

Research Article

Bearing Early Fault Diagnosis Based on an Improved Multiscale Permutation Entropy and SVM

Qunyan Jiang,¹ Juying Dai ,¹ Faming Shao,¹ Shengli Song,¹ and Fanjie Meng²

¹Department of Mechanical Engineering, College of Field Engineering and Army Engineering University, PLA, Nanjing 210007, China

²Department of Space Test and Launch, Noncommissioned Officer School, Space Engineering University, Beijing 102299, China

Correspondence should be addressed to Juying Dai; dinajy2001@aeu.edu.cn

Received 6 June 2022; Revised 21 July 2022; Accepted 26 July 2022; Published 2 September 2022

Academic Editor: Antonio Batista

Copyright © 2022 Qunyan Jiang 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.

Bearing fault is a process of gradual development and deepening. In the early stage of the fault, if it can be found out in time and taken reasonable prevention and elimination measures, we can avoid serious losses and safety accidents. Therefore, the feature extraction and analysis of early weak fault has important practical significance. In this paper, an improved multiscale permutation entropy (IMPE) method was proposed to overcome the shortcomings in the coarse-grained process. In order to solve the problem that only considering a single coarse-grained sequence under a certain scale may lead to the loss of feature information, this paper proposed to calculate the time series with equal overlapping segments, that was to consider all coarse-grained sequences under the same scale to reflect the feature information of fault signals more comprehensively. In order to solve the problem that feature extraction is not refined enough when using the first-order moment mean value calculation in traditional MPE calculation, a calculation method based on the skewness of the third-order moment was proposed. The calculation method is more sensitive to the complexity and fluctuation of signals and can better describe the feature details and extract the fault features effectively. IMPE was applied to feature extraction of early weak fault of rolling bearing and input into Support Vector Machines (SVMs) for faults classification. Aiming at SVM parameter optimization problem, an improved chaos firefly optimization algorithm was proposed. Experimental results show that the new method of early weak fault identification based on IMPE-SVM was effective in detecting rolling bearing faults with different severity.

1. Introduction

The rolling bearing supports the rotating shaft and the parts on the rotating shaft to maintain the normal working position and rotation accuracy of the rotating shaft. Its running state directly affects the overall performance of the rotating machinery. Most faults of rotating machinery are caused by rolling bearings, so it is of great significance for bearing fault diagnosis [1]. In the early stage of rolling bearing fault, the vibration signals collected by the sensor have no obvious characteristics and is often submerged in noise interference, which makes it very difficult for rolling bearing fault diagnosis [2]. Therefore, early fault diagnosis of rolling bearing has always been the focus and difficulty of researchers.

Early fault diagnosis methods can be divided into two categories: model-based fault diagnosis method and data

driven fault diagnosis method. Cf A [3] developed a Hierarchical Model Updating Strategy (HMUS) for Finite Element (FE) model updating with regard to uncorrelated modes. Tian et al. [4] established the dynamic model of inter-shaft bearing with localized defects with respect to time-varying displacement excitation to accurately describe the dynamic features of inter-shaft bearings with localized defect under operation. Fei et al. [5] proposed feature entropy distance method for the process character analysis and diagnosis of rolling bearing faults by the integration of four information entropies in time domain, frequency domain and time-frequency domain and two kinds of signals including vibration signals and acoustic emission.

How to extract the useful information from the signal is an urgent problem to be solved, which determines the accuracy and reliability of subsequent fault identification.

Permutation entropy (PE) is a nonlinear analysis method used to detect the randomness and dynamic variability of time series. It has the characteristics of high calculation efficiency, stable calculation value, strong anti-noise ability and suitable for online monitoring. It has been widely used in time series analysis and provides an effective tool for early fault weak feature extraction and analysis [6–8]. PE only carries out single-scale analysis of time series and ignores the useful fault information on other scales. Therefore, Aziz and ARIF proposed multiscale permutation to carry out multiscale analysis and processing of mechanical vibration signals to ensure the integrity of local and global information. Zhao et al. [9] presented a rolling bearing fault diagnosis approach by integrating wavelet packet decomposition (WPD) with multi-scale permutation entropy (MPE). Experimental study on a data set from the Case Western Reserve University bearing data center has shown that the presented approach can accurately identify faults in rolling bearings. Xiong et al. [10] proposed a method of complete ensemble empirical decomposition with adaptive noise (CEEMDAN) to solve the problem in virtue of its advantage of adaptive noise reduction. Combined with the multi-scale permutation entropy (MPE), which can reflect the random degree of time series in various scales and effectively detect the sudden dynamic change of the time series, an intelligent bearing fault recognition method is proposed by joint use of CEEMDAN, MPE, Fisher ratio and Gath-Geva (GG) clustering algorithm.

MPE is sensitive to the fluctuation of signals and can better describe feature details. However, there are some problems, such as the loss of feature information caused by considering only a single coarse-grained sequence at a certain scale. Aiming at the problem in MPE, Zheng et al. [11] proposed a new nonlinear dynamic method called generalized composite multiscale permutation entropy (GCMPE). GCMPE was compared with MPE by analyzing simulation data and also the influence of parameters on GCMPE calculation was studied. Huo et al. [12] proposed an improved entropy measure, termed Adaptive Multiscale Weighted Permutation Entropy (AMWPE). Then, a new rolling bearing fault diagnosis method is developed based on AMWPE and multi-class SVM. Li et al. [13] proposed a refined composite multiscale weighted permutation entropy (RCMWPE) method to efficiently characterize the operating state of bearings. The proposed method focuses on two aspects: the improved version reduces the dependence of entropy on the length of the original time series, and the error caused by considering the amplitude information is suppressed. Li et al. [14] proposed a method based on refined composite multiscale permutation entropy (RCMPE) and a support vector machine, the RCMPE algorithm is utilized to extract bearing feature information, and it is compared and analyzed with MPE and multiscale entropy (MSE). Through simulation and experimental verification of the signal, it is found that as the scale factor increases, RCMPE can retain more useful information. In this paper, an improved multiscale permutation entropy algorithm is proposed. By considering all coarse-grained time series at the same scale and using the third-order moment skewness calculation instead of

the first-order moment mean calculation, the characteristic information of the signal can be reflected more comprehensively and accurately to make up for the defect of MPE.

The essence of early fault diagnosis is fault identification and classification. Support vector machine (SVM) [15–17] is a good small sample set clustering algorithm. It also has excellent learning ability in the case of limited data, and has good generalization. It is widely used in fault diagnosis and other fields. However, in the process of fault identification, how to select appropriate parameters has become an important problem to be solved. Scholars have done a lot of work in algorithm optimization and parameter optimization. Keshtegar et al. [18] proposed two different Modified multi-extremum Response Surface basis Models (MRSM) for dynamic nonlinear responses of failure capacities for turbine blisk responses. The proposed MRSM is established using two regression processes including regressed the input variables by linear or exponential basis functions in first calibrating phase and regressed the second-order polynomial basis function using inputs data provided by first stage in second calibrating procedure. Lu et al. [19] proposed moving extremum surrogate modeling strategy (MESMS) in respect of multi-physics coupling with various dynamics/uncertainties to improve the dynamic reliability analysis of complex structures like turbine blisk. In this strategy, extremum thought is adopted to handle the dynamic process of input parameters and output response, and the importance sampling (IS) method is utilized to extract efficient samples and improve the efficiency of dynamic reliability estimation. Lu et al. [20] developed a surrogate model method, namely modified Kriging-based moving extremum framework (MKMEF), absorbing extremum thought, moving least squares (MLS) technique, Kriging model and collaborative evolution genetic algorithm (CEGA). For this proposed MKMEF, the extremum thought is used to transform dynamic output response into extremum values within a time domain, and the MLS method is applied to obtain efficient samples to derive Kriging model. Fei et al. [21] developed distributed collaborative improved support-vector regression (DCISR) method and multilevel nested model to effectively perform the reliability-based design optimization (REDO) of the assembly relationship to improve the assembly relationship design. In the DCISR method, the improved support-vector regression (ISR) is developed as the basis function of the DCISR model for reliability analysis, by adopting multi-population genetic algorithm (MPGA) to find the optimal model parameters. Zheng et al. [22] introduce an improved PSO-SVM algorithm based on distance pairing sorting support vector preselecting. They proposed that the training data set used distance pairing sorting support vector preselecting to obtain a support vector candidate set, the PSO parameter optimization process was put on the support vector candidate set. In this paper, a chaotic firefly optimization algorithm (CFOA) is introduced to optimize SVM parameters, and a new algorithm based on chaotic firefly optimization support vector machine (CFOA-SVM) is proposed.

Based on the above analysis, this paper uses the improved multiscale permutation entropy to extract the signal

features, and uses the SVM with the optimized parameters of the improved chaotic firefly optimization algorithm to classify the faults. A new bearing early fault identification method based on IMPE-SVM is proposed, and the proposed method is applied to the fault diagnosis experiments of rolling bearings with different severity. The results show that this method can effectively and accurately identify various types of faults.

The rest of this paper is organized as follows: In Section 2, we describe our proposed method in detail. The experimental results are given in Section 3, while Section 4 concludes this paper with remarks.

2. Methods

2.1. Overview of Our Method. The fault diagnosis method proposed in this paper mainly includes two parts: feature extraction and fault recognition. Firstly, the IMPE method was used to extract the multi-scale permutation entropy features of all signal samples, and then the permutation entropy feature vector was normalized, and the feature vector was divided into the training set and testing set. Then the SVM was trained, and the parameters were optimized by CFOA to obtain the optimal diagnosis model. Finally, the test set was input into the trained SVM for fault identification and classification. In order to verify the performance of the proposed fault diagnosis method in mechanical early weak fault detection and identification, the bearing fault identification experiments with different degrees of damage were carried out on the bearing open data set of Case Western Reserve University. The flow chart of fault diagnosis based on IMPE-SVM is shown in Figure 1.

2.2. Improved Multiscale Permutation Entropy. According to the principle of MPE, the coarse-grained process can be regarded as the average of the original time series in a length of window. There are some shortcomings in this process: firstly, different scale factors easily lead to instability of calculation results. In MPE calculation, the length of each coarse-grained time series is equal to $[N/\lambda]$, when the scale factor is too large, the length of coarse-grained data will be shortened with the increase of scale factor. Since calculation of permutation entropy depends on the data length, in order to reduce the MPE deviation caused by data length shortening, the average value of all coarse-grained time series under the same scale factor is taken as the final permutation entropy. However, in the process of down-sampling the data by a scale factor, the dynamic change of the original signal is weakened to a certain extent, and the estimated entropy value is less than the expected value, which may lead to inaccurate and unreliable results [23]. Secondly, non-overlapping computation based on continuous point segmentation may lead to incomplete feature extraction. In the coarse-grained process, the time series are divided into equal non overlapping segments. Taking scale 3 as an example, each sequence (x_1, x_2, x_3) , (x_4, x_5, x_6) , etc. are considered separately, then the average value of all the data points in each segment is calculated. In this calculation process, the

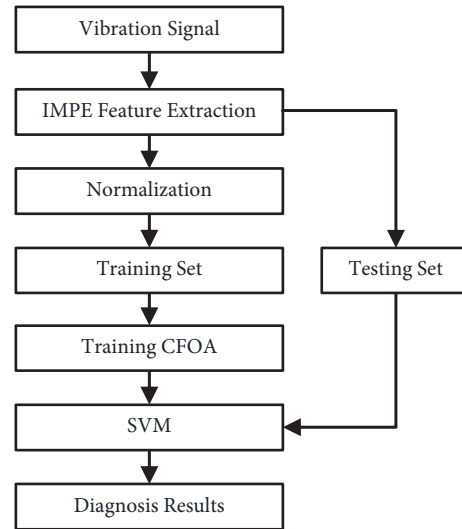


FIGURE 1: Overview of our method.

sequences (x_2, x_3, x_4) , (x_3, x_4, x_5) , etc. are not considered. This may result in the loss of some potentially useful information in the signal. Thirdly, the coarse-grained sequence based on mean value calculation can reflect the centralized trend of data, but it is difficult to reflect the general situation of a group of data.

2.2.1. Calculation Principle. In order to overcome the above shortcomings, an improved multi-scale permutation entropy (IMPE) feature extraction method was proposed. This method mainly improves the permutation entropy method based on the data skewness to obtain the mean value and calculate the overlapped data points, to better characterize the overall trend of the data, and to improve the reliability of permutation entropy in the random measurement and dynamic change detection of time series.

The skewness [24] is a characteristic number representing the degree of asymmetry of the probability distribution density curve with respect to the average value, the skewness of the samples is the third order standard moment of the samples. The formula is $S_k = E[(X - \mu)^3 / (\sigma)^3]$, where μ is the mean value, σ is the standard deviation, and E is the mean operation. The larger the absolute value of skewness is, the less ideal the performance of the mean value is, otherwise, the more reliable the mean value is [25]. Skewness is a dimensionless parameter, which is very sensitive to the change of signal amplitude and is not affected by working conditions, so it is especially suitable for early fault detection. Therefore, this paper proposes an improved multiscale permutation entropy based on skewness calculation to measure the difference of permutation entropy probability distribution under different states, to reveal the weak and complex change process of mechanical fault occurrence and development.

This paper has improved the traditional MPE non overlapping computing coarse-grained series, and considers multiple time series $z_{k,f}^\lambda = 1/\lambda \sum_{i=0}^{\lambda-1} x_{i+k+\lambda(f-1)}$ in each scale to form $y_k^\lambda = \{z_{k,1}^\lambda, z_{k,2}^\lambda, \dots\}$, in this way, for each scale factor λ ,

λ different time series were considered, which was different from the traditional MPE which only considers a single coarse-grained sequence and a single coarse-grained sequence at the same scale. To sum up, the improved multiscale permutation entropy algorithm is proposed as follows:

Step 1: for the scale factor λ , the original time series is divided into λ different coarse-grained sequences $y_{k,f}^\lambda$, $k = 1, 2, \dots, \lambda$; $f = 1, 2, \dots, [N/\lambda]$, which are obtained based on the following formula:

$$y_{k,f}^\lambda = \frac{1}{\lambda} \sum_{i=(f-1)\lambda+k}^{f\lambda+k-1} \left(\frac{x_i - \bar{x}_i}{x_i^d} \right)^3, \quad (1)$$

where $\bar{x}_i = 1/\lambda \sum_{h=1}^{\lambda-1} x_{i+h}$, $1 \leq f \leq N/\lambda$, $2 \leq k \leq \lambda$, x_i^d is the standard deviation.

Step 2: PE values of all coarse-grained sequences y_f^λ under each scale factor λ are calculated.

Step 3: calculate the mean value IMPE of all permutation entropy under each scale factor, which is the permutation entropy of the original time series under the scale factor λ .

$$\text{IMPE}(x(i), m, t, \lambda) = \frac{1}{\lambda} \sum \text{PE}(y_f^{(\lambda)}, m, t). \quad (2)$$

Step 4: $\lambda = \lambda + 1$, repeat steps 2 and 3 until $\lambda = \lambda_m$, where λ_m is the largest scale factor, then all PE values are plotted as a function of the scale factor λ , which is called IMPE.

Figure 2 shows the algorithm flow of IMPE. It can be seen that λ different coarse-grained sequences related to the scale factor λ are considered in the IMPE algorithm, that is, all the coarse-grained sequence information under the same scale is considered. In the traditional MPE algorithm, only one group of coarse-grained sequences $y_1^{(\lambda)}$ is considered in the calculation of each scale. Therefore, compared with MPE, IMPE can reflect the signal fault characteristic information more comprehensively. Moreover, the calculation of PE itself is based on the comparison of adjacent data points without considering the size of the data itself. Therefore, in formula (2), the mean value of the first-order moment, the unbiased square error of the second-order moment and the skewness of the third-order moment are calculated. In theory, it can reflect the characteristics of the whole coarse-grained sequence data, but compared with the mean value of the first-order moment and the unbiased square error of the second-order moment, the skewness calculated based on the third-order moment is more sensitive to the change of the data. Therefore, the multi-scale permutation entropy based on skewness is more sensitive to the complexity and volatility of signals, and can better describe the feature details.

2.2.2. Parameter Analysis and Optimization. According to the calculation principle of permutation entropy, the delay time t , embedding dimension m , signal length N and scale factor λ are the four main parameters affecting permutation

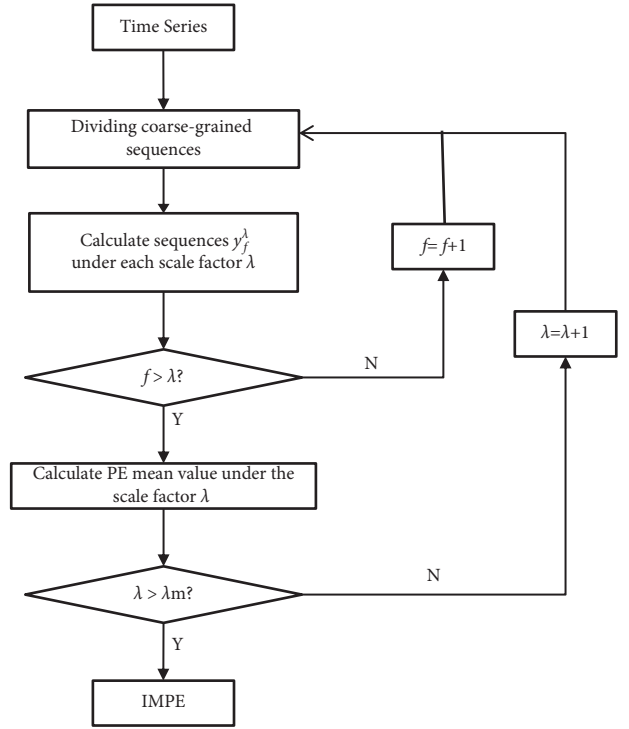


FIGURE 2: IMPE algorithm flow.

entropy algorithm. The permutation entropy calculation is highly dependent on the selection of embedding dimension m . When m is too large, the phase space calculation is complex and the computing time increased, and when m is too small, the reconstructed information is not enough to extract and detect the mutation signal. According to Bandt and Pompe [26], the range of m is 3 to 7. According to Matilla-García [27], the signal length should satisfy $N \geq 5m!$ when to obtain reliable statistical data. The scaling factor λ determines the permutation entropy characteristics of signals in the corresponding scale. Generally speaking, there is no unified standard for the maximum scale factor. As proposed by Zheng et al. [28], the maximum scale factor is usually greater than 10.

In order to research the influence of parameter selection on the calculation results of IMPE, Gaussian white noise was considered in the experiments. Although white noise is random in the time domain, its power spectral density is parallel to the horizontal axis and independent of frequency. In order to research the influence of embedding dimension m on MPE and IMPE calculation results, Gaussian white noise with data length of 2048 was selected, embedding dimensions were 3, 4, 5, 6, 7, delay factor was 1, and scale factor was 20. The results are shown in Figure 3. It can be seen from the figure that the permutation entropy of Gaussian white noise decreases monotonically with the increase of scale factor λ . MPE and IMPE in the same embedded dimension have similar intervals. However, in the whole scale factor range, the fluctuation range of MPE is larger than that of IMPE, that is, IMPE is more stable. Secondly, when the embedding dimension m is small, such as 3 or 4, the permutation entropy is larger. Moreover, with

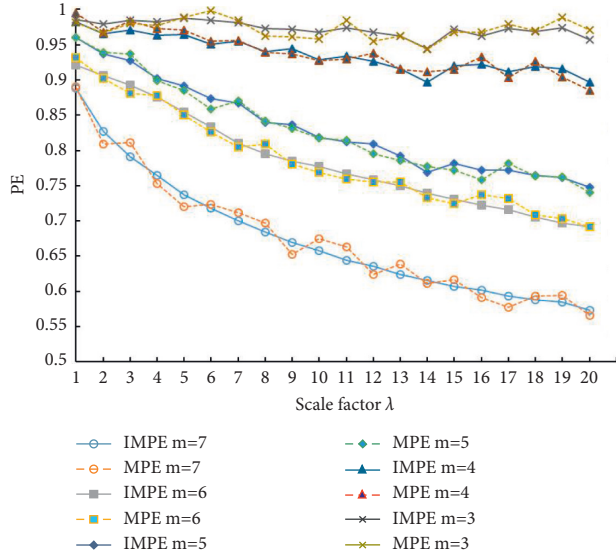


FIGURE 3: Permutation entropy of white Gaussian noise with different embedding dimensions.

the increase of scale factor, the change of PE value is not obvious, which cannot reflect the advantages of multi-scale analysis. However, if it is too large, such as 6 or 7, the permutation entropy calculation process makes the coarse-grained sequence homogenized, and it is easy to ignore the details of the time series. Therefore, the embedding dimension $m = 5$ was determined by experiments. If the signal length n is too large, it will affect the calculation efficiency, but if it is too small, the condition of $N \geq 5m!$ cannot be satisfied. Considering these constraints, the signal length 2048 is enough to obtain reliable and stable permutation entropy. When $m = 5$, the relationship between PE and scale factor is shown in Figure 4. It can be seen the influence of delay factor on time series is very small when the embedding dimension is fixed. Therefore, the delay factor $t = 1$ is determined in this paper. Finally, the scale factor λ is set to 20.

2.3. Support Vector Machine

2.3.1. Algorithm Principle. Support vector machine (SVM) is a better small sample set clustering supervised learning algorithm, which classifies data based on statistical learning and structural risk minimization principle, and minimizes the upper bound generalization error [29]. SVM overcomes the problems of too large network scale, over learning and poor generalization ability of general classification methods. It also has excellent learning ability in the case of limited data, and has achieved good results in fault diagnosis and other fields. Traditional SVM is an algorithm based on binary classification problem, which cannot be directly deal with multi classification problem [30]. Scholars use the calculation process of standard SVM to construct multiple decision boundaries in order to realize multi classification of samples. The common methods are “one against all” and “one against one” [31], This paper adopts one-to-one SVM.

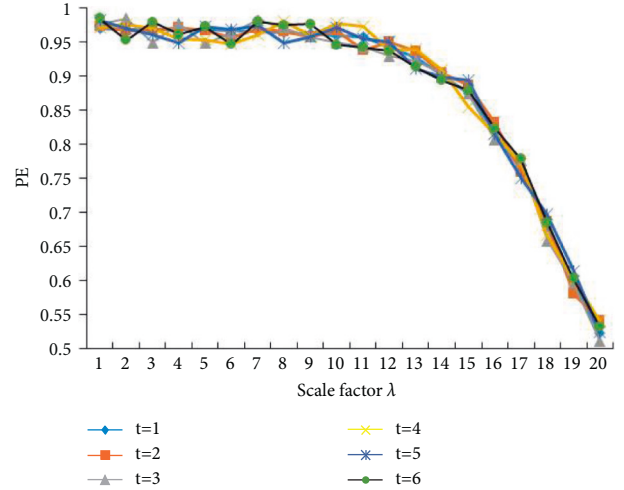


FIGURE 4: Permutation entropy of white Gaussian noise with different time delays.

For the obtained training sample set (x_i, y_i) , ($i = 1, \dots, n, x \in R^n, y \in R$), where x_i is the input space vector of the data sample and y_i is the target value. SVM classification model is as follows:

$$f(x) = W^T \beta(x) + b, \quad (3)$$

$$M = \frac{1}{2} \|W^T\|^2 + \delta F,$$

where $\beta(x)$ represents the high-dimensional space feature of mapping input space vector x , and W and b are parameters to be determined. $1/2 \|W^T\|^2$ is the regular term, δF is the empirical error, δ is the penalty factor, F is the loss function, this is equivalent to the approximate accuracy of the training data points, and the parameters W and b can be estimated by minimizing the regularization risk function [32]. The classification problem is transformed into a convex optimization problem, which can be expressed as follows:

$$\min \frac{1}{2} \|W^T\|^2 + \delta \sum_{i=1}^n \gamma_i^2, \quad (4)$$

$$\text{s.t. } y_i (W^T x_i + b) \geq 1 - \gamma_i.$$

In order to obtain the maximum generalized optimal classification, the problem of maximum classification interval and minimum misclassification samples must be considered. Lagrange multipliers and optimality constraints are introduced to solve the problem. At the same time, considering that most of the practical classification problems are nonlinear problems, the kernel function is introduced to transform it into a linear separable problem in high-dimensional space. Finally, the classification decision function is obtained as follows:

$$f(x) = \sum_{i=1}^n a_i K(x_i, x_p) + b, \quad (5)$$

where γ_i is the error, α_i is the Lagrange multiplier and $K(x_i, x_p)$ is the kernel function. The kernel function [33] is

introduced into SVM to solve the nonlinear problem effectively and reduce the calculation cost. The common kernel functions are linear kernel function, polynomial kernel function, Laplace kernel function, sigmoid kernel function and Gaussian radial basis function. At present, there is no good theoretical method to guide how to select the kernel function, which needs to be selected according to the specific problems and the experience of researchers. According to the existing researches, the selection of different kernel functions has little influence on the results, that is, SVM is not very sensitive to kernel function [34]. In this paper, the Gaussian radial basis function is selected and its expression is $K(x_i, x_p) = \exp(-\|x_i - x_p\|^2/2\sigma^2)$, where σ is the kernel parameter.

2.3.2. Improvement of SVM and Parameter Optimization.

Penalty factor δ and kernel parameter σ are the main parameters affecting the classification accuracy of SVM. The value of penalty factor δ balances the empirical risk and structural risk [35]. If δ is too small, it means that the limitation of empirical error is small, and the complexity of the model is reduced, but the empirical risk increases and the phenomenon of under fitting occurs. If δ is too large, the empirical risk will be reduced, but the structure will become complex and over fitting phenomenon will appear easily. The kernel parameter σ will affect the mapping function [36]. A suitable σ can make the Gaussian kernel as small as possible, thus reducing the dimension of high-dimensional space samples and reducing the empirical error of classifiers, which is conducive to accurate classification. If σ is too small, the distance between all mapped points is almost equal, that is, there is no clustering phenomenon, if σ is too large, two different points will become the same point in high-dimensional space after mapping, that is to say, all samples will be classified into the same class and cannot be distinguished. Therefore, in order to get a good classification effect, it is necessary to take a certain method to optimize the parameters.

In this paper, the firefly algorithm [37] was used to optimize the SVM parameters. In the traditional firefly algorithm, the optimization results are easy to oscillate repeatedly at the local or global extreme points [38], and the convergence is slow, which reduces the optimization accuracy. In order to improve the optimization ability and population diversity of the algorithm, the chaotic firefly optimization algorithm (CFOA) was used, this paper mainly optimizes the initial sequence chaotic mapping method to improve the diversity of the population and the ergodicity of the optimization, and improve the ability to fall into local minima. At present, logistic map is often used to generate chaotic sequence, and its optimization speed is easily affected by uneven traversal. In this paper, tent chaotic map (TCM) was used to generate initial sequence. Tent map [39] is a kind of nonlinear dynamic discrete chaotic mapping method which is widely used at present. It has uniform distribution function and good correlation. It has simple iteration and fast search speed. It can generate random and ergodic initial sequence when used in Firefly sequence initialization.

$$x_{i+1} = a - (1 + a)|x_i|, \quad (6)$$

where $a \in [0, 1]$, $x_i \in [0, 1]$.

After generating chaos mapping, the chaos space is mapped to the optimal solution space by

$$u_{d,b} = (U_b - D_b) \cdot x_{i+1} + D_b, \quad (7)$$

where U_b and D_b are the upper and lower limits of d dimension variables respectively.

According to the set chaotic mapping function, the chaotic sequence is generated from formula (2), and then the chaotic sequence is mapped to the optimization space by using formula (3) to search the optimal solution. The process of improved chaotic firefly optimization SVM parameter penalty factor δ and Gaussian kernel σ is shown in following [40]:

- Step 1: initialize parameters
- Step 2: generate initial firefly sequences using tent chaotic map
- Step 3: the parameters σ and δ that need to be optimized are taken as the initial positions of fireflies and mapped to the search space
- Step 4: the light intensity of firefly was taken as the objective function value and calculated
- Step 5: calculate the distance to attract fireflies
- Step 6: calculate the attractiveness of fireflies
- Step 7: update adaptive step size
- Step 8: if the stop condition is met, go to the next step; otherwise return to Step 4
- Step 9: complete the optimization and output the optimal parameters σ and δ

3. Experimental Results

3.1. Open Bearing Fault Data Set. The experimental data of bearing damage degree evaluation are from the bearing data of Case Western Reserve University [41]. The test bench is shown in Figure 5, it consists of a 2 HP motor, a torque sensor, a dynamometer and control electronics device. The fault of the drive end bearing and the fan end bearing are simulated on the test bench. The drive end bearing model is SKF6205, and the fan end bearing model is SKF6203, the bearing supports the motor shaft, the bearing pedestal at the fan end and the drive end of the motor. An acceleration sensor is placed above them to collect the vibration signal of the fault bearing. By using EDM technology, the single point local damage was processed manually. The fault points were set on the rolling body, outer ring raceway and inner ring raceway respectively, with diameters of 0.18 mm, 0.36 mm and 0.54 mm, these fault points simulate three different degrees of damage of rolling bearing, which are minor stage of failure, macro stage of failure and final stage of failure. The vibration signal is collected under four different load conditions, and the sampling frequency is 12 kHz. In this paper, the drive end bearing SKF6205 is selected for research. The specific bearing information is shown in Table 1. As the basic

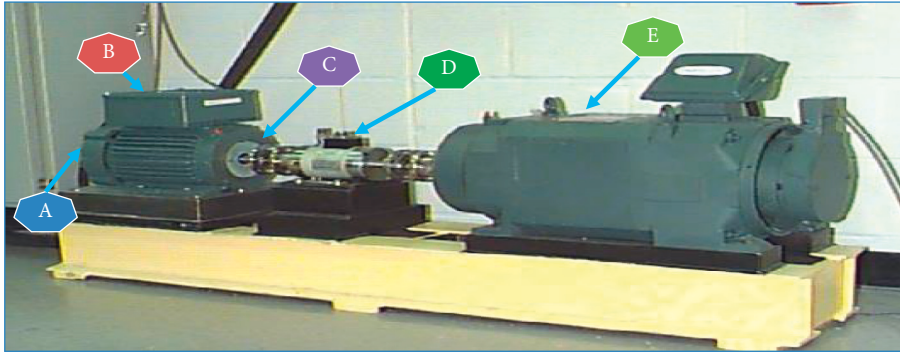


FIGURE 5: Bearing test ring. A: fan end bearing; B: electronic motor; C: drive end bearing; D: torque transducer; E: dynamometer.

TABLE 1: Drive end bearing parameters (mm).

Type	External diameter	Internal diameter	Thickness	Diameter of rolling element	Pitch diameter	Number of rolling
6205-2RS JEM SKF	52	25	15	7.94	39	9

element of rolling bearing, the fatigue pitting and spalling damage is a common failure mode of rolling bearing. Therefore, this paper takes the rolling element as the research object to carry out the fault diagnosis research of bearing with different damage degree. The motor speed is 1750 rpm and the load is 1492 W.

3.2. Definition of Bearing Failure. Bearing failure, especially fatigue pitting and wear, is usually a gradual development process from weak to severe. Figure 6 shows the fault development process of bearing, which can be roughly divided into four stages, namely, four types: the initial stage of failure, the stage of minor failure, the stage of macro failure and the stage of final failure. It can be seen that the bearing can work normally about 80% of the time in the basic rated life cycle. Once the fault occurs, the deterioration process of the fault is not linear development, but changes according to the exponential law. Therefore, it is very important to research the development process of bearing fault and grasp the change law of signal in the process of fault development, to carry out the assessment of bearing health status, to predict the remaining life, to formulate necessary maintenance and repair measures, to ensure the normal operation of machinery and reduce the occurrence of accidents.

The research results show that with the deepening of bearing fault, the physical quantities monitored will change. Taking vibration signal as an example, the characteristic change process is as follows:

The first stage is the initial stage of bearing failure. Currently, the total amount and frequency spectrum of noise and vibration velocity are normal, but the peak energy and spectrum have some symptoms, which reflect the initial abnormal of bearing. The characteristic frequency of bearing fault occurs in the range of 20 Hz–60 kHz in ultrasonic section.

The second stage is minor stage of failure, which is mainly the minor pitting or damage on the surface of parts.

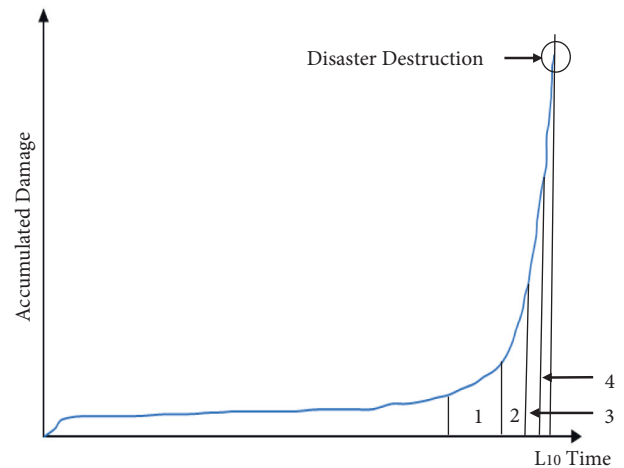


FIGURE 6: Four stages of rolling bearing fault development (L_{10} is the basic rated life of the bearing): (1) initial stage of failure, (2) minor stage of failure, (3) macro stage of failure, and (4) final stage of failure.

At this time, the noise increases slightly, the signal fault characteristics are weak, and the fault information is easy to be submerged by noise. The total vibration velocity and frequency spectrum change is not prominent, but the total peak energy has a greater increase, and the frequency spectrum is more prominent. The characteristic frequency of bearing fault occurs in the range of 500 Hz–2 kHz.

The third is the macro stage of failure. The slight fault gradually expands and becomes more obvious. The noise can be heard and the total vibration speed increases greatly. The fault characteristic frequency, harmonic and sideband of the bearing on the frequency spectrum are clearly visible. The peak energy becomes larger and the frequency spectrum is more obvious than the second stage. The bearing failure frequency occurs in the range of about 0-1 kHz. Generally, it is recommended to replace the bearing at the later stage of

the third stage. At this time, the rolling bearing failure characteristics such as pitting, spalling and wear can be observed by naked eyes.

The fourth stage is the final stage of the fault. The damage degree of the bearing is further deteriorated, large area spalling, the noise intensity is obvious, the total vibration velocity and amplitude are significantly increased, and the bearing fault characteristic frequency on the vibration speed spectrum is replaced by larger random high-frequency noise, the total peak energy increased rapidly and showed unstable changes. This will lead to functional failure, and the bearing must not be allowed to run in the final stage of fault development, otherwise catastrophic damage may occur.

Our research shows that if the whole service life of the rolling bearing is 1, then the bearing will work normally in the first 80% of its life cycle from the moment the bearing is put into service. The later life cycle time corresponds to the fault development stage of rolling bearing, and its remaining life is 10%~20% in the first stage, the second stage is 5%~10%, the third stage is 1%~5%, and the fourth stage is about 1 hour or 1%. Therefore, in order to avoid the catastrophic failure of bearing development to the final failure stage, it is necessary to monitor the working state of the bearing, in order to ensure the safe and reliable operation of the machinery, the fault feature information is extracted to detect the small fault and initial anomaly early, and necessary maintenance and repair measures are taken.

3.3. Experimental Analysis of Signal Spectrum Characteristics.

According to the calculation, the rotation frequency of the shaft is 29.2 Hz, and the fault characteristic frequency of the rolling element is $f_o = 137.6$ Hz according to the geometric parameters of the bearing and the rotation frequency of the shaft. Figure 7 shows the time domain waveform and spectrum of rolling element signal in normal state and three different damage degrees. Figure 7(a) shows the spectrum of normal signal. It can be seen from the figure that the main frequency of signal under normal condition of bearing is concentrated in low frequency band, the main characteristic frequencies are 87.89 Hz, 1037 Hz and 2104 Hz, which are 3 times, 37 times and 71 times of the shaft rotation frequency, respectively. Figure 7(b) shows the spectrum of slight damage of rolling element, and compares the spectrum of normal state, the bearing develops from normal to slight damage, and the frequency component increases and the frequency value begins to increase. The main frequency components are 3000–3500 Hz, of which 1400 Hz, 3176 Hz and 3434 Hz are the multiple frequency of rolling element characteristic frequency. Figure 7(c) shows the spectrum of moderate damage signal of rolling element. From slight damage to moderate damage, the main frequency components basically remain unchanged, the amplitude of a small part of low-frequency and low amplitude clutter increased, and the amplitude of high-frequency components decreased slightly. The main frequency components of 1400 Hz, 2730 Hz and 3287 Hz were the multiple frequency of rolling element characteristic frequency. Figure 7(d) shows the spectrum of rolling serious damage signal. From moderate

damage to damage injury, the clutter frequency of high-frequency part of main frequency components is less, and energy is concentrated on several prominent high-frequency components, among the main frequency components, 1406 Hz, 3029 Hz and 3445 Hz are the multiple frequency of rolling element characteristic frequency.

3.4. IMPE Feature Extraction. Based on the optimized parameter $m = 5$, $t = 1$, $\lambda = 20$, $N = 2048$ mentioned above, the multi-scale arrangement entropy of all samples in four states of bearing is calculated. Figure 8 shows the mean value of permutation entropy for different health states of bearings. It can be seen from the figure that when the scale factor is 1, the PE value of bearing normal state vibration signal is less than that of other three fault states. The reason may be that the dynamic characteristics of the vibration signal will change and the PE value and increase accordingly when the rolling bearing has a local fault. Therefore, permutation entropy method is suitable for rolling bearing fault monitoring, and is an effective means for health monitoring and fault detection. For example, in the above cases, a PE value of 0.75 can effectively distinguish the fault state from the normal bearing. It is also found that although the single scale permutation entropy can detect the normal and fault bearings, it cannot judge the severity of the fault.

In addition, it can be seen from Figure 8 that the multi-scale arrangement entropy values of three bearing vibration signals with different severity degrees have significant differences under different scale factors. In other words, the permutation entropy of the four states is arranged according to a certain size relationship under a certain scale, while the PE value size relationship of the four states is no longer tenable in another scale. This shows that the permutation entropy of different states has its own fluctuation range on multi-scale, and the fluctuation range of permutation entropy of different states has certain overlap and intersection. When the scale factor is 1, the PE value of each state is more obvious, but with the increase of scale factor, the PE value of each state is closer to the original value, which shows that increasing the scale factor blindly cannot improve the ability of IMPE to distinguish different states. This is because the overall influence of fault characteristics on permutation entropy is averaged by coarse granulation process. The larger the scale factor is, the more obvious the average effect is, and the smaller the difference of permutation entropy of fault characteristics is, therefore, in order to reduce the dimension of feature vector and reduce the computational complexity, the permutation entropy can be calculated by selecting the first 10 scales. To sum up, the improved multi-scale permutation entropy method can effectively reflect the fault characteristics of rolling bearing and reflect the difference of signal complexity under different states. However, as the previous analysis, sometimes the usage of coarse-grained sequences may reduce the usefulness of permutation entropy method, so it is not suitable to directly use multi-scale permutation entropy features for fault identification and classification.

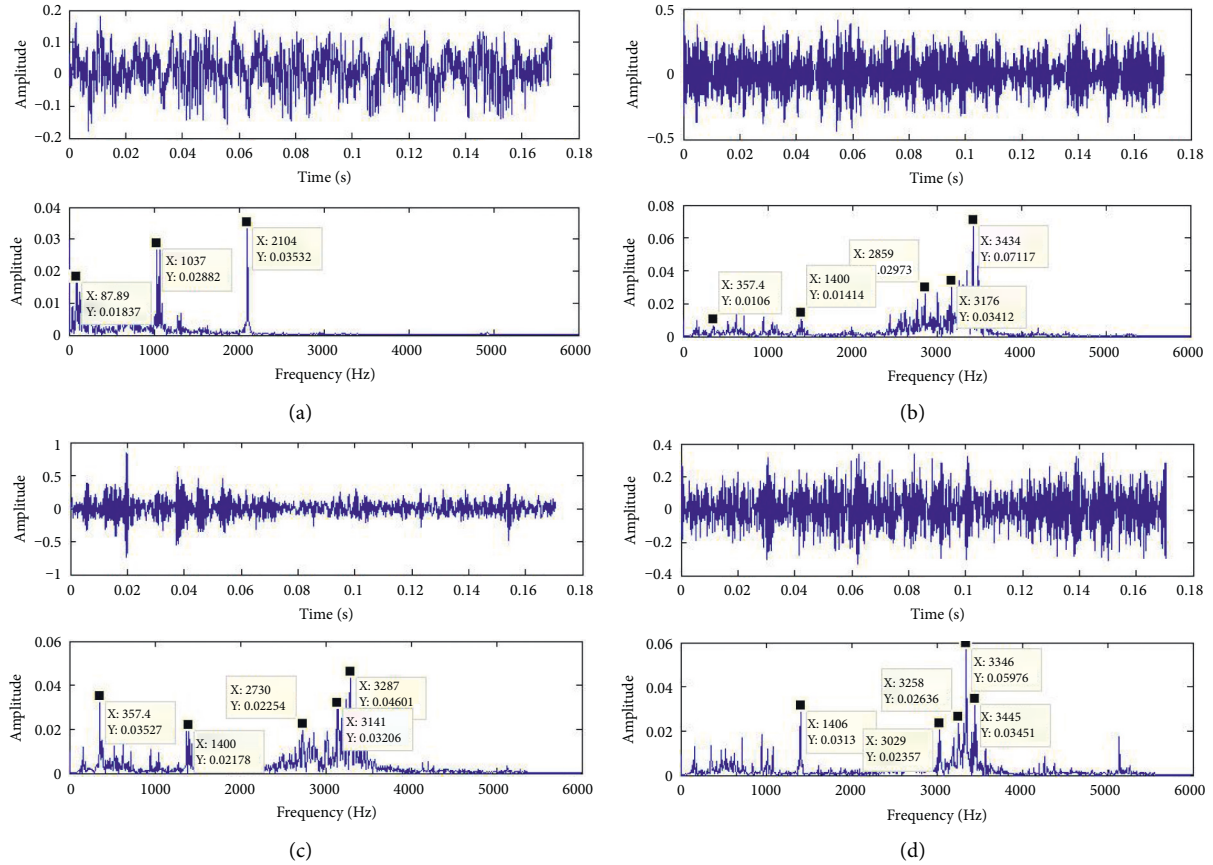


FIGURE 7: Time domain waveform and spectrum of rolling element with different damage degree. (a) Normal state. (b) Slight damage of rolling element. (c) Moderate damage of rolling element. (d) Serious damage of rolling element.

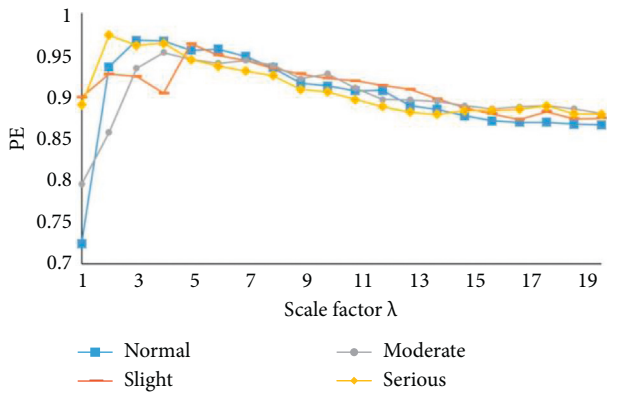


FIGURE 8: Permutation entropy of bearing vibration signals in different health states.

3.5. Diagnosis Results and Analysis. Four kinds of samples, including normal state, slight damage, moderate damage and serious damage, were selected for experimental study. The status category labels are 1,2,3,4, and 30 samples are selected for each state, with a total of 120 samples. There are 40 samples in 10 groups as training set, and 80 samples in 20 groups as test set.

The extracted IMPE features are input into SVM classifier for training, and the parameters of SVM are optimized by

CFOA. CFOA initial value are set by: the maximum population size is 30, the number of iterations is 200, the light absorption coefficient is 1, the step factor is 0.6, and the maximum attraction is 1. Fault detection accuracy is used as fitness function, CFOA is used to optimize SVM parameters. The optimization interval of penalty factor is 0.001–1.00, and that of kernel function is 1.00–100.00. After optimization, the optimal penalty factor $\delta = 10.8429$ and the parameter of Gaussian kernel function $\sigma = 0.4853$. The training data is used as the input training SVM, and the test data is input into the trained SVM for state classification. The final diagnosis results are shown in Figure 9. 78 of the 80 test samples were identified accurately, the correct rate of diagnosis was 97.5%, and the classification accuracy was high. All the normal samples are correctly identified. Among the three bearing samples with different damage degrees, the serious damage status is correctly identified, which is of great significance for timely detection of serious faults and necessary maintenance and repair. One slight damage sample was identified as a serious damage sample, and one moderate damage sample was identified as a slight damage sample. The experimental results show that the IMPE-SVM method proposed in this paper can accurately judge the bearing damage with different severity, and the diagnosis effect is good.

In order to verify the advantages of the proposed method, Wavelet, PE, MPE and IMPE features of different

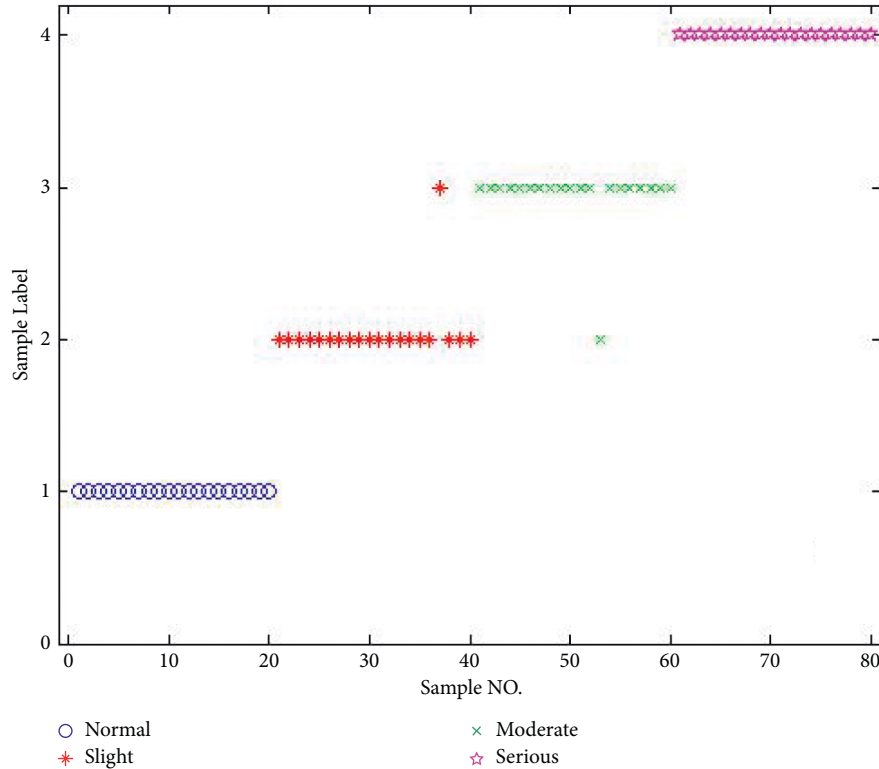


FIGURE 9: Fault classification results of bearing with different damage degree.

TABLE 2: Diagnostic results of different feature extraction methods.

Feature extraction	1	2	3	4	Total accuracy (%)
Wavelet [40]	78	79	75	76	77
Wavelet [41]	83	81	79	84	81.75
PE	95	85	90	95	91.25
MPE	100	95	90	95	95
IMPE	100	95	95	100	97.5

state samples were extracted for testing. The recognition and prediction results of different states are shown in Table 2. Firstly, through the verification of wavelet feature extraction, we use the method of literature [40, 41] to extract wavelet or modified wavelet features. From the results, it can be seen that the accuracy of the method based on wavelet is low, while the overall performance based on permutation entropy is excellent. It can be seen from the table that the classification accuracy of bearing in different health states is relatively high when IMPE is used as the feature, while the classification accuracy of individual state samples is low when MPE and PE are used as features, for example, MPE for class 3 moderate damage samples, PE for class 2 slight damage samples and class 3 moderate damage samples. In general, IMPE method is better than MPE and PE methods, which shows that IMPE can effectively extract weak fault features of bearings.

The genetic algorithm and particle swarm optimization algorithm are compared with the improved chaotic firefly optimization algorithm proposed in this paper. Table 3 shows the training and test results of different SVM

TABLE 3: Training and testing results of different parameter optimization methods.

Parameter optimization	Training accuracy (%)	Test accuracy (%)	Time (s)
GA	98.75	96.25	71.67
PSOA	100	97.5	15.67
FA	100	97.5	9.00
CFOA	97.5	98.75	14.33

parameter optimization methods. It can be seen from the table that the optimal training accuracy of SVM optimized by improved chaotic firefly algorithm is slightly lower than that of particle swarm optimization algorithm and genetic algorithm when IMPE is used as fault feature, but its prediction accuracy is the highest, and compared with the other two algorithms, the improved chaotic firefly optimization algorithm takes the shortest time.

Figure 10 shows the comparison of the classification ability of our SVM classifier and several other classifiers based on the characteristics of IMPE in this paper. Among them, CO-PNN [42] is the improved coyote optimization algorithm based probabilistic neural network which get the accuracy of 94.26%. The method of 2D-CNN [43] get the accuracy of 95.31%. The coyote optimization algorithm based probabilistic neural network (ICOA-PNN) [44] and the spiking neural network (SNN) [45] get the accuracy of 98.26% and 97.18%, respectively. Through comparison, we can see that our method has achieved the best results.

Although the above four methods are based on neural network, they do have more vitality in classification, but

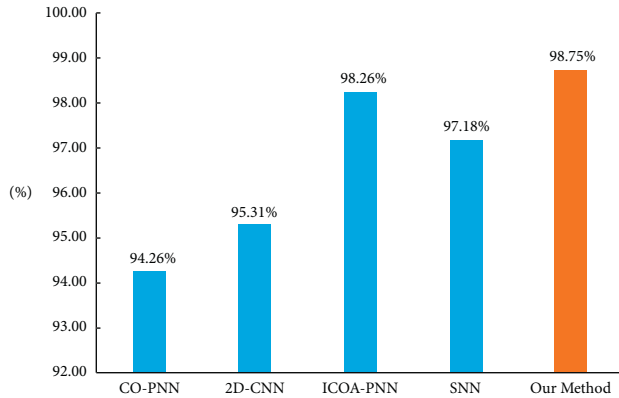


FIGURE 10: Comparison between our classifier and several other classifiers.

these methods often have better performance for image classification. Mechanical fault diagnosis is often based on one-dimensional signal. In those classification works, the weak classifier can also complete the task well. Importantly, from the mechanism of the algorithm, SVM has lower resource dependence, so it can be implemented on a simpler computing platform.

4. Conclusion

Permutation entropy is conducive to mining useful information from the complex and changeable weak signal with strong noise interference, which provides rich ideas, methods and technical means for early fault feature extraction of machinery. In view of the shortcomings of traditional MPE calculation process and the optimization of SVM parameters, the corresponding improvements are proposed. In order to solve the problem that only considering a single coarse-grained sequence under a certain scale may lead to the loss of feature information, this paper has proposed to calculate the time series with equal overlapping segments, that is, considering all coarse-grained sequences under the same scale, which can reflect the fault feature information of signals more comprehensively. In order to solve the problem that the feature extraction may not be refined enough by calculating the mean value of the first-order moment in traditional MPE calculation, a calculation method based on the skewness of the third-order moment has been proposed. The calculation method is more sensitive to the complexity and fluctuation of the signal, better describes the feature details, and effectively extracts the fault features. In order to optimize the key parameters of SVM classifier, an improved chaotic firefly algorithm has been proposed to optimize SVM parameters. IMPE has been applied to feature extraction of early weak fault of rolling bearings. Compared with traditional PE and MPE, it is found that the IMPE proposed in this paper can quantify different degrees of damage more effectively, is more sensitive to weak faults and has better stability. Experimental results show that the new method of early weak fault identification based on IMPE-SVM can effectively detect the rolling bearing faults with different severity.

Data Availability

Data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This research was funded by the National Natural Science Foundation of China (grant number: 61671470) and the Key Research and Development Program of China (grant number: 2016YFC0802904).

References

- [1] X. Zhu and Y. Wang, "A method of incipient fault diagnosis of bearings based on autocorrelation analysis and MCKD," *Journal of Vibration and Shock*, vol. 38, no. 24, pp. 183–188, 2019.
- [2] W. Sui, D. Zhang, and W. Wilson, "Fault diagnosis of rolling element bearings based on EMD and MKD," *Journal of Vibration and Shock*, vol. 34, no. 9, pp. 55–59, 2015.
- [3] A. Cf, A. Hl, and B. Rpl, "Hierarchical model updating strategy of complex assembled structures with uncorrelated dynamic modes," *Chinese Journal of Aeronautics*, vol. 35, 2021.
- [4] J. Tian, Y.-T. Ai, C.-W. Fei, F.-L. Zhang, and Y.-S. Choy, "Dynamic modeling and simulation of inter-shaft bearings with localized defects excited by time-varying displacement," *Journal of Vibration and Control*, vol. 25, no. 8, pp. 1436–1446, Apr 2019.
- [5] 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, Mar 2018.
- [6] G. Tang, X. Wang, and Y. He, "Rolling bearing fault diagnosis based on variational mode decomposition and permutation entropy," in *Proceedings of the International Conference on Ubiquitous Robots & Ambient Intelligence*, Xi'an, China, August 2016.
- [7] M. Kuai, G. Cheng, Y. Pang, and Y. Li, "Research of planetary gear fault diagnosis based on permutation entropy of CEEMDAN and ANFIS," *Sensors*, vol. 18, no. 3, p. 782, 2018.
- [8] S. Zhou, S. Qian, W. Chang, Y. Xiao, and Y. Cheng, "A novel bearing multi-fault diagnosis approach based on weighted permutation entropy and an improved SVM ensemble classifier," *Sensors*, vol. 18, no. 6, p. 1934, 2018.
- [9] L.-Y. Zhao, L. Wang, and R.-Q. Yan, "Rolling bearing fault diagnosis based on wavelet packet decomposition and multi-scale permutation entropy," *Entropy*, vol. 17, no. 12, pp. 6447–6461, 2015.
- [10] G. Xiong, C. Zhen, and I. Zhang, "Clustering method for intelligent identification of bearing faults based on CEEMDAN multi-scale permutation entropy," *Noise and vibration control*, vol. 40, no. 06, pp. 6–12, 2020.
- [11] J. Zheng, T. Liu, and R. Meng, "Rolling bearing fault diagnosis method based on generalized composite multiscale

- permutation entropy and PCA,” *Vibration and shock*, vol. 20, pp. 61–66, 2018.
- [12] Z. Huo, Y. Zhang, G. Jombo, and L. Shu, “Adaptive multiscale weighted permutation entropy for rolling bearing fault diagnosis,” *IEEE Access*, vol. 8, pp. 87529–87540, 2020.
- [13] Y. Li, Q. Gao, P. Li, J. Liu, and Y. Zhu, “Fault diagnosis of rolling bearing using a refined composite multiscale weighted permutation entropy,” *Journal of Mechanical Science and Technology*, vol. 35, no. 5, pp. 1893–1907, May 2021.
- [14] Y. Li, Q. Gao, B. Miao, W. Zhang, J. Liu, and Y. Zhu, “Application of the refined multiscale permutation entropy method to fault detection of rolling bearing,” *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 43, no. 5, pp. 280–314, 2021.
- [15] Z. Yin and J. Hou, “Recent advances on SVM based fault diagnosis and process monitoring in complicated industrial processes,” *Neurocomputing*, vol. 174, pp. 643–650, 2016.
- [16] L. Ren, W. Lv, S. Jiang, and Y. Xiao, “fault diagnosis using a joint model based on sparse representation and SVM,” *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 10, pp. 2313–2320, Oct 2016.
- [17] Y. Li, W. Zhang, Q. Xiong, D. Luo, G. Mei, and T. Zhang, “A rolling bearing fault diagnosis strategy based on improved multiscale permutation entropy and least squares SVM,” *Journal of Mechanical Science and Technology*, vol. 31, no. 6, pp. 2711–2722, Jun 2017.
- [18] B. Keshtegar, M. Bagheri, C.-W. Fei, C. Lu, O. Taylan, and D. K. Thai, “Multi-extremum-modified response basis model for nonlinear response prediction of dynamic turbine blisk,” *Engineering with Computers*, vol. 38, no. S2, pp. 1243–1254, 2021.
- [19] C. Lu, C.-W. Fei, H.-T. Liu, H. Li, and L.-Q. An, “Moving extremum surrogate modeling strategy for dynamic reliability estimation of turbine blisk with multi-physics fields,” *Aerospace Science and Technology*, vol. 106, Article ID 106112, 2020.
- [20] C. Lu, C.-W. Fei, Y.-W. Feng, Y.-J. Zhao, X.-W. Dong, and Y.-S. Choy, “Probabilistic analyses of structural dynamic response with modified Kriging-based moving extremum framework,” *Engineering Failure Analysis*, vol. 125, Article ID 105398, 2021.
- [21] C.-W. Fei, H. Li, H.-T. Liu, C. Lu, B. Keshtegar, and L.-Q. An, “Multilevel nested reliability-based design optimization with hybrid intelligent regression for operating assembly relationship,” *Aerospace Science and Technology*, vol. 103, Article ID 105906, 2020.
- [22] E. Zheng, S. Wu, and Z. Hu, “An improved PSO-SVM algorithm,” *Advances in Applied Mathematics*, vol. 10, no. 7, pp. 2305–2313, 2021.
- [23] H. Azami and J. Escudero, “Improved multiscale permutation entropy for biomedical signal analysis: interpretation and application to electroencephalogram recordings,” *Biomedical Signal Processing and Control*, vol. 23, pp. 28–41, 2016.
- [24] M. Källersjö, J.-S. Farris, A.-G. Kluge, and C. Bult, “Skewness and permutation,” *Cladistics*, vol. 8, no. 3, pp. 275–287, 1992.
- [25] P. O. Seglen, “The skewness of science,” *Journal of the American Society for Information Science*, vol. 43, no. 9, pp. 628–638, 1992.
- [26] C. Bandt and B. Pompe, “Permutation entropy: a natural complexity measure for time series,” *Physical Review Letters*, vol. 88, no. 17, Article ID 174102, 2002.
- [27] M. Matilla-García, “A non-parametric test for independence based on symbolic dynamics,” *Journal of Economic Dynamics and Control*, vol. 31, no. 12, pp. 3889–3903, 2007.
- [28] J. Zheng, J. Cheng, and Y. Yang, “Multi-scale permutation entropy and its applications to rolling bearing fault diagnosis,” *China Mechanical Engineering*, vol. 24, no. 19, pp. 2641–2646, 2013.
- [29] C. Schuldt, I. Laptev, and B. Caputo, “Recognizing human actions: a local SVM approach,” in *Proceedings of the 17th International Conference on Pattern Recognition (ICPR)*, pp. 32–36, British Machine Vis Assoc, Cambridge, UK, August 2004.
- [30] C. Jiao, *Comparative Study on Multi Classification Methods Based on Binary SVM*, Xi’an University of Electronic Science and Technology, Xi’an, China, 2011.
- [31] K. Ji and D. Han, “Application of SVM support vector machine in remote sensing image,” *Journal of Human City University*, vol. 25, no. 3, pp. 114–116, 2016.
- [32] L. Wang, *SVM Based on Mixed Kernel Function and its Application Research*, Dalian Maritime University, Dalian, China, 2016.
- [33] Q. Qin, *Hyperspectral Image Classification Based on SVM Kernel Function and Parameter Selection*, Huazhong University of Science and Technology, Wuhan, China, 2015.
- [34] M. Li, L. Yong, and J. Wang, “A new parameter optimization algorithm of SVM,” in *Proceedings of the International Conference on Advanced Intelligence & Awareness Internet*, Shenzhen, October 2011.
- [35] H. Guo, J. Yin, J. Zhao, Z. Huang, and Y. Pan, “Prediction of fatigue life of packaging EMC material based on RBF-SVM,” *International Journal of Materials and Product Technology*, vol. 49, no. 1, pp. 5–17, 2014.
- [36] I. Fister, I. Fister, X.-S. Yang, and J. Brest, “A comprehensive review of firefly algorithms,” *Swarm and Evolutionary Computation*, vol. 13, no. 1, pp. 34–46, 2013.
- [37] D. Wu and X. Ding, “T-S model identification based on improved firefly algorithm,” *Computer Simulation*, vol. 30, no. 3, pp. 327–330, 2013.
- [38] F. G. Hoffmann, J. G. Owen, and R. J. Baker, “mtDNA perspective of chromosomal diversification and hybridization in Peters’ tent-making bat (*Uroderma bilobatum*: phyllostomidae),” *Molecular Ecology*, vol. 12, no. 11, pp. 2981–2993, 2003.
- [39] S.-S. Gokhale and V.-S. Kale, “An application of a tent map initiated Chaotic Firefly algorithm for optimal overcurrent relay coordination,” *International Journal of Electrical Power & Energy Systems*, vol. 78, pp. 336–342, 2016.
- [40] S. Dash, A. Abraham, A. K. Luhach, J. Mizera-Pietraszko, and J. J. P. C. Rodrigues, “Hybrid chaotic firefly decision making model for Parkinson’s disease diagnosis,” *International Journal of Distributed Sensor Networks*, vol. 16, no. 1, Article ID 155014771989521, 2020.
- [41] W. A. Smith and R. B. Randall, “Rolling element bearing diagnostics using the Case Western Reserve University data: a benchmark study,” *Mechanical Systems and Signal Processing*, vol. 64–65, pp. 100–131, Dec 2015.
- [42] L. Bai and W. Xi, “early fault diagnosis of rolling bearing based empirical wavelet transform and spectral kurtosis,” in *Proceedings of the 2018 IEEE International Conference on Prognostics and Health Management (ICPHM)*, IEEE, WA, USA, June 2018.
- [43] J. Zheng, H. Pan, X. Qi, X. Zhang, and Q. Liu, “Enhanced empirical wavelet transform based time-frequency analysis

- and its application to rolling bearing fault diagnosis,” *Acta Electronica Sinica*, vol. 46, no. 2, pp. 358–364, 2018.
- [44] J. Ma, Z. Li, C. Li, L. Zhan, and G.-Z. Zhang, “Rolling bearing fault diagnosis based on refined composite multi-scale Approximate entropy and optimized probabilistic neural network,” *Entropy*, vol. 23, no. 2, p. 259, 2021.
- [45] J. Wang, D. Wang, S. Wang, W. Li, and K. Song, “fault diagnosis of bearings based on multi-sensor information fusion and 2D convolutional neural network,” *IEEE Access*, vol. 9, pp. 23717–23725, 2021.
- [46] Z.-A. Lin, Z.-A. Lei, and B. Zhz, “A spiking neural network-based approach to bearing fault diagnosis - ScienceDirect,” *Journal of Manufacturing Systems*, vol. 61, 2020.