diagnosis method for rolling bearing fault diagnosis.

Abbreviations

CNN:	Convolutional neural network
DBN:	Deep belief network
AR:	Autoregression
BPNN:	Back propagation neural network
PSO-SVM:	Particle swarm optimization-support vector machine
HHT:	Hilbert–Huang transform
IMF:	Intrinsic mode function
EMD:	Empirical mode decomposition
EDM:	Electrical discharge machining
VMD:	Variational mode decomposition
SVM:	Support vector machine
LSSVM:	Least squares support vector machine
PCA:	Principal component analysis
LDA:	Linear discriminant analysis
KPCA:	Kernel principal component analysis
JADE:	Joint approximative diagonalization of eigenmatrix
KJADE:	Kernel function joint approximate diagonalization of eigenmatrices
WOA:	Whale optimization algorithm
GA-LSSVM:	Genetic algorithm-least squares support vector machine
PSO-LSSVM:	Particle swarm optimization-least squares support vector machine
WOA-LSSVM:	Whale optimization algorithm-least squares support vector machine
VNWOA-LSSVM:	Von Neumann topology whale optimization algorithm-least squares support vector machine

Data Availability

The data used to support the findings of this study are available from the specific operational procedures upon request. Specific operational procedures, the bearing data used to support the findings of this study, have been deposited in the "http://csegroups.case.edu/bearingdatacenter/pages/download-data-file." Operating procedures: (1) Slide the mouse and find'Normal Baseline Data'. (2) Slide the mouse and find 'Normal_2'. (3) The page returns to the previous page. (4) Slide the mouse and find'48k Drive End Bearing Fault Data'. (5) Slide the mouse and find 'IR007_2'. (6) In the same way, click the mouse to download: B007_2, OR007@6_2, IR014_2, B014_2, OR014@6_2, IR021_2, B021_2, and OR021@6_2. (7) End.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This article was supported by Shanghai Polytechnic University Graduate Program Fund [EGD18YJ0003].

References

[1] C. Z. Li, J. D. Zheng, H. Y. Pan, and Q. Y. Liu, "Rolling bearing fault diagnosis method based on fine composite multi-scale scatter entropy and support vector machine," *China Mechanical Engineering*, vol. 30, no. 14, pp. 1713–1719, 2019.

[2] L. X. Ma, Y. L. Huang, H. C. Fan, and J. Y. Wang, "Motor bearing fault diagnosis on adaptive mutation particle Swarm optimization of SVM," *Electric Power Science and Engineering*, vol. 32, no. 2, pp. 66–71, 2016.

[3] Z. Huang, L. Chen, Y. Zhang, Z. Yu, H. Fang, and T. Zhang, "Robust contact-point detection from pantograph-catenary infrared images by employing horizontal-vertical enhancement operator," *Infrared Physics & Technology*, vol. 101, pp. 146–155, 2019.

[4] Y. S. Li, H. G. Zhang, B. N. Huang, and J. Han, "A distributed Newton–Raphson-based coordination algorithm for multi-agent optimization with discrete-time communication," *Neural Computing & Applications*, vol. 2018, 2018.

[5] Y. Li, H. Zhang, X. Liang, and B. Huang, "Event-triggered-based distributed cooperative energy management for multienergy systems," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2008–2022, 2019.

[6] R. Wang, Q. Sun, D. Ma, and Z. Liu, "The small-signal stability analysis of the droop-controlled converter in electromagnetic timescale," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 3, pp. 1459–1469, 2019.

[7] R. Xiong, F. Sun, Z. Chen, and H. He, "A data-driven multi-scale extended Kalman filtering based parameter and state estimation approach of lithium-ion polymer battery in electric vehicles," *Applied Energy*, vol. 113, pp. 463–476, 2014.

[8] H. Zhang, Y. Li, D. W. Gao, and J. Zhou, "Distributed optimal energy management for energy internet," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 6, pp. 3081–3097, 2017.

[9] X. Xia, "Forecasting method for product reliability along with performance data," *Journal of Failure Analysis and Prevention*, vol. 12, no. 5, pp. 532–540, 2012.

[10] G. Mukhopadhyay and S. Bhattacharya, "Failure analysis of a cylindrical roller bearing from a rolling mill," *Journal of Failure Analysis and Prevention*, vol. 11, no. 4, pp. 337–343, 2011.

[11] L. Q. Wang, H. X. Jia, D. Z. Zheng, and Z. H. Ye, "Research progress in high reliability ceramic bearing technology," *Aero Engine*, vol. 39, no. 2, pp. 6–13, 2013.

[12] Q. Y. Wan, *Research on Fault Diagnosis Method for Rolling Bearing of Helicopter Swashplate Based on DCAE-CNN*, Nanchang Hangkong University, Nanchang, China, 2019.

[13] H. Deng, L. Zhang, and X. Shu, "Feature memory-based deep recurrent neural network for language modeling," *Applied Soft Computing*, vol. 68, no. 1, pp. 432–446, 2018.

[14] B. Pan, Z. Shi, and X. Xu, "R-VCANet: a new deep-learning-based hyperspectral image classification method," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 5, pp. 1975–1986, 2017.

[15] S. Pang, J. J. del Coz, Z. Yu, O. Luaces, and J. Díez, "Deep learning to frame objects for visual target tracking," *Engineering Applications of Artificial Intelligence*, vol. 65, no. 1, pp. 406–420, 2017.

[16] J. T. Wen, C. H. Yan, J. X. Sun, and Y. L. Qiao, "Bearing fault diagnosis method based on compression acquisition and deep learning," *Journal of Scientific Instrument*, vol. 39, pp. 171–179, 2018.

[17] S. Ma and F. Chu, "Ensemble deep learning-based fault diagnosis of rotor bearing systems," *Computers in Industry*, vol. 105, no. 1, pp. 143–152, 2019.

[18] L. Manjurul and M. K. Jong, "Automated bearing fault diagnosis scheme using 2D representation of wavelet packet transform and deep convolutional neural network," *Computers in Industry*, vol. 106, no. 1, pp. 142–153, 2019.

[19] Q. H. Du and X. N. Yang, "Application of refined envelope analysis in diagnosis of rolling bearing defects," *Bearing*, vol. 3, pp. 31–34, 2004.

[20] H. Shao, H. Jiang, F. Wang, and Y. Wang, "Rolling bearing fault diagnosis using adaptive deep belief network with dual-tree complex wavelet packet," *ISA Transactions*, vol. 69, no. 1, pp. 187–201, 2017.

[21] J. B. Ali, N. Fnaiech, L. Saidi, B. Chebel-Morello, and F. Fnaiech, "Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals," *Applied Acoustics*, vol. 89, pp. 16–27, 2015.

[22] C. C. Liu, P. L. Nie, and Y. Tong, "Feature fault feature extraction based on morphological filtering and EMD-AR spectrum," *Noise and Vibration Control*, vol. 35, no. 3, pp. 159–162, 2015.

[23] Y. Yang, D. J. Yu, and J. S. Cheng, "Fault diagnosis method of rolling bearing based on empirical mode decomposition envelope spectrum," *China Mechanical Engineering*, vol. 15, no. 16, pp. 1469–1471, 2008.

[24] J. A. K. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural Processing Letters*, vol. 9, no. 3, pp. 293–300, 1999.

[25] F. L. Chen, *JADE Based Bearing Fault Diagnosis and Life Prediction Research*, Anhui University, Hefei, China, 2015.

[26] B. He, *Research on Bearing Fault Identification and Performance Degradation Evaluation Method Based on KJADE*, Anhui University, Hefei, China, 2017.

[27] H. Shao, H. Jiang, Y. Lin, and X. Li, "A novel method for intelligent fault diagnosis of rolling bearings using ensemble deep auto-encoders," *Mechanical Systems and Signal Processing*, vol. 102, no. 1, pp. 278–297, 2018.

[28] X. Yan and M. Jia, "A novel optimized SVM classification algorithm with multi-domain feature and its application to fault diagnosis of rolling bearing," *Neurocomputing*, vol. 313, pp. 47–64, 2018.

[29] M. Saimurugan, K. I. Ramachandran, V. Sugumaran, and N. R. Sakthivel, "Multi component fault diagnosis of rotational mechanical system based on decision tree and support vector machine," *Expert Systems with Applications*, vol. 38, no. 4, pp. 3819–3826, 2011.

[30] B. B. He, H. Y. Dai, and H. L. Shi, "Metro vehicle fault analysis based on time spectrum and aggregate empirical mode decomposition (EEMD) envelope spectrum analysis," *Urban Rail Transit Research*, vol. 18, no. 7, pp. 30–34, 2015.

[31] C. Liu, L. Zhu, and C. Ni, "Chatter detection in milling process based on VMD and energy entropy," *Mechanical Systems and Signal Processing*, vol. 105, pp. 169–182, 2018.

[32] M. E. Tipping and C. C. Nh, "Sparse kernel principal component analysis," *Advances in Neural Information Processing Systems*, vol. 13, pp. 633–639, 2001.

[33] H. Hoffmann, "Kernel PCA for novelty detection," *Pattern Recognition*, vol. 40, no. 3, pp. 863–874, 2007.
[34] Y. Liu, B. He, F. Liu, S. Lu, and Y. Zhao, "Feature fusion using kernel joint approximate diagonalization of eigen-matrices for rolling bearing fault identification," *Journal of Sound and Vibration*, vol. 385, pp. 389–401, 2016.
[35] Q. He, "Vibration signal classification by wavelet packet energy flow manifold learning," *Journal of Sound and Vibration*, vol. 332, no. 7, pp. 1881–1894, 2013.
[36] X. X. Ding and Q. B. He, "Two-class model based on nonlinear manifold learning for bearing health monitoring," in *Proceedings of the IEEE International Instrumentation and Measurement Technology Conference*, pp. 1–6, Taipei, Taiwan, May 2016.
[37] T. X. Wen and B. Zhang, "fault diagnosis of rolling bearing based on EMD and LSSVM," *World Science and Technology Research and Development*, vol. 36, pp. 76–79+88, 2014.
[38] X. F. Yue and H. H. Shao, "Fault diagnosis of rolling bearing based on DE-LSSVM," *Computer Measurement & Control*, vol. 23, no. 12, pp. 3933–3935, 2015.
[39] H. Xu and G. Chen, "An intelligent fault identification method of rolling bearings based on LSSVM optimized by improved PSO," *Mechanical Systems and Signal Processing*, vol. 35, pp. 167–175, 2013.
[40] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Advances in Engineering Software*, vol. 95, pp. 51–67, 2016.
[41] X. Y. Min, X. F. Xu, and Z. J. Wang, "Combining von neumann neighborhood topology with approximate-mapping local search for ABC-based service composition," in *Proceedings of the 2014 IEEE International Conference on Services Computing*, pp. 187–194, IEEE, Chicago, IL, USA, August 2014.
[42] J. Y. Xu, Y. Wang, and Z. C. Ji, "Fault diagnosis of rolling bearing based on WKELM optimized by whale algorithm," *Journal of System Simulation*, vol. 29, no. 9, pp. 2189–2197, 2017.
[43] K. A. Loparo, *Bearing Vibration Dataset*, Case Western Reserve University, Cleveland, OH, USA, 2003, http://www.eecs.cwru.edu/laboratory/bearing/download.Html.
[44] S. P. Wang, *Research on Fault Feature Extraction Method of Rolling Bearing Based on Popular Learning*, Dalian University of Technology, Dalian, China, 2013.
[45] A. L. Pei, *Research on Fault Feature Extraction Method of Rolling Bearing*, Shenyang Aerospace University, Shenyang, China, 2015.