

Research Article

Bearing Fault Diagnosis in the Mixed Domain Based on Crossover-Mutation Chaotic Particle Swarm

Tongle Xu ¹, Junqing Ji,¹ Xiaojia Kong,¹ Fanghao Zou,¹ and Wilson Wang²

¹School of Mechanical Engineering, Shandong University of Technology, Zibo 255049, China

²Department of Mechanical Engineering, Lakehead University, Thunder Bay, ON P7B 5E1, Canada

Correspondence should be addressed to Tongle Xu; jjunmqing12138@163.com

Received 19 November 2020; Revised 24 March 2021; Accepted 30 March 2021; Published 14 April 2021

Academic Editor: Hocine Cherifi

Copyright © 2021 Tongle Xu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The classification frameworks for fault diagnosis of rolling element bearings in rotating machinery are mostly based on analysis in a single time-frequency domain, where sensitive features are not completely extracted. To solve this problem, a new fault diagnosis technique is proposed in the mixed domain, based on the crossover-mutation chaotic particle swarm optimization support vector machine. Firstly, fault features are generated using techniques in the time domain, the frequency domain, and the time-frequency domain. Secondly, the weighted maximum relevance minimum redundancy (WMRMR) algorithm is adopted to reduce the dimension of the feature set and to establish the representative feature set. Thirdly, a new crossover-mutation strategy is suggested to reduce the local minima in optimization, and an optimization disturbance is added. Finally, the support vector machine is optimized using the improved chaotic particle swarm to improve fault classification diagnosis. The effectiveness of the proposed new bearing fault diagnostic technique is verified by experimental tests under different bearing conditions. Test results showed that the bearing fault classification accuracy of CMCP SO-SVM in the mixed domain was much higher than those in a single feature domain.

1. Introduction

The rolling bearings in hydraulic pumps operate in harsh environments with high pressure and high temperature, leading to operation degradation or even complete shut-down of the entire mechanical system [1, 2]. However, the fault representative features are nonlinear and nonstationary and seriously modulated by noise, and the traditional time domain and frequency domain fault diagnosis methods are not efficient to predict bearing faults, especially under changing and complex operating conditions [3, 4]. It still remains a challenging task on how to extract robust and representative features from the collected vibration signals for more reliable bearing fault diagnosis effect.

In recent years, support vector machine (SVM) and artificial neural networks have been widely concerned in the classification and recognition field. Although neural network has the characteristics of self-learning and nonlinear processing, it has some clear disadvantages, such as the “black

box” effect, expensive computing costs, and time-consuming data specimen training [5–7]. SVM is a machine learning method suitable for small sample applications. It has the characteristics of simple calculation and strong generalization ability and has been used in the fields of fault diagnosis [8], life prediction [9], and image processing [10]. Mandal and Kale [11] extended SVM to classify multiple types of faults and applied it to the multitype fault diagnosis process of the PV system. The experimental results showed that the fault diagnosis method based on multiple classifiers could accurately classify the four types of faults in the PV system. However, the multiclass SVM classification strategy is not applicable to the data processing of the single bearing fault diagnosis where this paper is focused. Yang et al. [12] designed a support vector machine fault diagnosis algorithm based on RBF kernel function and applied it to the fault diagnosis of metal roof. A least squares support vector machine was proposed by Suykens to transform the training process into a linear equation solution, which could simplify

the training process and improve the classification accuracy [13]. However, the classification accuracy of the SVM algorithm is limited due to a large amount of training data and the difficulty in determining the penalty factor and kernel function.

To solve these problems of SVM algorithm, optimization algorithms such as genetic algorithm (GA) and particle swarm optimization algorithm (PSO) are used by some scholars to optimize the penalty parameter and kernel function parameter. Ravasan et al. [14] used PSO to optimize the penalty factor and kernel function parameter of SVM to improve the generalization ability, robustness, and sparseness of SVM. Rajamohana and Umamaheswari [15] combined the improved binary particle swarm optimization with shuffled frog leaping algorithm to improve the optimization of the classifier. Wang et al. [16] used chaotic particle swarm algorithm to optimize SVM for receivers' fault diagnosis in avionics navigation system. Tharwat and Hassanien [17] used quantum particle swarm optimization to optimize SVM, and the model had the advantages of good generalization and strong classification ability comparing with grid search and GA. Jiang et al. [18] improved the adaptive weight and time factor of the PSO, and the optimized SVM classification model can adaptively adjust the local search and global search capabilities. Nevertheless, the PSO algorithm has weak local search ability and little information exchange, and it is difficult to escape local minima in optimization especially with high-dimensional sample sets. Therefore, a new crossover-mutation strategy and optimization perturbation are introduced into the chaotic particle swarm optimization (CPSO) algorithm to improve the communication between particles and avoid the particles being trapped in local optimization in this paper. Additionally, SVM is optimized by the cross-mutation chaotic particle swarm optimization (CMCPSO) algorithm to establish a new fault classification model (CMCPSO-SVM).

The above fault diagnosis techniques mainly focused on time-frequency characteristics of the decomposed signal, and the characteristics of the original signal in the time domain and the frequency domain were usually ignored. Thus, the fault representative features and deep features may not be recognized completely and accurately. Scholars have done a lot of research to solve the abovementioned problems. Jia et al. [19] proposed a new complex empirical mode decomposition method to process the fault characteristics of wind turbines in the time domain, frequency domain, and time-frequency domain. Xue and Zhou [20] proposed a hybrid fault diagnosis method for rolling bearings based on state characteristic data in a mixed domain, and the fault signals were classified under twelve working conditions. Results showed that the diagnosis method was more suitable for practical applications. Fei et al. [21] combined the four information entropies in the time domain, frequency domain, and time-frequency domain with the two signals of vibration signal and acoustic emission signal to diagnose bearing faults. The experimental results showed that the fault diagnosis method based on multidomain could better reflect the signal characteristics, and the classification effect was more obvious. Even so, the dimensionality of the fault set in

the abovementioned literature is too high, which leads to a large amount of calculation and a long time during the diagnosis processing. Consequently, to improve the calculation efficiency of the algorithm, the WMRMR algorithm is used to reduce the dimensionality of the multidomain fault feature set in this paper.

To solve the aforementioned problems, a new bearing fault diagnosis technique based on crossover-mutation chaotic particle swarm in the mixed domain is proposed in this paper. Firstly, the extracted features in the time domain, frequency domain, and time-frequency domain are combined to establish a mixed domain feature set. Secondly, the WMRMR technique is used to select the representative feature subset by sorting the features. Finally, the new crossover-mutation strategy and an optimization disturbance are added to the traditional chaotic particle swarm to optimize SVM, which can improve the accuracy of fault classification.

2. Establish Optimal Fault Feature Set

2.1. Extract Mixed Domain Features

2.1.1. Time Domain Feature Extraction. The time domain feature is a statistical feature that can intuitively discover important feature information through observation [22]. In this paper, twelve statistical parameters are selected to constitute the time domain feature set. Among them, the dimensional parameters T_1-T_8 include maximum value, minimum value, peak value, effective value, square root amplitude, absolute mean, mean square value, and kurtosis. Dimensionless parameters T_9-T_{12} are kurtosis parameter, pulse parameter, margin parameter, and peak parameter.

2.1.2. Frequency Domain Feature Extraction. The fault features represented in the frequency domain of vibration signals can be described according to the distribution of spectral information [23]. The frequency domain feature sets selected in this paper include average frequency, frequency variance, and centroid frequency, which are expressed as $F_{13}-F_{15}$, respectively.

2.1.3. Time-Frequency Domain Feature Extraction. Bearing fault related features are usually transient and time-varying. Therefore, rolling bearing faults cannot be found intuitively and reliably by the simple time-frequency domain features [24, 25]. Compared with the singular value of Hilbert spectrum and wavelet energy, the dual-LMD morphological filtering algorithm can not only effectively extract time-frequency domain features, but also greatly reduce the impact of noise on the extraction results [26]. Thus the time-frequency domain features will be extracted by the dual-LMD morphological filtering algorithm, which will be described next.

(1) Double LMD Morphological Filtering Algorithm. The LMD algorithm can adaptively decompose the bearing vibration signals into a series of product function (PF) of

instantaneous frequency components, where each PF component is a modulation combination of a single envelope signal and a pure frequency signal. In order to extract fault features in the time-frequency domain more accurately and effectively, a dual-LMD morphological filtering method is adopted [27]. The specific operating steps are as follows:

- (a) Several PF_1 components are obtained by LMD decomposition of the measurement signal $x(t)$. PF_1' is acquired by correlation analysis of the PF_1 components.
- (b) PF_1' is denoised by combined filtering, and PF_1'' is acquired by the sinusoidal structural elements process.
- (c) The combined reconstruction and noise reduction of PF_1'' are carried out to obtain the signal $x_1(t)$, and component PF_i'' is obtained by LMD morphological filtering decomposition of $x_1(t)$.

In order to improve the noise reduction effect and the accuracy of feature extraction, a new signal processing method of LMD decomposition—noise reduction—signal recombination—LMD decomposition is adopted to reduce the vector dimension and the sparsity of the SVM from the source.

(2) *Time-Frequency Domain Feature Based on PF Component Normalization*. The feature energy extraction steps of PF_i'' component are as follows:

- (a) The total energy of each component PF_i'' is calculated as follows:

$$E_i = \int_{-\infty}^{+\infty} |PF(t)|^2 dt, \quad i = 1, 2, 3, \dots, N. \quad (1)$$

- (b) The characteristic matrix C_N is composed of elements E_i , such that

$$C_N = [E_1, E_2, E_3, \dots, E_N]. \quad (2)$$

The larger value of the total energy E_i calculated by equation (1) will cause difficulty in the subsequent feature selection [28]. Therefore, the feature vector matrix C_N is normalized by $E = (\sum_{i=1}^N |E_i|^2)^{(1/2)}$, and the feature vector matrix C_N' is computed as $C_N' = [(E_1/E), (E_2/E), (E_3/E), \dots, (E_N/E)]$. Since the energy of the fault feature after LMD decomposition is mainly concentrated in the first five PF components, the feature vector set in the time-frequency domain will be composed of the first five vectors of C_N' , or $[TF_{16}, \dots, TF_{20}] = [C_1', \dots, C_5']$.

(3) *Time-Frequency Feature Characteristics of PE Based on PF Components*. The PE is a parameter that can express the irregularities and nonlinearities of the system quantitatively, and the PE can reflect the system characteristics [29]. The PE is defined as

$$H_{PE} = - \sum_{j=1}^k P_j \ln P_j, \quad (3)$$

where P_j is the probability of symbol sequence occurrence, $k = N - (m - 1) \times \tau$, N is the sequence length of the original signal, and m and τ are the respective embedding dimension and delay time.

In general, bearing fault feature information is mainly concentrated in the first five PF components, so the first five PF component permutation entropies will be extracted as the time-frequency domain set, which is written as $[TF_{21}, \dots, TF_{25}] = [PE_1, \dots, PE_5]$.

2.2. Weighted Maximum Relevance Minimum Redundancy Algorithm. The maximum relevance minimum redundancy (MRMR) algorithm can select features by mutual information, correlation or distance similarity scores. Similar to the maximum dependent feature selection algorithm [30], MRMR aims to punish the feature correlation by the feature redundancy [31].

The mutual information between random variables x and y is determined by

$$I(x; y) = \iint p(x, y) \log \frac{p(x, y)}{p(x)p(y)} dx dy, \quad (4)$$

where $p(x)$ and $p(y)$ are probability densities and $p(x, y)$ is the joint probability density.

The MRMR evaluation index is determined by

$$\begin{cases} \max D(S, c), D = \frac{1}{S_N} \sum_{x_i \in S} I(x_i; c), \\ \min R(S), R = \frac{1}{S_N^2} \sum_{x_i, x_j \in S} I(x_i; x_j), \end{cases} \quad (5)$$

where S is the feature set of the signal; S_N is the number of features; c is the target category of the signal; $I(x_i; c)$ is the mutual information between x_i and c ; and $I(x_i; x_j)$ is the mutual information between the samples x_i and x_j . D is the mean value of $I(x_i; c)$ in S , which reflects the correlation between the feature set and target category of the signal. R is the mutual information size of the features, which reflects the redundancy between features.

In order to minimize the redundancy between features and maximize the relevance, the MRMR evaluation criterion is defined as

$$\max_z \Phi(D, R), \Phi_z = D - R, \quad z = 1, 2. \quad (6)$$

The traditional MRMR method cannot select optimal features of complex feature sets with high relevance and redundancy. Therefore, the weighting coefficient of the evaluation criterion will be optimized. The revised evaluation criterion is expressed as

$$\max \Phi(D, R), \quad \Phi = \delta D - (1 - \mu)R, \quad (7)$$

where $0 < \delta \leq 1$. If $\delta = 0.5$, equations (6) and (7) are identical.

3. Construction of CMPSO-SVM Model

The SVM is a linear classifier that can maximize the distance between categories according to certain criteria. SVM can search for the optimal hyperplane $w^T x + b = 0$, so as to classify positive and negative information. To ensure the accuracy of classification, penalty factor p and relaxation factor ξ_i ($\xi_i \geq 0$, $i = 1, 2, \dots, n$) are introduced. Furthermore, the objective function and constraints are transformed into dual form through the Lagrange multiplier algorithm. In addition, compared with other kernel functions, the radial basis kernel function (RBF) is simpler and more universal [32, 33]. As illustrated in equation (8), it only needs to determine one parameter θ , which will be used as the kernel function in this SVM:

$$H(x_i, y_i) = \exp\left(-\frac{\|x_i - y_i\|^2}{\theta^2}\right). \quad (8)$$

3.1. Cross-Mutation Chaotic Particle Swarm Optimization. The PSO has the advantages of convenient operation, fast convergence, and good optimization effect [34]. The traditional mathematical theory is as follows.

Assuming that a swarm contains M -dimensional particles and the size of the particle swarm is N , the spatial position x_i and velocity v_i of each particle in the swarm can be expressed as

$$\begin{cases} x_i = (x_{i1}, x_{i2}, \dots, x_{ia}, \dots, x_{iM}), \\ v_i = (v_{i1}, v_{i2}, \dots, v_{ia}, \dots, v_{iM}). \end{cases} \quad (9)$$

In each optimization iteration of the particle swarm, the position and velocity need to be tracked dynamically. The velocity v_{ij}^a and position x_{ij}^a of the i -th particle in the a -th iteration can be represented as

$$\begin{cases} v_{ij}^a = wv_{ij}^{a-1} + c_1 \text{rand}_1(p_{ij} - x_{ij}^{a-1}) + c_2 \text{rand}_1(p_{gj} - x_{gj}^{a-1}), \\ x_{ij}^a = x_{ij}^{a-1} + v_{ij}^a, \end{cases} \quad (10)$$

where w is the inertia weight; rand_1 and rand_2 are random numbers in $[0, 1]$; c_1 and c_2 are the self-learning weight coefficient and the social learning weight coefficient, respectively; and $c_1, c_2 \in [1.5, 2]$. p_{ij} is the j -th component historical optimal search position of the i -th particle, and p_{gj} is the j -th component best position of the global optimization.

However, due to the weak local search ability and lack of information communication, the PSO algorithm is easy to fall into a cycle of local minima for complex sample sets. Some methods were proposed by scholars to solve this problem, such as chaotic motion, which has advantages of sensitive initial conditions and good randomness to optimize the PSO algorithm [35]. So the classic logistic equation is used to construct the chaotic motion sequence:

$$x_{n+1} = \mu x_n (1 - x_n), \quad i = 1, 2, \dots, n, \quad (11)$$

where μ is the chaos coefficient and n is the sample size.

The CPSO algorithm can optimize the inertial weight w , self-learning weight coefficient c_1 , and the social learning weight coefficient c_2 by equation (11). On the contrary, since the fitness of the particle swarm is determined by the position of the particle, the particle swarm variance σ^2 can be used to assess the ‘‘premature’’ situation in the CPSO algorithm. When σ^2 is less than the given judgment value F , the particle is considered being fallen into a local minima, and the position and velocity of the particle have to be redefined, and σ^2 can be represented as

$$\sigma^2 = \sum_{i=1}^n \frac{S_i - \bar{S}}{S_i}, \quad (12)$$

where S_i is the fitness of the i -th particle and \bar{S} is the average of all fitness in the particle swarm.

However, the CPSO algorithm is likely to fall into the local minima for samples with complex interference. In order to solve this shortcoming of the CPSO algorithm, a crossover-mutation chaos particle swarm optimization algorithm (CMPSO) is proposed.

3.1.1. Adaptive Crossover-Mutation Strategy. Based on the crossover-mutation process of genetic algorithm, a new adaptive crossover-mutation strategy [36] is adopted to improve the chaos particle swarm optimization in this paper. The purpose is to make each particle fully communicate with others, and a new communication mechanism is established to prevent local minima.

3.1.2. Crossover Strategy. In order to increase the communication between the particles, a certain number of particles are randomly selected to perform the cross operation according to the cross probability P_c . Then the cross particles will replace the original particles, without changing the number of particles. The adaptive crossover probability expression is represented as

$$P_c = \begin{cases} F_1 \cdot \frac{S_{\max} - S_c}{S_{\max} - \bar{S}_c}, & S_c \geq \bar{S}_c, \\ F, & S_c < \bar{S}_c, \end{cases} \quad (13)$$

where S_{\max} is the maximum fitness of the selected particles, S_c is the fitness of participating cross particles, \bar{S}_c is the average of S_c , and F_1 is a constant and usually $F_1 = 1$.

3.1.3. Mutation Strategy. To prevent local minima of the CPSO algorithm, an adaptive mutation strategy is suggested in this work. The adaptive mutation probability P_m will be used for mutation operation to improve the problem of swarm singleness and enrich the diversity of the swarm. The adaptive mutation probability formula is expressed as

$$P_m = \begin{cases} F_2 \cdot \frac{S_{\max} - S_m}{S_{\max} - \bar{S}_m}, & S_m \geq \bar{S}_m, \\ F_2, & S_m < \bar{S}_m, \end{cases} \quad (14)$$

where S_m is the fitness of participating mutant particles, \bar{S}_m is the average of S_m , and F_2 is a constant ($F_2 = 0.5$).

3.1.4. Optimization Disturbance. When the globally optimal and locally optimal particles are found, better particles in their neighborhood may exist. Thus, an optimization perturbation will be added into the global and local optimal solutions to improve the global optimization ability of the algorithm. The particle velocity formula is expressed as

$$v_{ij}^a = wv_{ij}^{a-1} + c_1 \text{rand}_1(p_{ij} - (1 \pm \chi \text{rand})x_{ij}^{a-1}) + c_2 \text{rand}_1(p_{gj} - x_{gj}^{a-1}). \quad (15)$$

By simulation, it is found that $\chi = 0.01$ can provide best performance of the algorithm.

3.2. Cross-Mutation Chaotic Particle Swarm Optimization Support Vector Machine. The key part of SVM is the optimization of the penalty factor p and the parameter θ , which are directly related to the accuracy of the fault classification. In this paper, the crossover-mutation chaotic particle swarm is used to optimize SVM. The optimization steps are illustrated in Figure 1, where k is the number of iterations and k_{\max} is the maximum number of iterations.

- (1) Initialize chaotic particle swarm parameters, including inertial weight w , self-learning and social learning weight coefficients c_1 and c_2 , maximum number of iteration k_{\max} , swarm size, optimization disturbance χ , penalty factor p , and parameter θ .
- (2) Initialize the position and velocity of the particles. The chaotic swarm is generated by logistic mapping, the optimal position of particles is determined, and the fitness is calculated.
- (3) Update the optimal position and velocity of particles. The update position, velocity, and the fitness variance σ^2 of particles are calculated, respectively.
- (4) Crossover-mutation operations are performed based on P_m and P_c to enrich the diversity of particles and form the new swarm. Furthermore, the individual and global optimal fitness values are updated.
- (5) Optimize p and θ in the SVM using the optimal parameters obtained in steps (3) and (4), and construct the CMCP SO-SVM model.

4. Performance Evaluation for Bearing Fault Diagnosis

4.1. Overview. The proposed CMCP SO-SVM technique will be implemented for rolling bearing fault diagnosis in this section. Its effectiveness will be examined by the use of the

comparison with other related techniques. The overall flow chart of rolling bearing fault diagnosis is shown in Figure 2, which can be divided into four steps.

The first step is the vibration signal acquisition using an accelerator and a data acquisition model. The second step is mixed domain feature extraction by the statistical analysis, spectral analysis, and dual-LMD methods, and a mixed domain feature set is established.

The third step is the selection of representative features. The weight factor δ is selected by maximum relevance minimum redundancy algorithm, and the step size is 0.1. The low-dimensional representative feature set Z_i is obtained by equation (7).

The fourth step is to apply the proposed CMCP SO-SVM technique for bearing fault diagnosis. The selected representative features are divided into training samples and test samples, and they are input into the CMCP SO-SVM classification model to diagnose rolling bearing faults.

4.2. Experimental Setup. Figure 3 shows the experimental setup built by Qianpeng Company for bearing testing. It is mainly composed of a motor, bearing chocks, and discs, and the related parameters are summarized in Table 1.

Tested bearings have simulated defect on the inner ring, outer ring, and rolling element of the normal bearings, tested under load and no-load conditions, respectively. The faults of the inner ring, outer ring, and ball failure under the load condition are recorded as LIR, LOR, and LBR, respectively, and the faults of inner ring, outer ring, and ball failure under no-load condition are denoted as IR, OR, and BR, respectively, as shown in Figure 4. In addition, the vibration signals are collected by the use of a DH620 piezoelectric accelerometer. The first 102,400 sampling points under each working condition are divided into 50 groups on average, and each group contains 2048 samples. To avoid continuity between sample sets, the first 25 groups in each operating condition are used for training, and the other 25 groups are used for testing.

Step 1. The bearing representative features under six operating conditions are extracted in the time domain, frequency domain, and time-frequency domain, respectively. The results of dual-LMD extraction in time-frequency domain are shown in Figure 5, and the results of energy normalization are shown in Table 2. In order to rank the selected features, the features in the three domains are numbered, where, the maximum value, minimum value, peak value, effective value, square root amplitude, absolute mean value, mean square value, kurtosis, kurtosis index, pulse index, edge index, and peak index of time domain features are marked as T_1 – T_{12} , respectively. The average frequency, frequency variance, and centroid frequency of frequency domain features are marked as F_{13} – F_{15} . The first five vectors obtained from the PF energy in the time-frequency domain are denoted as $[\text{TF}_{16}, \dots, \text{TF}_{20}] = [C'_1, \dots, C'_5]$, and the first five vectors of the PF permutation entropy are denoted as $[\text{TF}_{21}, \dots, \text{TF}_{25}] = [\text{PE}_1, \dots, \text{PE}_5]$.

Step 2. The weight factor δ of the WMRMR algorithm is recorded as $\delta_i = 0.1, \dots, 1$, and the step length is 0.1. The

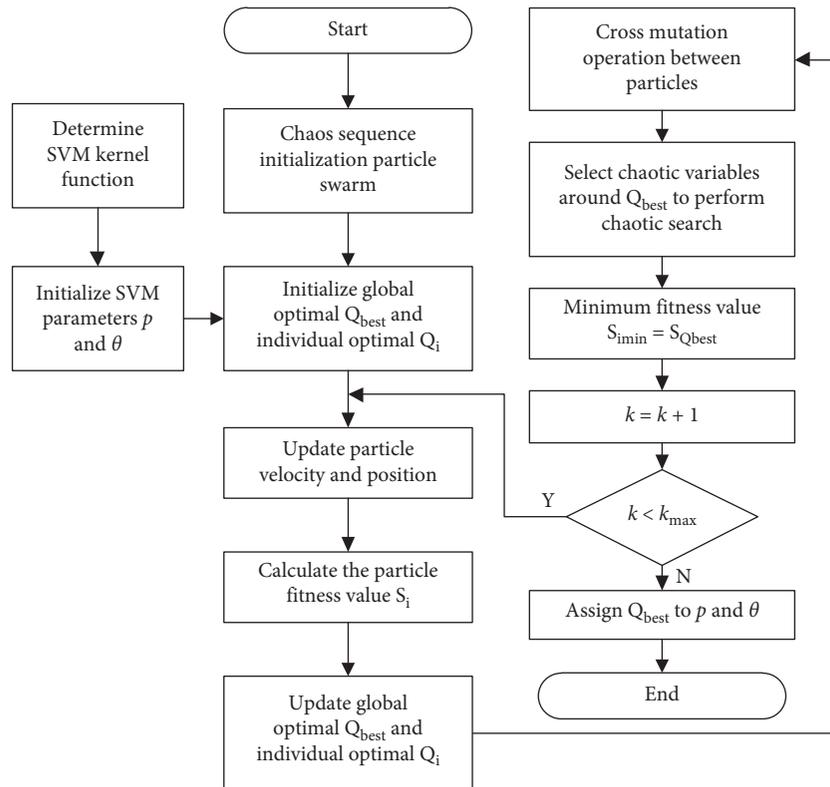


FIGURE 1: The process of SVM optimized by CMCP SO.

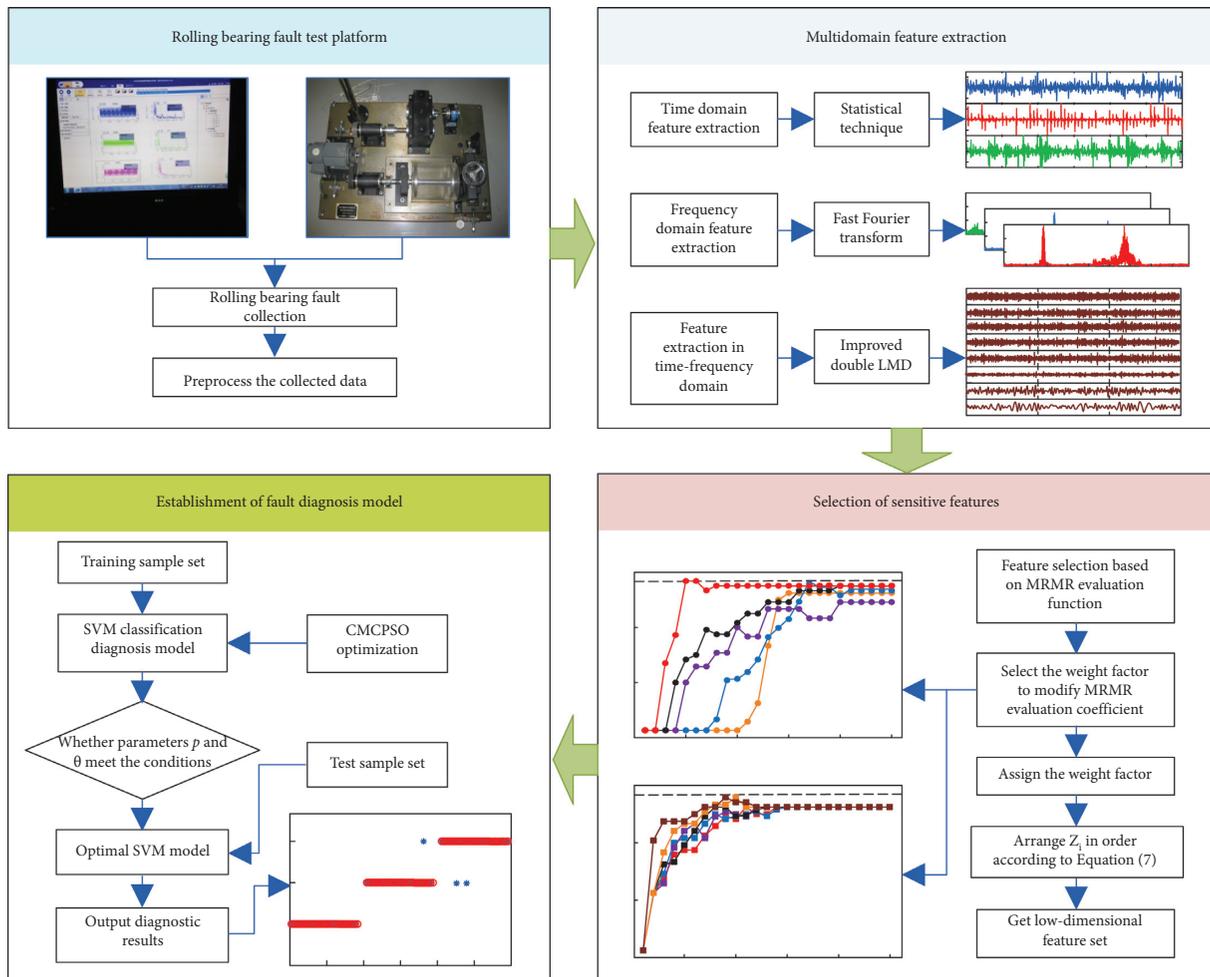


FIGURE 2: The flow chart of rolling bearing fault diagnosis in the mixed domain.

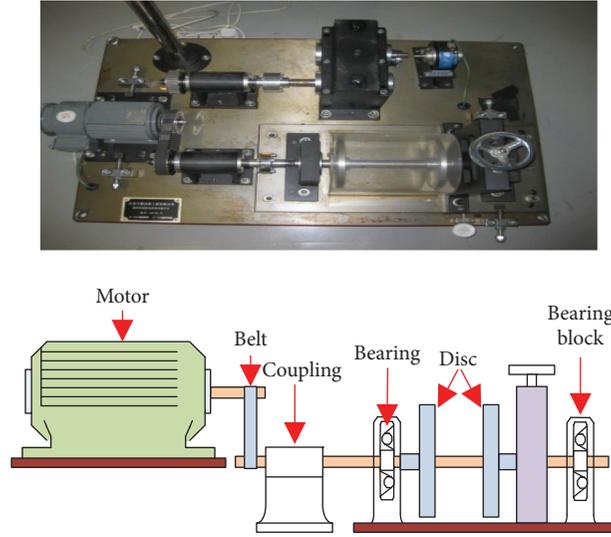


FIGURE 3: The bearing fault experiment platform.

TABLE 1: Experimental parameters.

Rotating speed	Bearing	Parameter (mm)
940r/min	Outer diameter	52
	Inner diameter	25
	Pitch circle diameter	39
	Basic rated dynamic load	12.6
	Basic rated static load	3.3

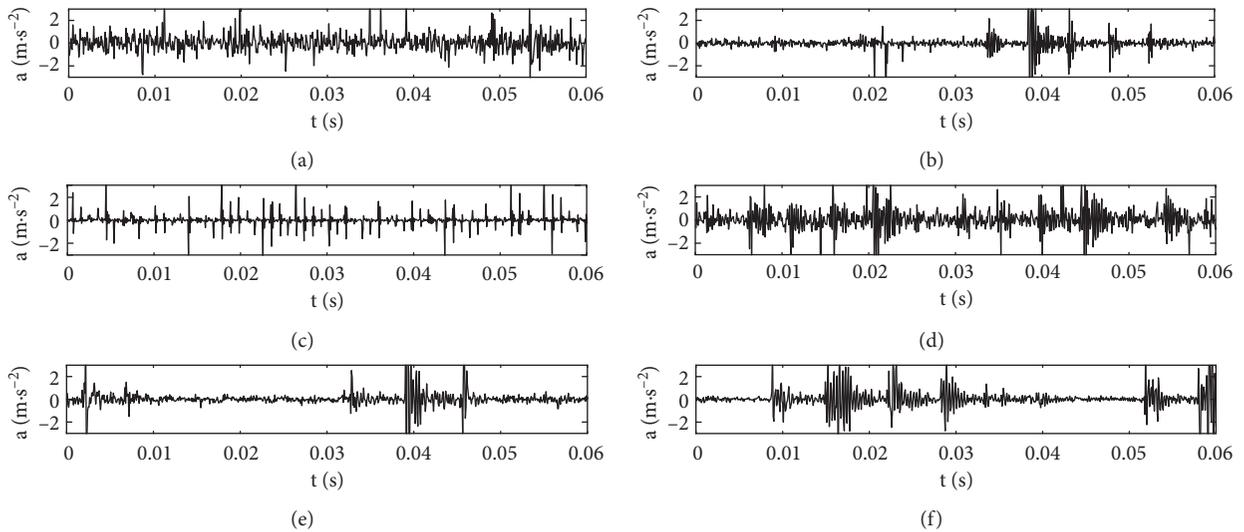


FIGURE 4: Fault signals under different working conditions. (a)~(f) are fault signals of BR, LBR, OR, LOR, IR, and LIR, respectively.

feature set corresponding to different δ_i values is calculated using the WMRMR algorithm, and the candidate feature set Z_i corresponding to the weight factor δ_i is obtained by sequence arrangement in equation (7). The results are summarized in Table 3.

Step 3. The selected low-dimensional feature subset Z_i is input into the CMCP SO-SVM classification model for

training. Bearing fault diagnostic classification is undertaken by the diagnostic process in Section 3.

Step 4. The accuracy of classification results with different weight factors δ is compared to determine the optimal weight factor δ_o . The feature subset Z_o corresponding to δ_o is selected as the optimal low-dimensional feature subset. Furthermore, Z_o is input into the CMCP SO-SVM

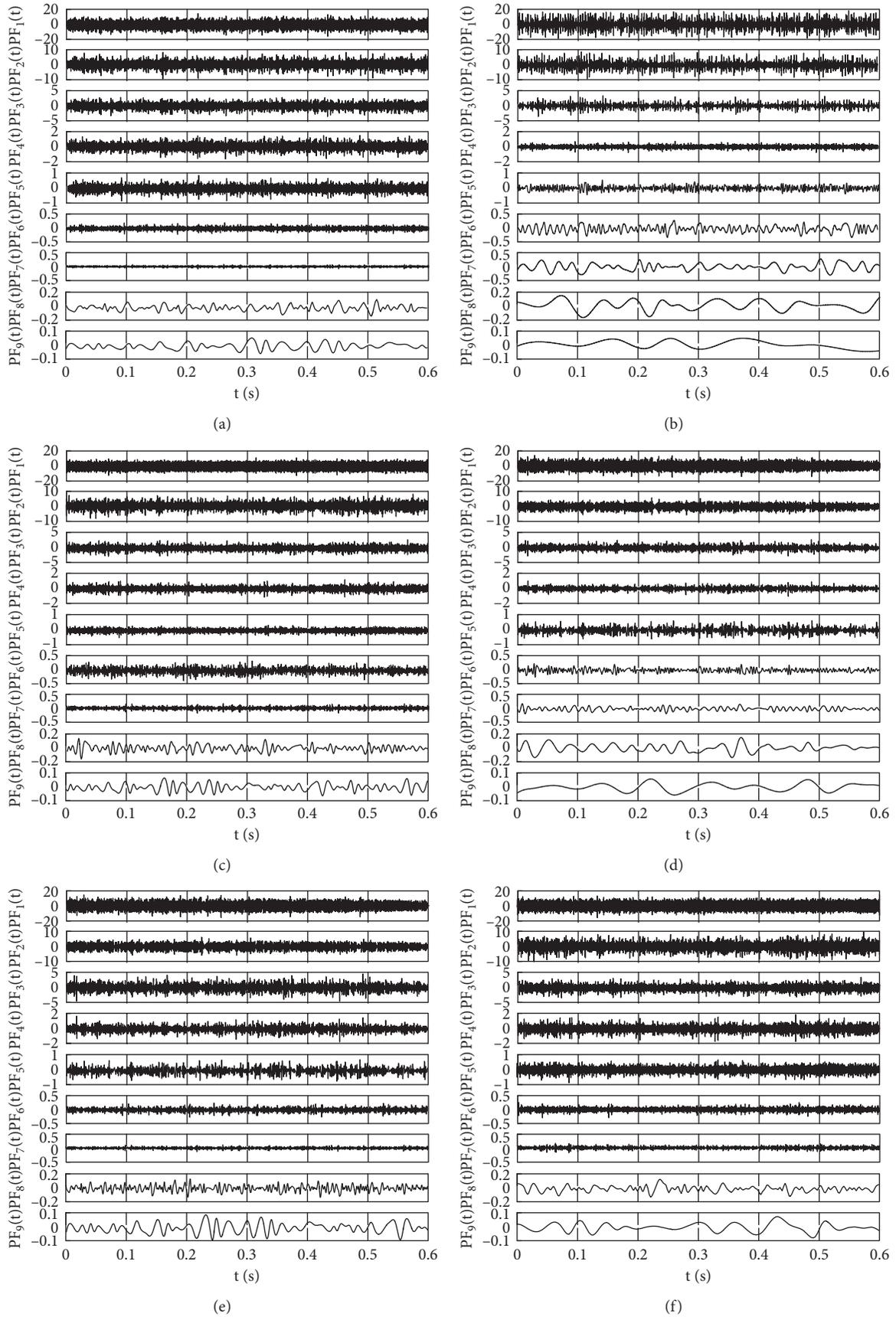


FIGURE 5: The double LMD extraction results. (a)~(f) are the double LMD extraction results of BR, LBR, OR, LOR, IR, and LIR, respectively.

TABLE 2: Energy normalization results of dual-LMD decomposition.

Operating status	Fault type	E1	E2	E3	E4	E5	E6	E7	E8	E9
No-load	Outer ring failure	0.5004	0.1578	0.1302	0.1270	0.0921	0.0651	0.0406	0.0427	0.0241
	Ball failure	0.5589	0.1731	0.1187	0.0754	0.0393	0.0224	0.0065	0.0031	0.0026
	Inner ring failure	0.4514	0.2081	0.1524	0.1055	0.0652	0.0434	0.0079	0.0065	0.0050
With load	Outer ring failure	0.3425	0.3047	0.1484	0.1337	0.0343	0.0253	0.0071	0.0027	0.0013
	Ball failure	0.3435	0.1998	0.1453	0.1428	0.0702	0.0391	0.0045	0.0023	0.0014
	Inner ring failure	0.4001	0.2835	0.1318	0.0925	0.0732	0.0165	0.0048	0.0029	0.0016

TABLE 3: Feature ranking.

δ Value	Feature order
0.1	TF ₁₉ , T ₄ , F ₁₃ , F ₁₄ , F ₁₅ , TF ₁₇ , T ₁₂ , TF ₁₈ , TF ₂₀ , TF ₂₅ , T ₅ , T ₁ , T ₂ , TF ₂₄ , TF ₂₃ , TF ₂₂ , TF ₂₁ , TF ₁₆ , T ₆ , T ₈ , T ₉ , T ₃ , T ₇ , T ₁₀ , T ₁₁
0.2	TF ₁₉ , T ₄ , F ₁₃ , F ₁₄ , F ₁₅ , TF ₁₇ , T ₁₂ , TF ₁₈ , T ₈ , T ₉ , T ₃ , T ₁₀ , T ₁₁ , TF ₂₀ , TF ₂₅ , T ₅ , T ₇ , TF ₂₃ , TF ₂₄ , T ₁ , T ₂ , TF ₂₂ , TF ₂₁ , TF ₁₆ , T ₆
0.3	TF ₁₉ , T ₄ , F ₁₃ , F ₁₄ , F ₁₅ , TF ₂₂ , TF ₁₇ , T ₃ , T ₁₂ , T ₂ , T ₈ , T ₉ , TF ₁₈ , TF ₂₁ , TF ₂₄ , T ₇ , TF ₁₆ , TF ₂₀ , TF ₂₅ , T ₁ , T ₅ , T ₆ , TF ₂₃ , T ₁₀ , T ₁₁
0.4	TF ₁₉ , T ₄ , F ₁₃ , F ₁₄ , T ₁₁ , TF ₁₇ , TF ₂₄ , TF ₂₁ , T ₂ , T ₈ , T ₉ , TF ₂₂ , TF ₂₃ , F ₁₅ , TF ₁₈ , T ₁ , T ₅ , TF ₁₆ , TF ₂₀ , TF ₂₅ , T ₃ , T ₆ , T ₇ , T ₁₀ , T ₁₂
0.5	TF ₁₉ , T ₄ , F ₁₃ , T ₁₁ , TF ₁₆ , TF ₁₇ , T ₅ , T ₆ , TF ₂₁ , TF ₂₂ , TF ₂₄ , F ₁₃ , TF ₂₀ , T ₂ , T ₃ , F ₁₅ , T ₈ , F ₁₄ , T ₇ , T ₉ , T ₁₀ , T ₁ , T ₁₂ , TF ₂₃ , TF ₁₈ , TF ₂₅
0.6	TF ₁₉ , TF ₂₂ , T ₉ , T ₁₀ , TF ₂₃ , T ₄ , T ₁₂ , TF ₂₁ , T ₈ , TF ₂₀ , TF ₂₄ , T ₃ , T ₇ , T ₈ , TF ₁₆ , T ₁ , F ₁₃ , T ₅ , T ₂ , F ₁₄ , F ₁₅ , T ₆ , T ₁₁ , TF ₁₇ , TF ₁₈ , TF ₂₅
0.7	TF ₁₉ , T ₁₁ , TF ₁₇ , T ₁₂ , T ₁₀ , TF ₂₂ , T ₃ , TF ₂₁ , TF ₂₄ , T ₇ , T ₁ , TF ₂₃ , T ₆ , T ₂ , F ₁₃ , F ₁₄ , F ₁₅ , TF ₁₆ , TF ₂₀ , T ₄ , T ₅ , T ₈ , T ₉ , TF ₁₈ , TF ₂₅
0.8	TF ₁₉ , T ₇ , T ₁₁ , T ₁₂ , TF ₁₇ , T ₁₀ , TF ₂₁ , T ₃ , TF ₂₂ , T ₁ , TF ₂₃ , TF ₂₄ , T ₆ , T ₂ , F ₁₃ , F ₁₄ , F ₁₅ , TF ₁₆ , TF ₂₀ , T ₄ , T ₅ , T ₈ , T ₉ , TF ₁₈ , TF ₂₅
0.9	TF ₁₉ , T ₇ , T ₁₁ , T ₁₂ , TF ₁₇ , TF ₂₁ , T ₁₀ , T ₃ , TF ₂₄ , T ₁ , TF ₂₂ , TF ₂₃ , T ₆ , T ₂ , F ₁₃ , F ₁₄ , F ₁₅ , TF ₁₆ , TF ₂₀ , T ₄ , T ₅ , T ₈ , T ₉ , TF ₁₈ , TF ₂₅
1	TF ₁₉ , T ₇ , T ₁₁ , TF ₁₇ , T ₁₂ , TF ₂₁ , T ₁₀ , T ₃ , T ₁ , TF ₂₄ , TF ₂₂ , TF ₂₃ , T ₆ , T ₂ , F ₁₃ , F ₁₄ , F ₁₅ , TF ₁₆ , TF ₂₀ , T ₄ , T ₅ , T ₈ , T ₉ , TF ₁₈ , TF ₂₅

classification model for fault classification. Figure 6 shows different feature subset classification accuracy rates corresponding to different δ values.

As can be seen from Figure 6, the classification accuracy tends to be stable with the increase of feature number, but the convergence speeds are different. When $\delta = 0.1, 0.3, 0.4, 0.5$, its classification accuracy can be stabilized at about 98.5%, and the number of features with stable convergence is 21, 15, 8, and 13, respectively. When $\delta \in [0.6, 0.7, 0.8, 0.9, 1]$, the number of features with curve convergence is about 14, and the classification accuracy is around 97%. Correspondingly, it can be seen that the classification accuracy is the highest when $\delta = 0.4$. Therefore, the first eight features are selected as the optimal feature subset \tilde{Z} when $\delta = 0.4$ in this paper, and the corresponding features are C'_4 , effective value, average frequency, frequency variance, pulse index, C'_2 , PE_4 , and PE_1 .

To verify the effectiveness of the CMCPPO-SVM classification model, the low-dimensional optimal feature subset \tilde{Z} is input to the PSO-SVM and CPSO-SVM diagnosis models for fault diagnosis to prove the classification reliability of the new fault diagnosis model. Figure 7 shows the classification results using the three diagnosis models, which are summarized in Table 4.

It can be seen from Figure 7 that the best fitness value of the three optimized SVM classifiers should be stable at about 98. The actual fitness values of the PSO-SVM classifier is mainly concentrated in [60, 70], which is quite different from the expected value. The fitness values of the CPSO-SVM classifier are mainly concentrated in [70, 80]. Although it is higher than those of the PSO-SVM classifier, it still cannot reach the expected target. Apparently, the fitness value of the proposed CMCPPO-SVM classifier can reach about 94, with less variance and closer to expectations, which is superior to the other two methods. From the classification results of the three classifiers in Table 4, it can be seen that the proposed CMCPPO-SVM outperforms other related classifiers with an average accuracy 98.67%, compared with the PSO-SVM (88%) and CPSO-SVM (90.51%), respectively.

Since the iteration number of the classifier is also an important factor which can affect the classification performance, the fault diagnosis accuracy and time under different iteration number working conditions are listed in Table 5. It is seen that the fault diagnosis accuracy is different under four different maximum iteration number conditions, and the average diagnostic accuracy is 97.10%, 98.67%, 96.86%, and 96.40%, respectively. Therefore, the classification

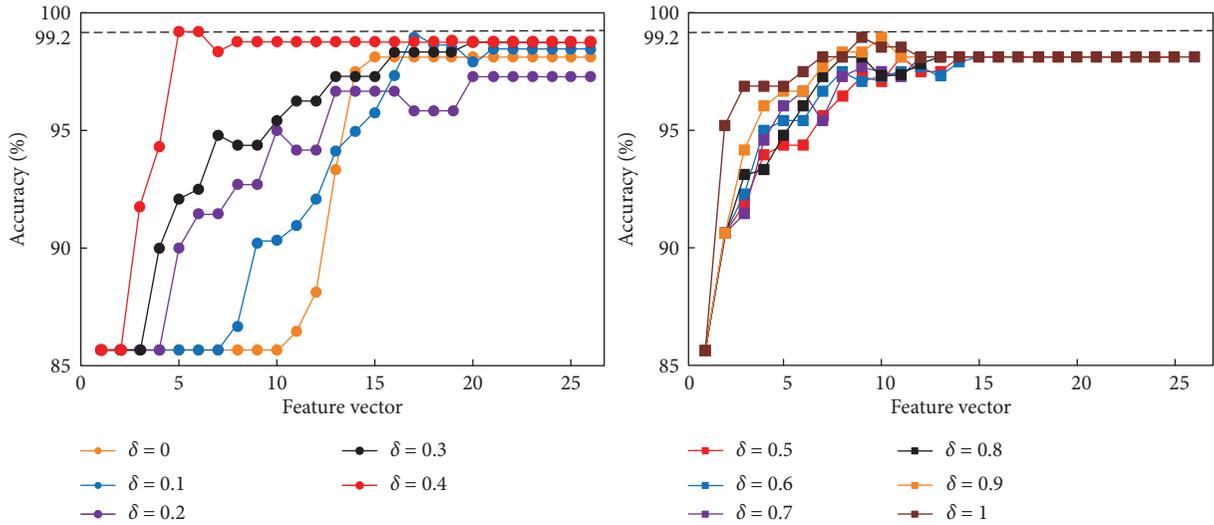


FIGURE 6: Different feature subset classification accuracy rate curves corresponding to different δ values.

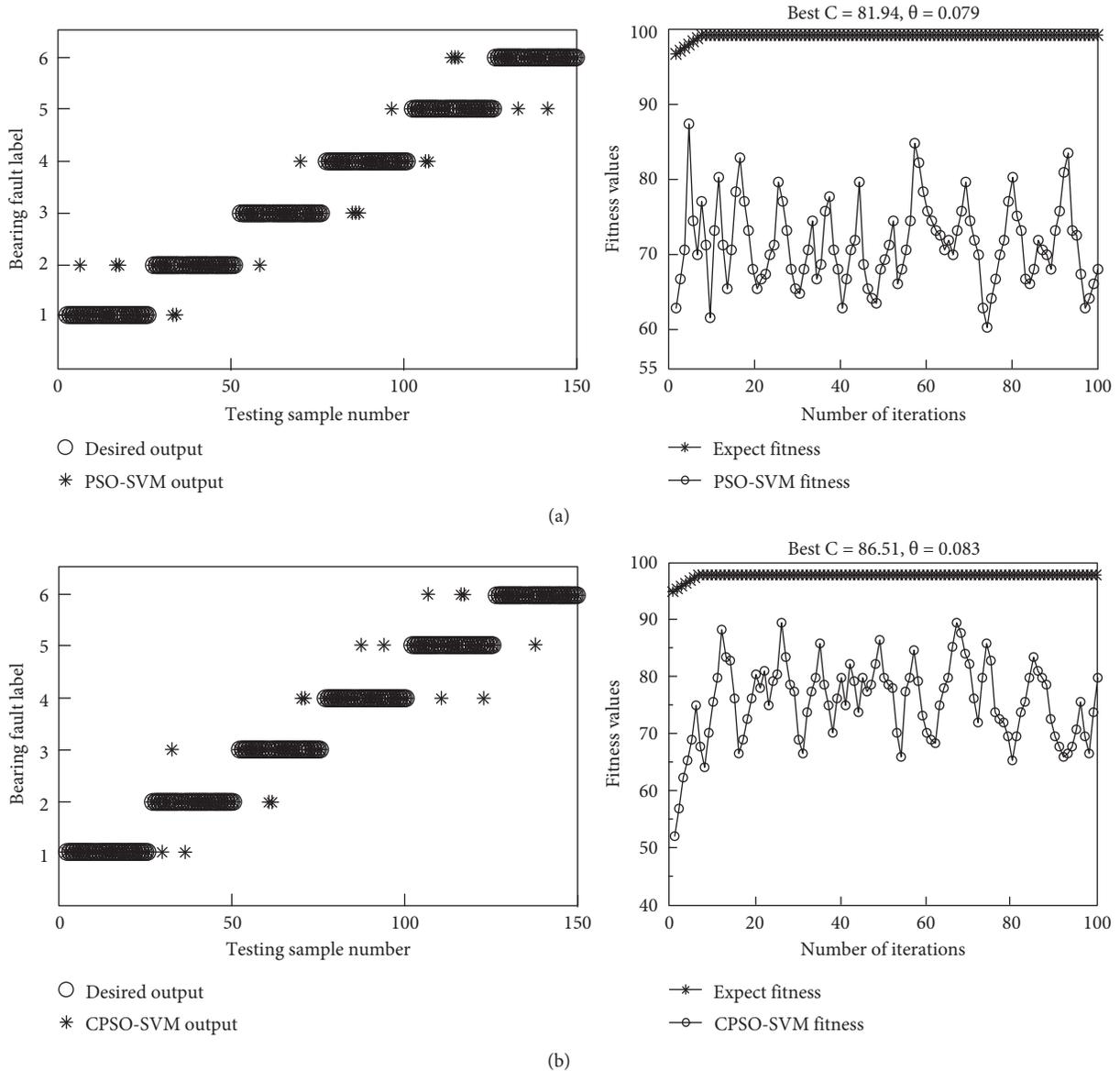


FIGURE 7: Continued.

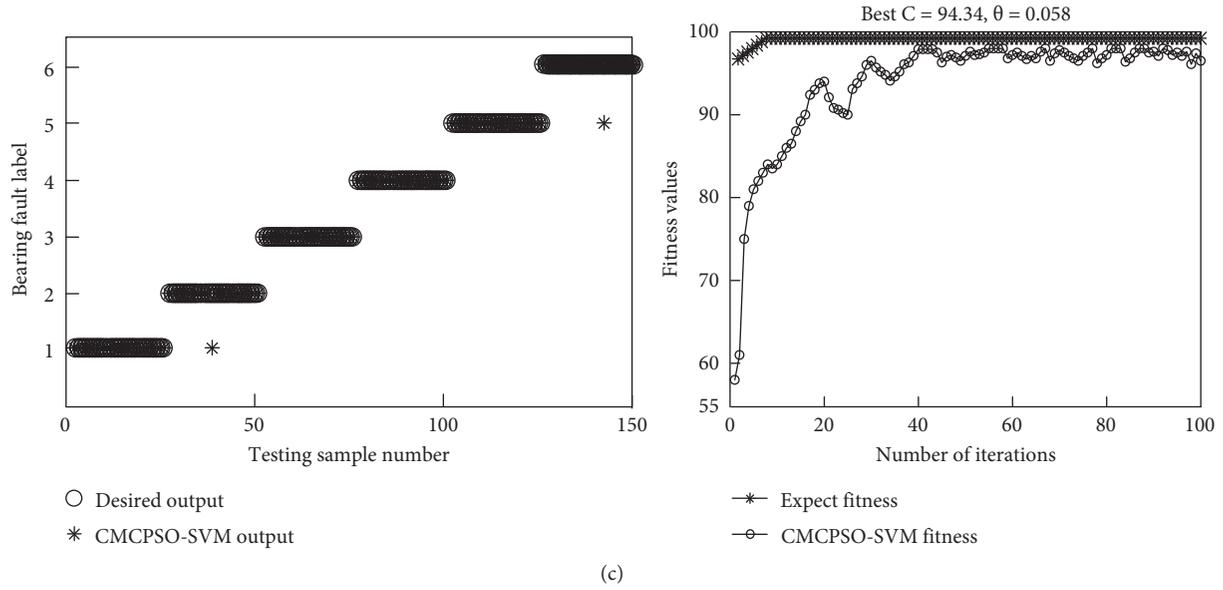


FIGURE 7: The classification results and the change law of fitness value. (a)~(c) are the classification results and the change law of fitness value of PSO-SVM, CPSO-SVM, and CMCP SO-SVM classification models, respectively.

TABLE 4: Classification accuracy of the three diagnosis models.

Fault type	Bearing fault label	Number of training samples	Number of test samples	Classifier type		
				PSO-SVM (%)	CPSO-SVM (%)	CMCP SO-SVM (%)
BR	1	25	25	88	100	100
LBR	2	25	25	92	88	96
OR	3	25	25	92	88	100
LOR	4	25	25	84	92	100
IR	5	25	25	80	80	100
LIR	6	25	25	92	96	96
Total/average	-	150	150	88	90.51	98.67

TABLE 5: Fault diagnosis accuracy and time under different iteration number working conditions.

Fault type	Number of iterations							
	50		100		150		200	
	Accuracy	Time	Accuracy	Time	Accuracy	Time	Accuracy (%)	Time (s)
BR	97.26	30.25	100	55.37	97.47	75.35	95.25	105.62
LBR	97.01	33.51	96	52.96	96.29	77.64	96.32	112.31
OR	96.59	37.65	100	52.13	97.02	79.26	96.54	99.86
LOR	98.37	38.01	100	50.64	97.64	70.66	95.91	102.45
IR	97.05	34.55	100	54.75	96.65	80.51	96.56	103.34
LIR	96.34	33.19	96	52.12	96.11	78.96	97.79	105.75
Average	97.10	34.53	98.67	52.99	96.86	77.06	96.40	104.89

accuracy is the highest, and the reliability is the best when the maximum iteration number is 100.

There are many types of classifiers, such as neural network classifier and Bayesian classifier. Therefore, experimental comparison between SVM classifiers is not enough to illustrate the superiority of CMCP SO-SVM classifier. To

further examine the effectiveness of the proposed CMCP SO-SVM classifier, the radial basis function neural network (RBFNN), extreme learning machine (ELM), K nearest neighbor algorithm (KNN), BP neural network algorithm (BPNN), and CMCP SO-SVM classifier are selected to compare the accuracy of fault diagnosis. Figure 8 shows the

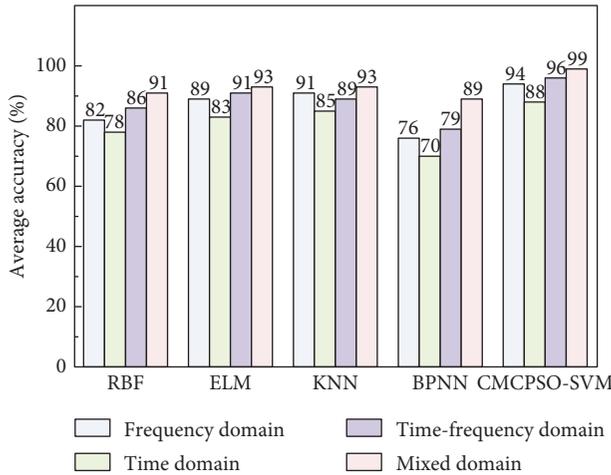


FIGURE 8: Average classification accuracy of different classification algorithms in different domains.

average classification performance. It can be seen from the diagnosis results that the accuracy of the CMCP SO-SVM algorithm is higher than other algorithms in the mixed domain and its average accuracy is about 15% higher than that of BPNN.

5. Conclusion

In order to solve the problem of sensitive features incomplete extraction in time-frequency domain, a fault diagnosis method based on crossover-mutation chaotic particle swarm optimization support vector machine in the mixed domain is proposed in this paper. The innovation points of this paper are as follows:

- (1) The sensitive characteristic values of the fault signals are selected in mixed domain, and the WMRMR algorithm is used to reduce the dimension of the sensitive feature set to obtain the optimal low-dimensional feature subset.
- (2) A new CMCP SO-SVM classifier is proposed. The new crossover-mutation strategy and an optimization disturbance are added to the CPSO algorithm to avoid the local optimization and improve the classification accuracy.
- (3) Different algorithms are used to classify bearing faults under different working conditions. Results showed that the average accuracy of CMCP SO-SVM was 98.67%, which was higher than that of the PSO-SVM and CPSO-SVM. In addition, the algorithm proposed in this paper, and the other four commonly algorithms were used for fault classification in different domains. From experimental analysis, it can be seen that the fault diagnosis accuracy of the CMCP SO-SVM algorithm in the mixed domain is the highest.
- (4) In this paper, CMCP SO-SVM was used to diagnose single faults of bearings under different working conditions, and the ideal results could be achieved.

And for the complicated working environment with variable bearing faults, the multiclass SVM classification strategy based on the CMCP SO-SVM might be a good choice to diagnose the bearing compound faults. As for future work, the combining multiple binary SVM classifiers could be utilized for bearing compound faults diagnosis.

Data Availability

All data are provided in full in the numerical simulation and discussion section of this article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

All authors contributed equally and all of them read and approved the final manuscript.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (grant nos. 51805299 and 51465009).

References

- [1] I. Attoui, B. Oudjani, N. Boutasseta et al., "Novel predictive features using a wrapper model for rolling bearing fault diagnosis based on vibration signal analysis," *International Journal of Advanced Manufacturing Technology*, vol. 106, no. 7-8, pp. 3409-3435, 2020.
- [2] S. R. Saufi, Z. A. B. Ahmad, M. S. Leong, and M. H. Lim, "Low-speed bearing fault diagnosis based on ArSSAE model using acoustic emission and vibration signals," *IEEE Access*, vol. 7, pp. 46885-46897, 2019.
- [3] T. D. Popescu and D. Aiordachioaie, "Fault detection of rolling element bearings using optimal segmentation of vibrating signals," *Mechanical Systems and Signal Processing*, vol. 116, pp. 370-391, 2019.
- [4] J. Singh, A. K. Darpe, and S. P. Singh, "Rolling element bearing fault diagnosis based on Over-Complete rational dilation wavelet transform and auto-correlation of analytic energy operator," *Mechanical Systems and Signal Processing*, vol. 100, pp. 662-693, 2018.
- [5] J. Grezmak, P. Wang, C. Sun, and R. X. Gao, "Explainable convolutional neural network for gearbox fault diagnosis," *Procedia CIRP*, vol. 80, pp. 476-481, 2019.
- [6] I. I. E. Amarouayache, M. N. Saadi, N. Guersi et al., "Bearing fault diagnostics using EEMD processing and convolutional neural network methods," *International Journal of Advanced Manufacturing Technology*, vol. 107, no. 9-10, pp. 4077-4095, 2020.
- [7] G. Jin, T. Zhu, M. W. Akram, Y. Jin, and C. Zhu, "An adaptive anti-noise neural network for bearing fault diagnosis under noise and varying load conditions," *IEEE Access*, vol. 8, pp. 74793-74807, 2020.
- [8] 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.

- [9] G. Zhang, H. Liu, P. Li et al., "Load prediction based on hybrid model of VMD-mRMR-BPNN-LSSVM," *Complexity*, vol. 2020, no. 4, pp. 1–20, 2020.
- [10] C. Demirkesen and H. Cherifi, "A comparison of multiclass SVM methods for real world natural scenes," in *Proceedings of the International Conference on Advanced Concepts for Intelligent Vision Systems*, Juan-les-Pins, France, October 2008.
- [11] R. K. Mandal and P. G. Kale, "Assessment of different multiclass SVM strategies for fault classification in a PV system," *Proceedings of the 7th International Conference on Advances in Energy Research*, Springer, Singapore, 2021.
- [12] L. Yang, L. Su, Y. Wang et al., "Metal roof fault diagnosis method based on RBF-SVM," *Complexity*, vol. 2020, Article ID 9645817, 2020.
- [13] S. Suyken and A. K. Johan, "Least squares support vector machines," *International Journal of Circuit Theory & Applications*, vol. 27, no. 6, pp. 605–615, 2002.
- [14] S. M. H. Bamakan, H. D. Wang, and A. Z. Ravasan, "Parameters optimization for nonparallel support vector machine by particle swarm optimization," *Promoting Business Analytics and Quantitative Management of Technology: 4th International Conference on Information Technology and Quantitative Management*, vol. 91, pp. 482–491, 2016.
- [15] S. P. Rajamohana and K. Umamaheswari, "Hybrid approach of improved binary particle swarm optimization and shuffled frog leaping for feature selection," *Computers & Electrical Engineering*, vol. 67, pp. 497–508, 2018.
- [16] E. Wang, C. Jia, G. Tong, P. Qu, X. Lan, and T. Pang, "Fault detection and isolation in GPS receiver autonomous integrity monitoring based on chaos particle swarm optimization-particle filter algorithm," *Advances in Space Research*, vol. 61, no. 5, pp. 1260–1272, 2018.
- [17] A. Tharwat and A. E. Hassanien, "Quantum-behaved particle swarm optimization for parameter optimization of support vector machine," *Journal of Classification*, vol. 36, 2019.
- [18] S. Q. Zhang, Y. T. Hu, A. Q. Jiang et al., "bearing fault diagnosis based on DTCWT and AWTFPSO-optimized SVM," *China Mechanical Engineering*, vol. 28, no. 03, pp. 327–333, 2017.
- [19] R. Jia, F. Ma, J. Dang, G. Liu, and H. Zhang, "Research on multidomain fault diagnosis of large wind turbines under complex environment," *Complexity*, vol. 2018, Article ID 2896850, 13 pages, 2018.
- [20] X. Xue and J. Zhou, "A hybrid fault diagnosis approach based on mixed-domain state features for rotating machinery," *Isa Transactions*, vol. 66, pp. 284–295, 2016.
- [21] C.-W. Fei, Y.-S. Choy, G.-C. Bai, and W.-Z. Tang, "Multi-feature entropy distance approach with vibration and acoustic emission signals for process feature recognition of rolling element bearing faults," *Structural Health Monitoring*, vol. 17, no. 2, pp. 156–168, 2018.
- [22] H. B. Yang, H. L. Zhang, and S. L. Liu, "A visual optimal selection method of time statistical feature and its application to rolling bearing fault diagnosis," *Machine Design and Manufacturing Engineering*, vol. 46, no. 8, pp. 123–126, 2017.
- [23] X. Li, W. Zhang, Q. Ding, and J.-Q. Sun, "Multi-Layer domain adaptation method for rolling bearing fault diagnosis," *Signal Processing*, vol. 157, pp. 180–197, 2019.
- [24] M. N. Yasir and B. H. Koh, "Data decomposition techniques with multi-scale permutation entropy calculations for bearing fault diagnosis," *Sensors*, vol. 18, no. 4, 2018.
- [25] I. R. Quinde, J. C. Sumba, L. E. Ochoa, A. J. V. Guevara, and R. Morales-Menendez, "bearing fault diagnosis based on optimal time-frequency representation method," *IFAC-PapersOnLine*, vol. 52, no. 11, pp. 194–199, 2019.
- [26] Y. G. Xu, K. Zhang, C. Y. Ma et al., "Optimized LMD method and its applications in rolling bearing fault diagnosis," *Measurement Science and Technology*, vol. 30, no. 12, 2019.
- [27] J. H. Liang, L. P. Wang, J. Wu et al., "Elimination of end effects in LMD based on LSTM network and applications for rolling bearing fault feature extraction," *Mathematical Problems in Engineering*, vol. 2020, Article ID 7293454, 16 pages, 2020.
- [28] T. Xu, Z. Yin, D. Cai, and D. Zheng, "Fault diagnosis for rotating machinery based on local mean decomposition morphology filtering and least square support vector machine," *Journal of Intelligent & Fuzzy Systems*, vol. 32, no. 3, pp. 2061–2070, 2017.
- [29] Y. Tian, Z. Wang, and C. Lu, "Self-adaptive bearing fault diagnosis based on permutation entropy and manifold-based dynamic time warping," *Mechanical Systems and Signal Processing*, vol. 114, pp. 658–673, 2019.
- [30] S. Chernbumroong, S. Cang, and H. Yu, "Maximum relevancy maximum complementary feature selection for multi-sensor activity recognition," *Expert Systems with Applications*, vol. 42, no. 1, pp. 573–583, 2015.
- [31] H. M. Dai, A. Q. Xu, W. F. Li et al., "Fault diagnosis of rolling bearings in mixed domain based on WMRMR," *Journal of Vibration and Shock*, vol. 34, no. 19, pp. 57–61, 2015.
- [32] S. Tyagi and S. K. Panigrahi, "An SVM-ANN hybrid classifier for diagnosis of gear fault," *Applied Artificial Intelligence*, vol. 31, no. 3, pp. 209–231, 2017.
- [33] D. Goyal, A. Choudhary, B. S. Pabla, and S. S. Dhama, "Support vector machines based non-contact fault diagnosis system for bearings," *Journal of Intelligent Manufacturing*, vol. 31, no. 5, pp. 1275–1289, 2020.
- [34] M. Z. Lv, X. M. Su, C. Z. Chen et al., "Application of SVM optimized by IPSO in rolling bearing fault diagnosis," *Machinery & Electronics*, vol. 37, no. 1, pp. 42–48, 2019.
- [35] G. S. Tan, Y. J. Shi, D. D. Liu et al., "fault diagnosis of transformers based on support vector machine with improved particle swarm optimization," *Journal of Kunming University of Science and Technology (Natural Science)*, vol. 44, no. 5, pp. 54–61, 2019.
- [36] D. Zhang and G. N. Li, "Fault diagnosis of S700K switch machine based on improved WNN analyses power curve," *Journal of Railway Science and Engineering*, vol. 15, no. 8, pp. 2123–2130, 2018.