

Research Article

An Integrated Health Condition Detection Method for Rotating Machinery Using Refined Composite Multivariate Multiscale Amplitude-Aware Permutation Entropy

Fuming Zhou ¹, Wuqiang Liu ¹, Ke Feng ¹, Jinxing Shen ¹ and Peiping Gong²

¹Field Engineering College of Army Engineering University, Nanjing 210007, China

²Training Base of Army Engineering University, Xuzhou 210004, China

Correspondence should be addressed to Ke Feng; 2020659722@qq.com and Jinxing Shen; 565423803@qq.com

Received 20 August 2020; Revised 26 November 2020; Accepted 1 December 2020; Published 22 December 2020

Academic Editor: Yong Chen

Copyright © 2020 Fuming Zhou 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.

With a view to realizing the fault diagnosis of rotating machinery effectively, an integrated health condition detection approach for rotating machinery based on refined composite multivariate multiscale amplitude-aware permutation entropy (RCmvMAAPE), max-relevance and min-redundancy (mRmR), and whale optimization algorithm-based kernel extreme learning machine (WOA-KELM) is presented in this paper. The approach contains two crucial parts: health detection and fault recognition. In health detection stage, multivariate amplitude-aware permutation entropy (mvAAPE) is proposed to detect whether there is a fault in rotating machinery. Afterward, if it is detected that there is a fault, RCmvMAAPE is employed to extract the initial fault features that represent the fault states from the multivariate vibration signals. Based on the multivariate expansion and multiscale expansion of amplitude-aware permutation entropy, RCmvMAAPE enjoys the ability to effectively extract state information on multiple scales from multichannel series, thereby overcoming the defect of information loss in traditional methods. Then, mRmR is adopted to screen the sensitive features so as to form sensitive feature vectors, which are input into the WOA-KELM classifier for fault classification. Two typical rotating machinery cases are conducted to prove the effectiveness of the raised approach. The experimental results demonstrate that mvAAPE shows excellent performance in fault detection and can effectively detect the fault of rotating machinery. Meanwhile, the feature extraction method based on RCmvMAAPE and mRmR, as well as the classifier based on WOA-KELM, shows superior performance in feature extraction and fault recognition, respectively. Compared with other fault identification methods, the raised method enjoys better performance and the average fault recognition accuracy of the two typical cases in this paper can all reach above 98%.

1. Introduction

As one of the widely applied mechanical equipment, rotating machinery plays a vital role in industrial production. Nevertheless, it usually operates in harsh environments such as heavy load and high speed, which greatly increases the risk of faults. These faults may result in equipment shutdown and even casualties cause if they are not dealt with in time [1, 2]. Due to the particularity of industrial machinery, direct disassembly overhaul will affect normal production. Hence, research on nondisassembly health condition detection technology of rotating machinery has always been a hotspot. When encountering faults, some changes will occur in the

internal structure of rotating machinery, which affects the frequency and amplitude of vibration signals. It indicates that the vibration signals contain a wealth of information related to the operating states of rotating machinery [3, 4]. Consequently, analyzing vibration signals is a feasible method for fault diagnosis [5].

The essences of vibration signals-based fault diagnosis are the fault feature extraction and pattern recognition issues. Among which, how to extract the features which can represent the working states from the vibration signals is the key in fault diagnosis. In the past decades, time-frequency analysis is widely applied in feature extraction of vibration signals. Many time-frequency analysis methods such as

empirical mode decomposition (EMD) [6], local mean decomposition (LMD) [7], wavelet packet transform (WPT) [8], and variational mode decomposition (VMD) [9] are applied to fault diagnosis of rotating machinery. Unfortunately, the vibration signals of rotating machinery usually exhibit nonlinear and nonstationary characteristics, which cause the above methods to have some defects in practical applications. For instance, WPT needs to choose the suitable wavelet kernel function [8] and VMD need to set the penalty factor α and the number of intrinsic mode functions (IMFs) K before processing the vibration signals [10], thereby the self-adaptive capacity of them is poor. EMD enjoys good adaptability, but it has defects such as mode mixing and end effect. In addition, the application of time-frequency analysis methods alone requires the operators to have a certain knowledge reserve, which limits the efficiency and application scope of these methods. Therefore, developing an efficient and accurate fault feature extraction tool is urgent and necessary.

Recently, the entropy-based theory has been widely adopted as feature extraction tool in the field of fault diagnosis due to its excellent performance in measuring the nonlinear complexity of time series [11]. Entropy methods that are commonly applied include approximate entropy (AE) [12], sample entropy (SE) [13], fuzzy entropy (FE) [14], and permutation entropy (PE) [15]. Among them, AE is highly dependent on the data length and is prone to undefined entropy value. SE and FE are time-consuming, so they are not suitable for processing signals with a large amount of data, while PE is favored by many scholars because of its high computational efficiency and strong anti-noise ability. Zhang et al. [16] adopted PE to detect bearing faults and proposed a bearing fault diagnosis model based on PE, ensemble empirical mode decomposition, and optimized SVM. Kuai et al. [17] proposed a fault diagnosis method for planetary gears based on PE, CEEMDAN, and ANFIS. CEEMDAN is applied to decompose the vibration signal of planetary gears, and PE is used to extract the characteristics of the obtained IMFs. Finally, ANFIS is used as a classifier to complete fault identification. Nevertheless, PE also exists some inherent defects. For example, it loses sight of the influence of amplitude information of signals on the entropy value, which may lose the crucial information. To address this problem, Azami et al. [18] presented the amplitude-aware permutation entropy (AAPE), which is not only sensitive to the frequency but also sensitive to the amplitude of signals. The excellent performance of AAPE has been verified through the simulation and biological signals experiments.

However, AAPE also possesses some shortcomings that cannot be ignored. Firstly, AAPE only measures the complexity of the measured signal on one temporal scale, thereby cannot capture the long correlation of the signal [19]. To address this question, based on multiscale entropy theory [19], multiscale amplitude-aware permutation entropy (MAAPE) was proposed to extract the fault information of rolling bearings [20]. Unfortunately, MAAPE enjoys poor stability, especially for short-time series. The defect will cause MAPPE to produce unreliable entropy values on high

scales. Secondly, AAPE cannot extract fault features from multichannel vibration signals, which limits its ability to extract fault information for large equipment. For large equipment, the long transmission path will reduce the vibration impulse to a certain extent. In other words, the fault information will be lost. Therefore, the vibration signal collected by single channel is usually not enough to provide enough fault information to identify the fault type [21]. It is necessary to improve AAPE so that it can extract fault features from multichannel vibration signals synchronously.

With a view to solving the aforementioned defects, refined composite multivariate multiscale amplitude-aware permutation entropy (RCmvMAAPE) is presented in this paper. Compared with the existing AAPE methods, the proposed RCmvMAAPE possesses two main improvements. Firstly, refined composite multiscale method is employed to substitute the traditional multiscale method in MAAPE to overcome the entropy instability problem [22]. In addition, on the basis of multidimensional embedding reconstruction theory [23], AAPE is expanded to multivariate AAPE (mvAAPE) to measure the complexity of multichannel vibration signals. Based on the above improvements, RCmvMAAPE overcomes the abovementioned defects and can stably measure the complexity of multichannel signals on multiple scales. The performance of RCmvMAAPE is comprehensively tested utilizing a variety of synthetic signals in this paper, and the results indicate that RCmvMAAPE can availablely measure the complexity of multivariate signals. In view of the advantages of RCmvMAAPE, this paper employs it to extract the fault features of multichannel vibration signals of rotating machinery.

As we know, the fault features distributed on multiple scales extracted by RCmvMAAPE are a high-dimensional feature vector. Among which, some sensitive features can effectively represent the fault information, but some redundant features not only affect the accuracy of subsequent fault classification but also reduce the diagnosis efficiency. For this reason, it is necessary to compress the high-dimensional fault features to improve the fault recognition rate. The max-relevance and min-redundancy (mRmR) is a typical features selection method based on spatial search, which uses mutual information to measure the relevance and redundancy of features [24]. The maximum correlation indicates that the feature has a large correlation with the sample category, that is, it can reflect the sample category information to the greatest extent. Minimal redundancy means that the correlation between features is the smallest, that is, the redundancy of features is the smallest. This paper adopts mRmR to select the sensitive features to form sensitive features vectors that represent the fault state of rotating machinery.

Afterward, different fault states of rotating machinery will be identified according to the sensitive feature vectors, namely, pattern recognition. At this stage, a classifier with high computational efficiency and good generalization performance is needed. Kernel extreme learning machine (KELM) [25] is a machine learning method that combines ELM and kernel function. While retaining the high calculation efficiency of ELM, the introduction of kernel function

enables KELM to enjoy stronger generalization ability compared with commonly used classifiers such as BP neural network (BP) [26], support vector machine (SVM) [27], and extreme learning machine (ELM) [28] when dealing with linear inseparable problems; meanwhile, KELM is sensitive to parameter setting due to the existence of kernel function. To choose the best parameters, we need to employ a suitable optimization algorithm to determine the best parameters of KELM. Commonly used optimization algorithms consist of particle swarm optimization (PSO) [29], ant colony optimization (ACO) [30], and whale optimization algorithm (WOA) [31]. Among which, WOA has attracted more and more attention due to its uncomplicated operation, less adjustment parameters, and strong capability to jump out of local optimum. Therefore, WOA is utilized to iteratively select the optimal parameter of KELM to build a classifier based on WOA-KELM. The low-dimensional sensitive feature vectors are input into WOA-KELM so as to judge the fault type of the rotating machinery.

Consequently, a new integrated health detection method for rotating machinery is proposed, which includes two parts: fault detection and fault identification. In the fault detection stage, mvAAPE is employed to extract the features of the vibration signals to determine whether the rotating machinery is malfunctioning. By introducing the key link of fault detection, the unnecessary disassembly and maintenance of the equipment can be avoided, and the damage to the equipment can be reduced. In the fault identification stage, the presented method based on RCmvMAAPE, mRmR, and WOA-KELM is applied to diagnose different fault types and fault severity of rotating machinery. Two examples are conducted to prove the performance of the proposed method and its superiority compared to other existing methods.

The rest of the paper is arranged as follows: in Sections 2 and 3, the basic theory of RCmvMAAPE and WOA-KELM is introduced in detail; Section 4 displays the steps of the proposed approach; two typical cases are adopted for experiments to verify the excellent performance of the proposed approach in Section 5; finally, this paper is summarized in Section 6.

2. The Basic Theory of RCmvMAAPE

2.1. Multivariate Amplitude-Aware Permutation Entropy

2.1.1. AAPE. AAPE is a method based on PE, which is a powerful tool for analyzing nonlinear time series. Therefore, it is necessary to introduce the concept of PE firstly. The original theory of PE is reviewed in [15].

For a given time series $X = \{x_i\}$, $i = 1, 2, \dots, N$, at any time point t , the m dimensional reconstruction vector can be obtained as

$$X_t^{m,d} = [x_t, x_{t+d}, \dots, x_{t+(m-2)d}, x_{t+(m-1)d}], \quad t = 1, 2, \dots, N - (m-1)d, \quad (1)$$

where m denotes the embedding dimension and d denotes the time delay.

For each reconstruction vector, in accordance with the size of the elements in ascending order, the permutation $\pi_{r_0, r_1, \dots, r_{m-1}}$ can be acquired, which fulfills that

$$\left[x_{t+(j_1-1)d}, x_{t+(j_2-1)d}, \dots, x_{t+(j_{m-1}-2)d}, x_{t+(j_m-1)d} \right], \quad (2)$$

where j_* represents the index of the column of each element in the reconstructed component. Accordingly, there are $m!$ possible permutation patterns, of which the i -th permutation is marked as π_i .

The relative frequency of π_i can be expressed as

$$p(\pi_i) = \frac{g(\pi_i)}{N - (m-1)d}, \quad (3)$$

where $g(\pi_i)$ represents the function that counts the number of π_i in $X_t^{m,d}$. The value of $g(\pi_i)$ will increase by 1 if the permutation order of the internal elements of $X_t^{m,d}$ is π_i .

Consequently, based on the calculation theorem of Shannon entropy, PE can be defined as

$$PE(X, m, d) = - \sum_{i=1}^{m!} p(\pi_i) \ln p(\pi_i). \quad (4)$$

Nevertheless, PE enjoys some nonnegligible deficiencies, which led to its inability in describing the irregularity of the series. Firstly, from the theoretical point of view, the original PE algorithm only considers the effect of the ordinal structure of the time series on the entropy value, but the amplitude information of each mapped element in the series is ignored. Secondly, when there are elements with equal amplitude, their influence on the entropy value cannot be accurately estimated. In view of the aforementioned defects of PE, Azami proposed AAPE to significantly enhance the performance of PE [18]. The basic principle of the AAPE algorithm is as follows:

Supposing that the starting value of $p(\pi_i)$ is 0, for the reconstruction vector $X_t^{m,d}$, when the time t adds from 1 to $N - m + 1$, the value of $p(\pi_i)$ is updated whenever the permutation is π_i .

$$p^{\text{update}}(\pi_i) = p(\pi_i) + \left(\frac{\alpha}{m} \sum_{k=1}^m |x_{t+(k-1)d}| + \frac{1-\alpha}{m-1} \sum_{k=2}^d |x_{t+(k-1)d} - x_{t+(k-2)d}| \right), \quad (5)$$

where $\alpha \in [0, 1]$ denotes the adjustment coefficient which is utilized to adjust the weight of the time series amplitude

average and the deviation between the amplitudes. Thus, the probability of $p(\pi_i)$ is

$$p(\pi_i) = \frac{p^{\text{update}}(\pi_i)}{\sum_{t=1}^{N-m+1} \left((\alpha/m) \sum_{k=1}^m |x_{t+(k-1)d}| + (1 - \alpha/m - 1) \sum_{k=2}^m |x_{t+(k-1)d} - x_{t+(k-2)d}| \right)}. \quad (6)$$

The AAPE of time series x can be defined as

$$\text{AAPE}(X, m, d, \alpha) = - \sum_{i=1}^{m!} p(\pi_i) \ln p(\pi_i). \quad (7)$$

2.1.2. *mvAAPE*. To describe the complexity of multichannel time series, it is necessary to extend the AAPE to

$$Z_t^{m,d} = [x_{c,t}, x_{c,t+d}, \dots, x_{c,t+(m-2)d}, x_{c,t+(m-1)d}], t = 1, 2, \dots, N - (m-1)d. \quad (8)$$

- (2) Arrange the reconstruction time series $Z_i^{m,d}$ in ascending order as $[x_{c,i+(j_1-1)d} \leq x_{c,i+(j_2-1)d} \leq \dots \leq x_{c,i+(j_{m-1}-1)d} \leq x_{c,i+(j_m-1)d}]$. At the same time, there are $m!$ potential permutations $\pi_i, 1 \leq i \leq m!$.
- (3) For c -th channel, supposing that the starting value of $p(\pi_{c,i})$ is 0, for the reconstruction series $Z_i^{m,d}$, when t gradually increases from 1 to $N - m + 1$, the value of $p(\pi_{c,i})$ will be renewed as $\pi_{c,i}$ appears.

multivariate analysis so as to put forward multivariate amplitude-aware permutation entropy (mvAAPE). The definition of mvAAPE is described as follows:

- (1) Given a p -channel series $X = \{x_{c,1}, x_{c,2}, \dots, x_{c,i}, \dots, x_{c,N}\}, c = 1, 2, \dots, p$, phase space reconstruction is performed as follows:

$$p^{\text{update}}(\pi_{c,i}) = p(\pi_{c,i}) + \left(\frac{\alpha}{m} \sum_{k=1}^m |x_{c,t+(k-1)d}| + \frac{1 - \alpha}{m - 1} \sum_{k=2}^d |x_{c,t+(k-1)d} - x_{c,t+(k-2)d}| \right). \quad (9)$$

- (4) Calculate the relative frequency of i -th permutation in c -th channel $\pi_{c,i}$ as follows:

$$p(\pi_{c,i}) = \frac{p^{\text{update}}(\pi_{c,i})}{\sum_{c=1}^p \sum_{t=1}^{N-m+1} \left((\alpha/m) \sum_{k=1}^m |x_{c,t+(k-1)d}| + (1 - \alpha/m - 1) \sum_{k=2}^m |x_{c,t+(k-1)d} - x_{c,t+(k-2)d}| \right)}. \quad (10)$$

For p -channel time series, $p(\pi_{c,i})$ satisfies $\sum_{c=1}^p \sum_{j=1}^{m!} p(\pi_{c,i}) = 1$.

- (5) The probability of the i -th pattern π_i in p -channel time series X can be calculated as follows:

$$p(\pi_i) = \sum_{c=1}^p p(\pi_{c,i}). \quad (11)$$

- (6) Based on the definition of Shannon entropy, mvAAPE is expressed as

$$\text{mvAAPE}(X, m, \alpha, d) = - \sum_{i=1}^{m!} p(\pi_i) \ln p(\pi_i), \quad (12)$$

where mvAAPE actually extends the application of AAPE from univariate analysis to multivariate analysis. However, mvAAPE only analyzes the multichannel time series on one temporal scale, while the measured time series often contains information on multiple scales. Therefore, the key information will lose if only a single scale analysis is conducted. In response to this problem, mvMAAPE that is able to analyze time series on multiple scales is proposed.

2.2. *mvMAAPE*. The principle of mvMAAPE is as follows:

- (1) For p -channel series $U = \{u_{k,1}, u_{k,2}, \dots, u_{k,i}, \dots, u_{k,L}\}, k = 1, 2, \dots, p$, the multivariate coarse-grained time series at scale factor τ is defined as follows:

$$y_{k,j}^\tau = \frac{1}{\tau} \sum_{b=(j-1)\tau+1}^{j\tau} u_{k,i}, 1 \leq j \leq \frac{L}{\tau}, 1 \leq k \leq p. \quad (13)$$

When $\tau > 1$, the multivariate series is divided into coarse-grained time series of length $[L/\tau]$.

- (2) Calculate the mvAAPE of τ multivariate coarse-grained time series and the result is as follows:

$$\text{mvMAAPE}(U, m, \alpha, d, \tau) = \text{mvAAPE}(y_{k,j}^\tau, m, \alpha, d), \quad (14)$$

where mvMAAPE overcomes the shortcomings that PE does not consider the amplitude information; meanwhile, the combination with multivariate analysis improves the utilization of multichannel information, which is essentially an assessment of the irregularity of multichannel data. The

evaluation principle can be summarized into two aspects: (1) if the entropy value of the multivariate series X is greater than that of series Y on most scale factors, it can be shown that X is more random than Y and more prone to dynamic mutations. (2) If the entropy value of X decreases significantly with the increase of the scale factor, it indicates that the information included in X mainly appears on a smaller scale factor, such as a random white noise signal. mvMAAPE considers the interrelationship of each time series in multichannel data and comprehensively evaluates each dimension of multichannel series. Therefore, mvMAAPE can effectively detect the mutation change of multichannel series.

2.3. Refined Composite Multivariate Multiscale Amplitude-Aware Permutation Entropy

2.3.1. Basic Principle. The mvMAAPE realizes multivariate and multiscale analysis by extending the mvAAPE method to multiple scales, so as to obtain more useful information. However, the coarse-graining method adopted by mvMAAPE has serious defects, which leads to incomplete information analysis. For instance, the calculation of mvMAAPE only considers the coarse-graining series starting from $u_{k,1}$ and ignores the coarse-graining series such as $u_{k,2}$ at scale factor τ . However, the remaining $\tau - 1$ time series also contain the key information, and the direct neglect will lead to insufficient analysis and affect the analysis effect. Therefore, the refined composite multiscale coarse-graining approach is employed to achieve accurate and sufficient analysis. The implementation principle of the coarse-graining method is presented in Figure 1.

The Detailed Procedures of RCmvMAAPE are Described as follows:

- (1) For p -channel series $U = \{u_{k,1}, u_{k,2}, \dots, u_{k,i}, \dots, u_{k,L}\}$, $k = 1, 2, \dots, p$, the coarse-grained multivariate time series are computed on a given scale factor τ and the elements of the a -th coarse-grained time series $Y_a^\tau = \{y_{k,i,1}^\tau, y_{k,i,2}^\tau, \dots\}$ are computed by

$$y_{k,i,a}^\tau = \frac{1}{\tau} \sum_{b=a+(i-1)\tau}^{a+i\tau-1} u_{k,b}, \quad (15)$$

where $1 \leq i \leq L/\tau$, $1 \leq k \leq p$, $1 \leq a \leq \tau$. For the scale factor τ , there will be τ diverse coarse-grained multivariate time series.

- (2) For each coarse-grained multivariate series, the marginal relative frequencies $p(\pi_j)$ are computed. Then, the average relative frequencies $\overline{p(\pi_j)}$ can be acquired by

$$\overline{p(\pi_j)} = \frac{1}{\tau} \sum_{a=1}^{\tau} p_a(\pi_j). \quad (16)$$

- (3) The RCmvMAAPE of original multivariate time series is computed as follows:

$$RCmvMAAPE(U, m, \alpha, d, \tau) = - \sum_{j=1}^{m!} \overline{p(\pi_j)} \ln \overline{p(\pi_j)}. \quad (17)$$

In the RCmvMAAPE approach, there are three key parameters, namely, the m , α , and d . For the embedding dimension m , if the value is too small, the reconstructed vector includes too few states and the algorithm will lose its validity and significance, whereas if m is too large, the phase space reconstruction will homogenize the time series, which not only increases the amount of calculation but also cannot reflect the slight change of the time series. According to references [18, 29], the AAPE for univariate analysis usually sets the embedding dimension to 3–7, and the optimal parameters of the univariate analysis method and multivariate analysis are generally consistent, so this article sets the embedding dimension to $m = 5$. The adjustment coefficient α is usually set to 0.5 according to reference [18], so this article sets $\alpha = 0.5$. Time delay has little effect on the performance of the algorithm, so in this article, $d = 1$.

2.3.2. Performance Analysis. To validate the performance of RCmvMAAPE, other multivariate analysis approaches are compared with it to reflect its advantages in extracting the complexity of multichannel signals. White Gaussian noise (WGN) and $1/f$ noise are two signals that are widely adopted to evaluate the univariate and multivariate analysis method. Compared with WGN signals, the power spectrum of $1/f$ noise is more complicated and includes more mode information. The generation of WGN is randomly distributed, so the probability of its state transition matrix appearing is approximately equal. On the contrary, $1/f$ noise is a long-range correlation signal, and the irregularity of $1/f$ noise is lower than that of WGN. Consequently, the complexity of $1/f$ noise is higher than that of WGN. Considering the universality, WGN and $1/f$ noise are employed to create a multichannel signal with three different channels to analyze RCmvMAAPE, mvMAAPE, RCmvMSE, and RCmvMPE. They are (a) three channel WGN; (b) three channel $1/f$ noise; (c) two channel WGN and one channel $1/f$ noise; and (d) two channel $1/f$ noise and one channel WGN. There are 25 groups (length 2048) of the synthesized signals in each case.

For sake of verifying the advantages of the proposed approach in measuring the complexity of multivariate signals, RCmvMAAPE, mvMAAPE, RCmvMPE, and RCmvMSE of four kinds of multivariate synthetic signals are calculated. The mean standard deviation diagrams of the four methods are shown in Figure 2. Compared with mvMAAPE, RCmvMPE, and RCmvMSE, the standard deviation of RCmvMAAPE is significantly smaller than mvMAAPE and RCmvMSE, which indicates that the stability and robustness of RCmvMAAPE are stronger than mvMAAPE and RCmvMSE. It can be clearly seen from the figure that RCmvMAAPE can effectively separate four multivariate synthetic signals, proving that RCmvMAAPE

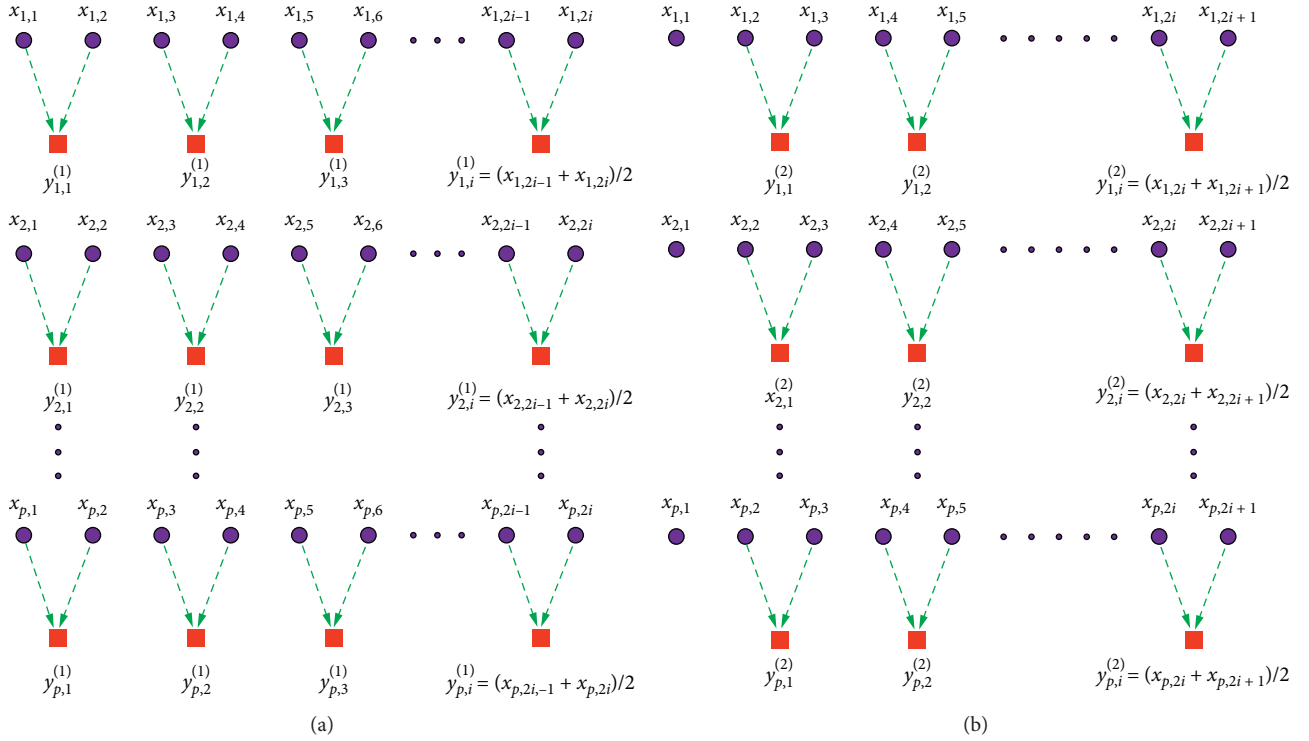


FIGURE 1: Illustration of refined composite coarse-grained approach for multivariate data with scale factor 2. (a) First coarse-grained time series; (b) second coarse-grained time series.

has better separation performance. What's more, the fluctuation of the RCmvMPE curve is greater than that of RCmvMAAPE, especially the fluctuation of (d) is obvious. This phenomenon shows that RCmvMAAPE is more stable when analyzing multivariate data and is not prone to large errors. In addition, when the scale factor is 14–20, RCmvMSE cannot effectively distinguish between (b) and (d). Similarly, mvMAAPE cannot effectively distinguish (a) and (c); meanwhile, the entropy value of four multivariate signals has extremely large fluctuation, which also verifies that the traditional coarse-graining method is prone to large errors. In a word, compared with the other three multivariate analysis methods, RCmvMAAPE enjoys better separation performance and robustness, thereby can better characterize the complexity of multivariate signals.

3. The Principle of the WOA-KELM

3.1. Kernel Extreme Learning Machine. Kernel extreme learning machine is a training algorithm based on single-hidden layer feedforward neural network. It does not require to repeatedly adjusting the hidden layer parameters [28]. In addition, the conventional single-hidden layer feedforward neural network parameter training problem is transformed into solving linear equations, and the smallest norm least-squares solution obtained is used as the network output

weight. The whole training process is completed once. Therefore, the training speed is greatly improved and the generalization performance is better.

For input and output data, the goal of ELM is to simultaneously minimize training error and output weight norm, which can be expressed as follows:

$$\begin{cases} \min \sum \|\beta \cdot h(x_i) - t_i\|^2, \\ \min \|\beta\|, \end{cases} \quad (18)$$

where β is the connection weight vector between the hidden layer and the output layer and $h(x_i)$ is the kernel mapping of the hidden layer.

The optimization problem of equation (18) is simplified to the following constraint problem:

$$\begin{cases} \min L = \frac{1}{2} \|\beta\|^2 + C \frac{1}{2} \sum_{i=1}^N \xi_i^2, \\ h(x_i)\beta = t_i - \xi_i, i = 1, 2, \dots, N, \end{cases} \quad (19)$$

where ξ_i stands for training error and C denotes the penalty factor. Using the theory of orthogonal projection, the training process of ELM is equivalent to solving the following dual optimization problems:

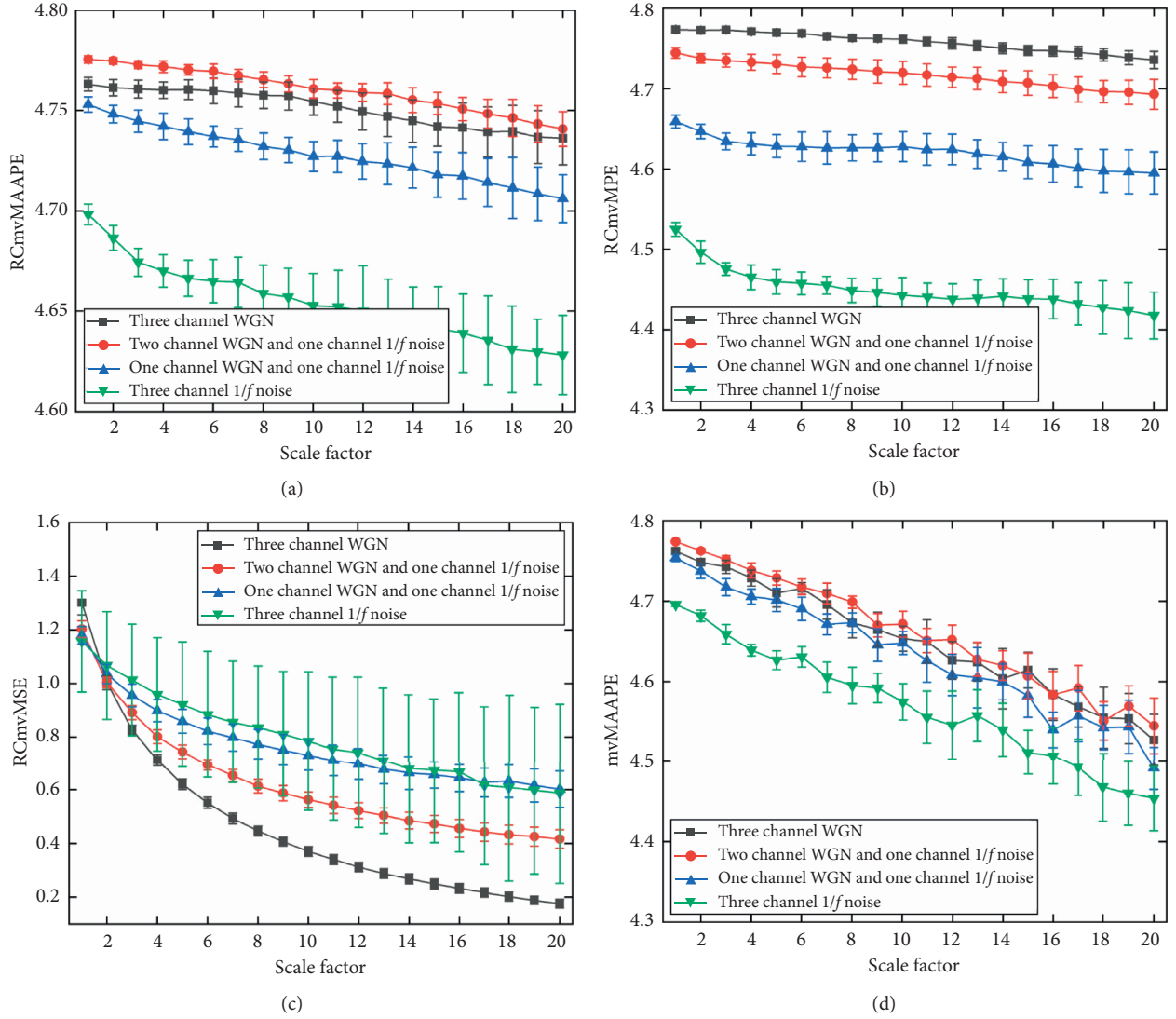


FIGURE 2: RCmvMAAPE, RCmvMPE, RCmvMSE, and mvMAAPE of multivariate synthetic signals. (a) RCmvMAAPE; (b) RCmvMPE; (c) RCmvMSE; and (d) mvMAAPE.

$$L_{\text{ELM}} = \frac{1}{2} \|\beta\|^2 + C \frac{1}{2} \sum_{i=1}^N \xi_i^2 - \sum_{i=1}^N \alpha_i (h(x_i)\beta - t_i + \xi_i), \quad (20)$$

where α_i is the Lagrangian operator, and the derivative of it is

$$\frac{\partial L_{\text{ELM}}}{\partial \beta} = 0 \implies \beta = \sum \alpha_i (h(x_i))^T = H^T \alpha, \quad (21)$$

$$\frac{\partial L_{\text{ELM}}}{\partial \xi_i} = 0 \implies \alpha_i \xi_i = 0, \quad (22)$$

$$\frac{\partial L_{\text{ELM}}}{\partial \alpha_i} = 0 \implies h(x_i)\beta - t_i + \xi_i = 0, \quad (23)$$

where $\alpha = [\alpha_1, \dots, \alpha_N]^T$.

Substituting formulas (20) and (21) into formula (22), the formula (23) can be equivalently written as follows:

$$\left(\frac{I}{C} + HH^T\right)\alpha = T. \quad (24)$$

The corresponding output function of ELM is described as follows:

$$f(x) = h(x)\beta = h(x)H^T \left(\frac{I}{C} + HH^T\right)^{-1} T. \quad (25)$$

It can be seen from the formula (25) that the parameter I/C is added to the main diagonal in the unit diagonal HH^T , thereby its eigenvalue cannot be 0. Then, the weight vector is

computed. ELM is more stable and has strong generalization ability in this way.

The kernel function is introduced into ELM and the KELM algorithm is proposed. Mercer condition is applied to define the kernel function matrix of KELM as follows:

$$\begin{aligned}\Omega &= HH^T, \\ \Omega_{i,j} &= h(x_i) \times h(x_j) = K(x_i, x_j),\end{aligned}\quad (26)$$

where $K(x_i, x_j)$ denotes the kernel function and the elements of the kernel matrix $\Omega_{i,j}$ in row i and column j , $i, j \in (1, 2, \dots, N)$.

Therefore, it can be concluded that the actual output of the KELM model is

$$f(x) = h(x)H^T \pi \left(\frac{I}{C} + HH^T \right)^{-1} T = \begin{bmatrix} K(x, x_1) \\ \dots \\ K(x, x_N) \end{bmatrix} \left(\frac{I}{C} + \Omega \right)^{-1} T. \quad (27)$$

3.2. Whale Optimization Algorithm. Whale optimization algorithm (WOA) is a novel heuristic search optimization algorithm [31]. Its advantages lie in its uncomplicated operation, less adjustment parameters, and strong capability to jump out of local optimum. The algorithm mainly imitates three behaviors of humpback whale, including encircling prey, hunting prey, and searching prey.

WOA supposes that the current best candidate solution is the target quarry or close to the best. After defining the best search agent, other search agents will therefore try to renew their best-located search agents. The update formula of WOA position is as follows:

$$\begin{aligned}D &= |CX^*(t) - X(t)|, \\ X(t+1) &= X^*(t) - AD,\end{aligned}\quad (28)$$

where A and C are the coefficients; t is the number of iterations; $X(t)$ represents the current position vector of the whale; and $X^*(t)$ denotes the best whale position vector so far. The mathematical expressions of A and C are as follows:

$$\begin{aligned}A &= 2ar_1 - a, \\ C &= 2 \cdot r_2, \\ a &= 2 \left(1 - \frac{t}{T_{\max}} \right),\end{aligned}\quad (29)$$

where T_{\max} represents the maximum number of iterations and r_1 and r_2 are random numbers in the interval $[0, 1]$. The value of a decreases linearly from 2 to 0, and t is the number of iterations.

When hunting, humpback whales not only swim to the prey in spiral shape but also contract the encircling circle. The position of whales is updated with 50% probability between the contraction mechanism and the spiral model.

$$X(t+1) = \begin{cases} X^*(t) - A \cdot D & \text{if } p < 0.5, \\ D' \cdot e^{bl} \cdot \cos(2\pi l) + X^*(t) & \text{if } p \geq 0.5, \end{cases} \quad (30)$$

where $D' = |X^*(t) - X(t)|$ denotes the distance between the whale and its prey; the constant b is used to define the spiral shape; and l is a random number in $[-1, 1]$.

When the humpback whale attacks the prey, by linearly reducing the value of parameter a , the fluctuation range of A is continuously decreased and the value of A in the interval $[-a, a]$ decreases continuously as a decreases. When the value of A is in the interval $[-1, 1]$, the solution position of the whale's next search agent will be any position between the current position and the prey position. By simulating the behavior of the humpback whale attacking the prey, the development capability of local search is shown. When the random value of A is greater than 1 or less than -1 , the humpback whale search agent moves away from the prey to search, thereby finding a more suitable prey, which shows the exploration function of the whale optimization algorithm in the global search.

3.3. Whale Optimization Algorithm-Based Kernel Extreme Learning Machine (WOA-KELM). Considering that the performance of the KELM is easily affected by penalty factors and kernel parameters, a new method for optimizing the kernel extreme learning machine by whale optimization algorithm is raised. The optimization procedure is presented in Figure 3, and the detailed step is as follows:

- (1) Input training set and testing set samples and normalize the two sample sets, respectively.
- (2) Initialize the position of whale population and set the population number to N . The maximum iteration number is T_{\max} .
- (3) Initialize the parameters of KELM and select the corresponding fitness function.
- (4) The fitness of each whale is computed and sorted according to the fitness value, so as to continuously update the whale population.
- (5) When the fitness value meets the conditions or reaches the maximum number of iterations, the optimization process is terminated.
- (6) According to the optimal penalty factor and kernel function parameter, the KELM fault diagnosis model is established.

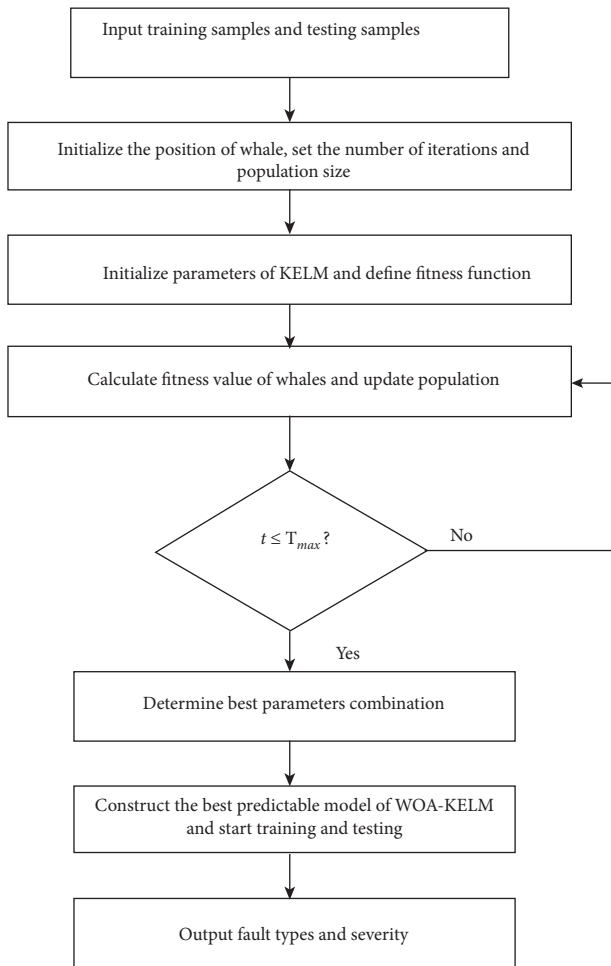


FIGURE 3: The flow chart of WOA-KELM algorithm.

- (7) The trained KELM health condition detection model is employed to output the fault type and severity of the testing data.

4. The Proposed Approach

In this study, considering that RCmvMAAPE possesses excellent performance of processing multivariate time series, it is used to extract the fault features of rotating machinery. Combining mRmR and WOA-KELM, an integrated health condition detection method for rotating machinery is proposed. The method includes fault detection and health condition recognition.

4.1. Fault Detection. The ability of mvAAPE to measure the complexity of multivariate nonlinear data and the probability of dynamic mutation is the basis for fault diagnosis. Since mvAAPE is proposed based on mvPE, it inherits the ability of mvPE to detect failures. The inconsistent entropy values of mvAAPE corresponding to different states are a prerequisite for fault screening.

The mvAAPE values of the rotating machinery vibration signals in all fault states are greater than that in the normal state, and the difference is obvious. Therefore, mvAAPE can

be applied for fault screening. In order to determine the screening criteria intuitively, a threshold based on mvMAAPE is set. When the mvMAAPE value of the vibration signal of rotating machinery in an unknown state is less than the threshold, the state is determined to be healthy. Conversely, if it is greater than the threshold, it is determined that there is a fault.

4.2. Health Condition Recognition. After fault detection, if it is detected that there is a fault in rotating machinery, further analysis is required to judge the type and severity of the fault. Firstly, RCmvMAAPE is employed to acquire the nonlinear complex information of fault multichannel vibration signals to form the initial fault feature vectors. However, the RCmvMAAPE values at all scales may include redundant information, so it is necessary to compress the feature dimensions to obtain sensitive feature vectors. The mRmR is a dimensionality reduction algorithm for nonlinear data, which uses mutual information to measure the correlation and redundancy of features, so as to realize the importance ranking of features. Therefore, the mRmR is utilized to screen the initial fault features to obtain sensitive feature vectors. Finally, the whale optimization algorithm is utilized to optimize the kernel function parameter and penalty factor of KELM to construct the optimal classification model and accomplish the health condition recognition of rotating machinery.

The flowchart of the raised approach is shown in Figure 4 and the implementation procedures of the integrated health condition detection method are listed as follows:

- (1) Multichannel vibration signals of rotating machinery under diverse working conditions are collected.
- (2) Divide the collected vibration data into multiple nonoverlapping samples of length N .
- (3) Compute the mvAAPE value of the vibration signal and establish a threshold based on mvAAPE to determine the health condition of the rotating machinery. If the mvAAPE value of the vibration signal to be detected is less than the threshold value, it indicates that the rotating machinery is healthy. The output is normal and the diagnosis terminates. Otherwise, the next step is conducted to judge the fault type and severity of the rotating machinery.
- (4) RCmvMAAPE is utilized to extract fault information from fault vibration signals of rotating machinery to generate the initial fault features.
- (5) The mRmR method is employed to screen the sensitive feature from the initial fault feature to form the sensitive feature vectors.
- (6) The training set samples are utilized to train the WOA-KELM-based multiclassifier.
- (7) The testing set samples are fed to the trained multiclassifier for prediction. The fault type and severity are recognized in line with the output of WOA-KELM multifault classifier.

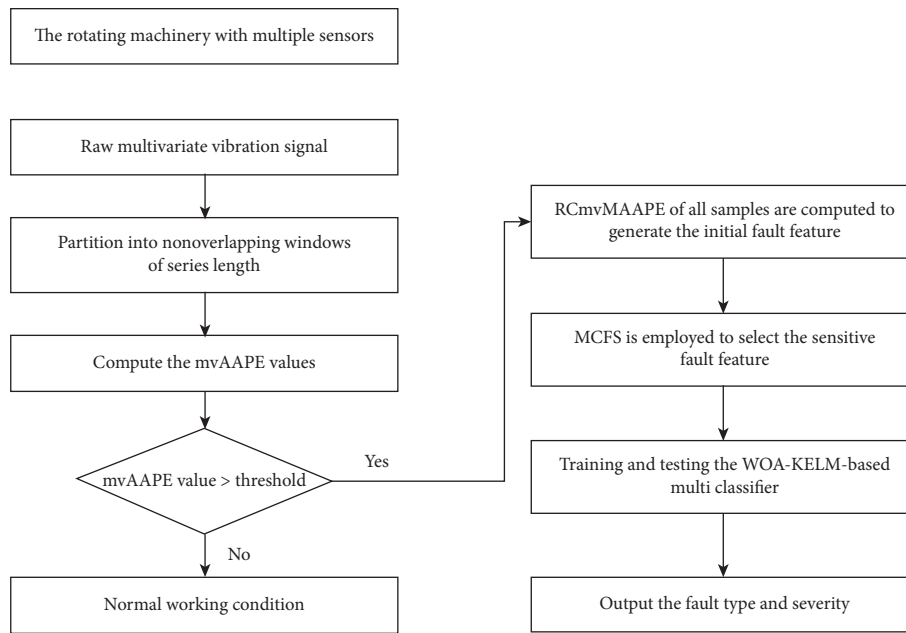


FIGURE 4: The flowchart of the proposed integrated health condition detection method.

5. Experimental Analysis and Results

In order to study the health condition detection method for rotating machinery raised in this paper to verify its universality and effectiveness for fault identification of general rotating machinery, experiments and analysis are conducted using two typical examples, namely, rolling bearings and gearboxes. The rolling bearing dataset was provided by CWRU [32]. The gearbox experiment data were collected on the QPZZ-II vibration analysis platform produced by Jiangsu Qianpeng Diagnostic Engineering Co., Ltd.

5.1. Health Condition Detection Experiment of Rolling Bearing

5.1.1. Experimental Rig and Data Introduction. The data were collected by the high-precision multichannel sensor installed on the bearing experimental rig. The specific structure of the bearing experimental rig is presented in Figure 5. The experimental rig includes a motor, a torque transducer/encoder, control electronics, and a dynamometer. The installation position of the acceleration sensors is at the 12 o'clock position at both the drive end and fan end of the motor housing, which are connected with the magnetic casing. The collected experimental data are the vibration waveforms of the motor, which are collected by the 16-channel data recorder. Single-point faults are set on SKF rolling bearings by electrical discharge machining. The fault diameter is 0.1778 mm, 0.3556 mm, and 0.5334 mm, respectively, and the fault depth is 0.2794 mm. The three fault diameters represent the different severity of the bearing fault. The experimental environment is set as follows: the motor load is 0 hp, the motor speed is 1797 r/min, and the sampling frequency is 12 kHz. In this article, the data used include 10 categories, normal bearings, inner race faults, outer race faults, and ball faults. The fault diameter of each fault state is

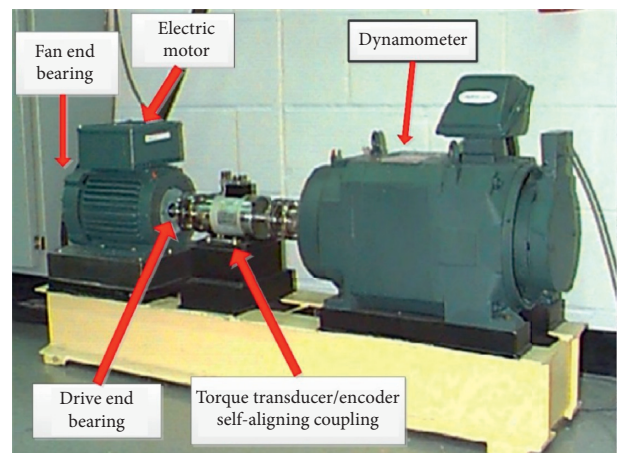


FIGURE 5: The rolling bearing test platform.

0.1778 mm, 0.3556 mm, and 0.5334 mm (label as NM, IRF1, IRF2, IRF3, ORF1, ORF2, ORF3, BF1, BF2, and BF3, respectively). For each fault state, the synchronous vibration signal at the drive end and fan end is used as dual-channel data. Generally, in the field of bearing fault diagnosis, the vibration signals are basically collected at the drive end. Since the data quality of the driver end is higher, which contains less noise and can directly reflect the vibration of the output part, however, for the fault diagnosis of mechanical equipment, high accuracy of fault identification is our goal. Therefore, it is necessary for us to use all available information to improve the utilization rate of information. The data of the fan end contains part of the fault information and the use of the data can significantly improve the characteristic quality, thus improving the fault recognition rate.

In this study, the vibration data of each working condition were divided into 58 samples without overlap, and the

number of sampling points of each sample was set to 2048. In order to be consistent with the engineering application under the actual condition, 28 samples for various working conditions are randomly selected for training, and the remaining 30 samples are the testing set. The effectiveness of the raised approach is validated by randomly selecting training and testing samples. The specific introduction of the dataset is presented in Table 1.

5.1.2. Fault Detection. The time domain waveforms of rolling bearing under ten working conditions are shown in Figure 6. Due to the lack of regularity, it is hard to directly recognize diverse working conditions based on their original vibration signals. According to previous analysis, PE has the ability to detect faults, the mvAAPE is obtained based on the theory of multidimensional embedding, and reconstruction also enjoys the same function. Therefore, mvAAPE can be used to detect whether the equipment is faulty. Figure 7 shows the mvAAPE values for all samples. As presented in Figure 7, the mvAAPE values in the fault states are generally large and the mvAAPE of the normal state is small, which is significantly different from the mvAAPE values of the fault states. Consequently, this method can be used to screen the normal state of the bearing. The value at the blue dotted line is defined as the mvAAPE threshold (2.9973). By comparing the mvAAPE value of the vibration signals with the threshold, the normal and fault states can be clearly distinguished. However, the samples of different fault types have poor separability, so mvAAPE cannot be used as the standard to judge the fault type and severity. A further analysis is needed to obtain more reliable characteristics.

The fault samples have the maximum mvAAPE value, which demonstrates that they are more complicated than normal samples. When the bearing is in normal operation, the vibration mainly comes from the interaction and coupling between the mechanical parts and the ambient noise, thereby the vibration signal shows certain regularity. Therefore, the mvAAPE value of normal condition is lower than that of the fault condition. When a fault occurs in the running process of the bearing, the vibration of the bearing will produce periodic pulse components. The high frequency vibration is mixed with the bearing vibration, which makes the frequency component and bandwidth of vibration signal more complex.

The first procedure in fault diagnosis is health detection. For a complicated mechanical system, it is necessary to judge whether there is a fault in the component firstly and then identify the type and severity of the fault. If the system does not detect the fault, it indicates that the system is running normally, and there is no need to disassemble and repair it.

5.1.3. Fault Recognition. Once a bearing fault is detected, the raised approach is used to distinguish the diverse fault types and severity. To validate the advantages of multivariate analysis, univariate analysis methods such as RCMAAPE are employed to test the bearing vibration signals at the drive end. By comparing with the univariate feature extraction method, the advantages of multichannel analysis in terms of

information utilization are intuitively verified. Each method uses data from 9 fault conditions for experiments. The entropy results of univariate analysis method RCMAAPE and multivariate analysis methods RCmvMAAPE, RCmvMPE, RCmvMSE, and mvMAAPE are shown in Figures 8(a)–8(e).

Compared with other multivariate analysis methods shown in Figures 5(b)–5(d), the entropy deviation of RCmvMAAPE is smaller and the stability is higher. First of all, when the scale factor is 5–16, RCmvMPE has poor discrimination of NM, IRF3, and ORF3. In addition, mvMAAPE is generally poorly distinguished, and the entropy deviation of each fault state is very large, which indicates its performance is unstable and easily causes large errors. Except for NM and ORF2, the RCmvMSE curves of the other states are similar on most scales, and the degree of overlap is high, making it difficult to distinguish them. For the other two univariate analysis methods, entropy deviation is significantly greater than that of the multivariate analysis method, and the degree of entropy curve overlap is also greater than that of the multivariate analysis method. This is mainly because the univariate analysis method only uses the vibration information of one channel, so the utilization rate of information is relatively low, while the multivariate analysis method realizes the effective use of information by comprehensively considering the vibration information of multiple channels, thus improving the stability and robustness of the analysis. Therefore, based on the above-mentioned analysis, RCmvMAAPE is more effective in feature extraction than RCmvMPE, RCmvMSE, mvMAAPE, and RCMAAPE, while the quality of the extracted features is also higher.

According to the abovementioned analysis, although the features extracted by the RCmvMAAPE method have high quality and can represent the fault state well, the fault features on the partial scale enjoy low separability and cannot achieve satisfactory distinguishing effect. For the sake of reducing the redundancy between features and enhancing the separability of fault features, the mRmR approach is utilized to reduce the dimension of original features. The distribution of multiscale features after the rearrangement is visually described in Figure 9. The dimensionality of the new multiscale fault features is selected as 9 according to the correlation with the main fault information and the importance of the features. Finally, the obtained new fault features are input into the WOA-KELM classifier to determine the fault type and severity. Figure 10 shows the failure classification results for one trial. It can be clearly observed from the figure that all the faults have been accurately identified and the classification accuracy has reached 100%, which indicates that the proposed approach can available distinguish the types and severity of faults.

In addition, for the sake of avoiding the influence of random factors such as contingency on the experimental results, 20 trials are repeated to obtain more accurate and reliable classification results. Moreover, four other entropy-based methods are also used to diagnose rolling bearing faults. The detailed classification results of the five approaches for 20 trials are presented in Figure 10 and Table 2.

TABLE 1: The detailed introduction of experiment sample.

Fault location	Fault diameter (mm)	Abbreviation	Training sample number	Testing sample number	Class label
Normal	0	NM	28	30	0
	0.1778	IRF1	28	30	1
	0.3556	IRF2	28	30	2
Inner race	0.5334	IRF3	28	30	3
	0.1778	ORF1	28	30	4
	0.3556	ORF2	28	30	5
Outer race	0.5334	ORF3	28	30	6
	0.1778	BF1	28	30	7
	0.3556	BF2	28	30	8
Ball	0.3556	BF2	28	30	8
	0.5334	BF3	28	30	9

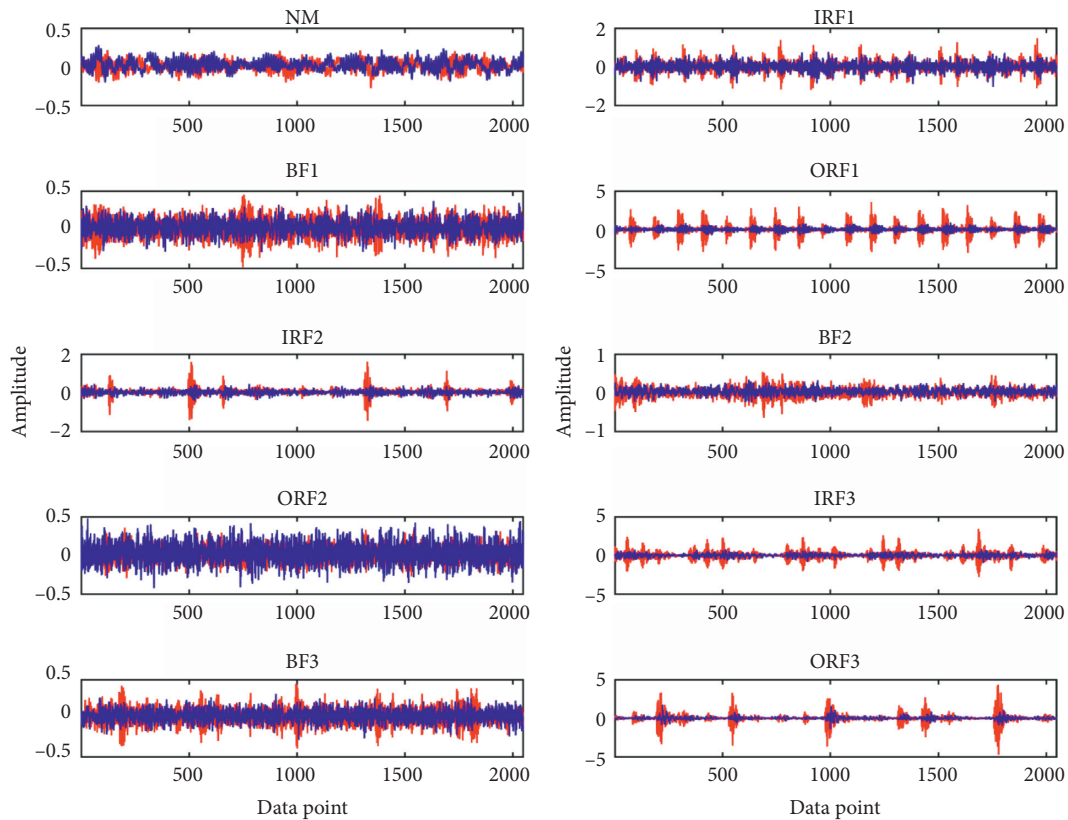


FIGURE 6: The waveforms of diverse classes of rolling bearing, where red denotes data of drive end and blue denotes fan end.

From Figure 11 and Table 2, it is obvious that the average classification accuracy of the raised approach is higher than that of other approaches, and the average accuracy rate is 99.96%. Moreover, the accuracy of the multivariate analysis methods (RCmvMAAPE, RCmvMPE, RCmvMSE, and mvMAAPE) is generally higher than that of the univariate analysis method (RCMAAPE), which is consistent with the previous analysis. Therefore, the comparison results indicate that the raised approach can effectively extract fault features and obtain high fault recognition rate.

To verify the necessity of mRmR feature selection, two-dimensional projections of two random features selected without adopting the mRmR method are presented in Figure 12(a), while the first two sensitive features obtained applying the mRmR method are visualized as Figure 12(b). By comparing Figures 12(a) and 12(b), it can be clearly

found that RCmvMAAPE combined with mRmR has a better recognition effect than using RCmvMAAPE alone. Moreover, nine random features ($\tau = 8, 19, 1, 17, 9, 3, 20, 14, 6$) are directly inputted into WOA-KELM to identify the fault type and the identification results are presented in Table 3. According to the results in Table 3, it can be clearly found that the fault recognition accuracy rate gained without using the mRmR method is lower than that gained with adopting the mRmR method. In addition, it can be noticed that the recognition accuracy of RCmvMAAPE is still higher than that of other methods without using mRmR. Thus, the experimental results again verify that RCmvMAAPE can extract fault features from multichannel signals effectively and improve the quality of fault information. The mRmR method can select sensitive low-dimensional features from high-dimensional fault

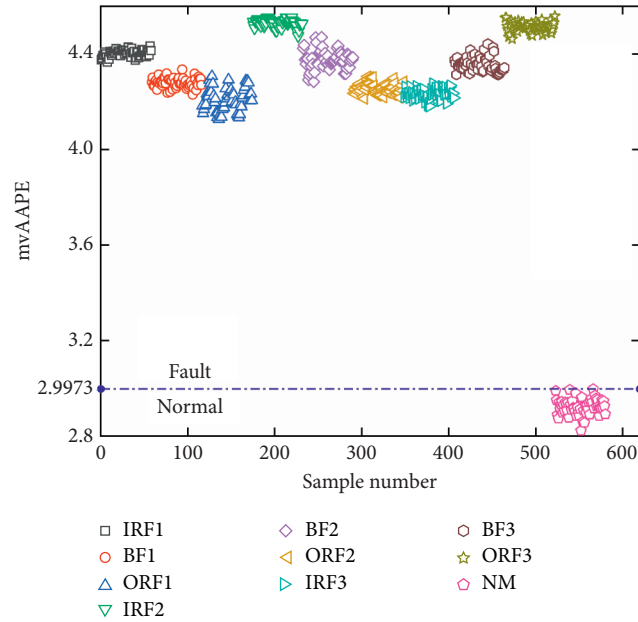


FIGURE 7: The multivariate amplitude-aware permutation entropy (mvAAPE) distribution of all samples.

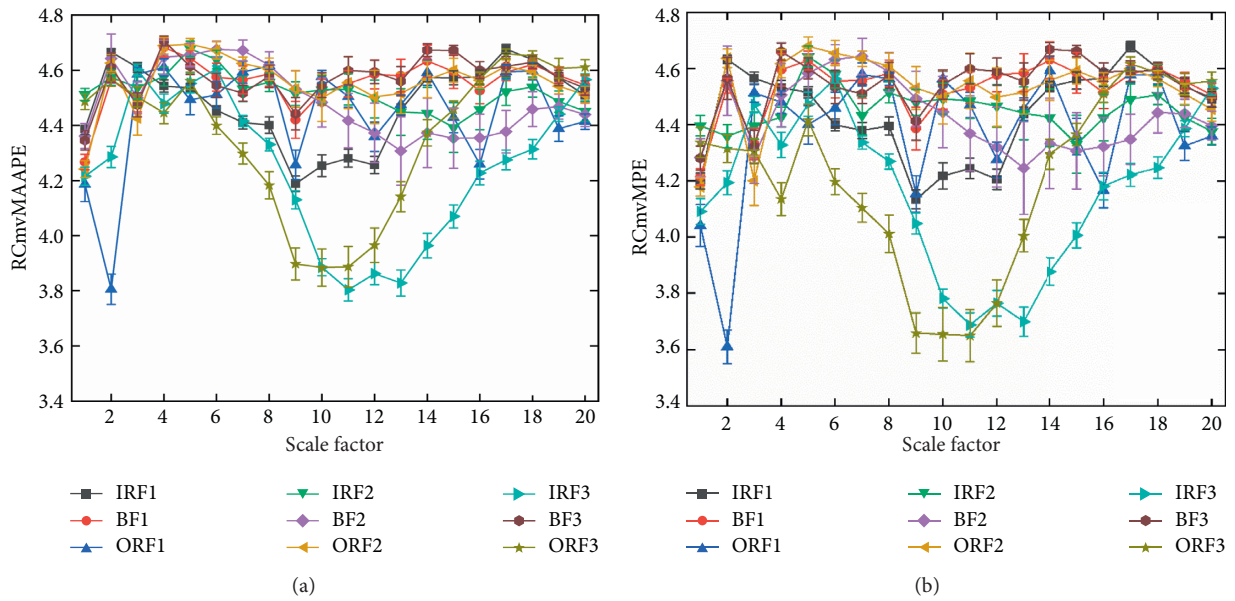


FIGURE 8: Continued.

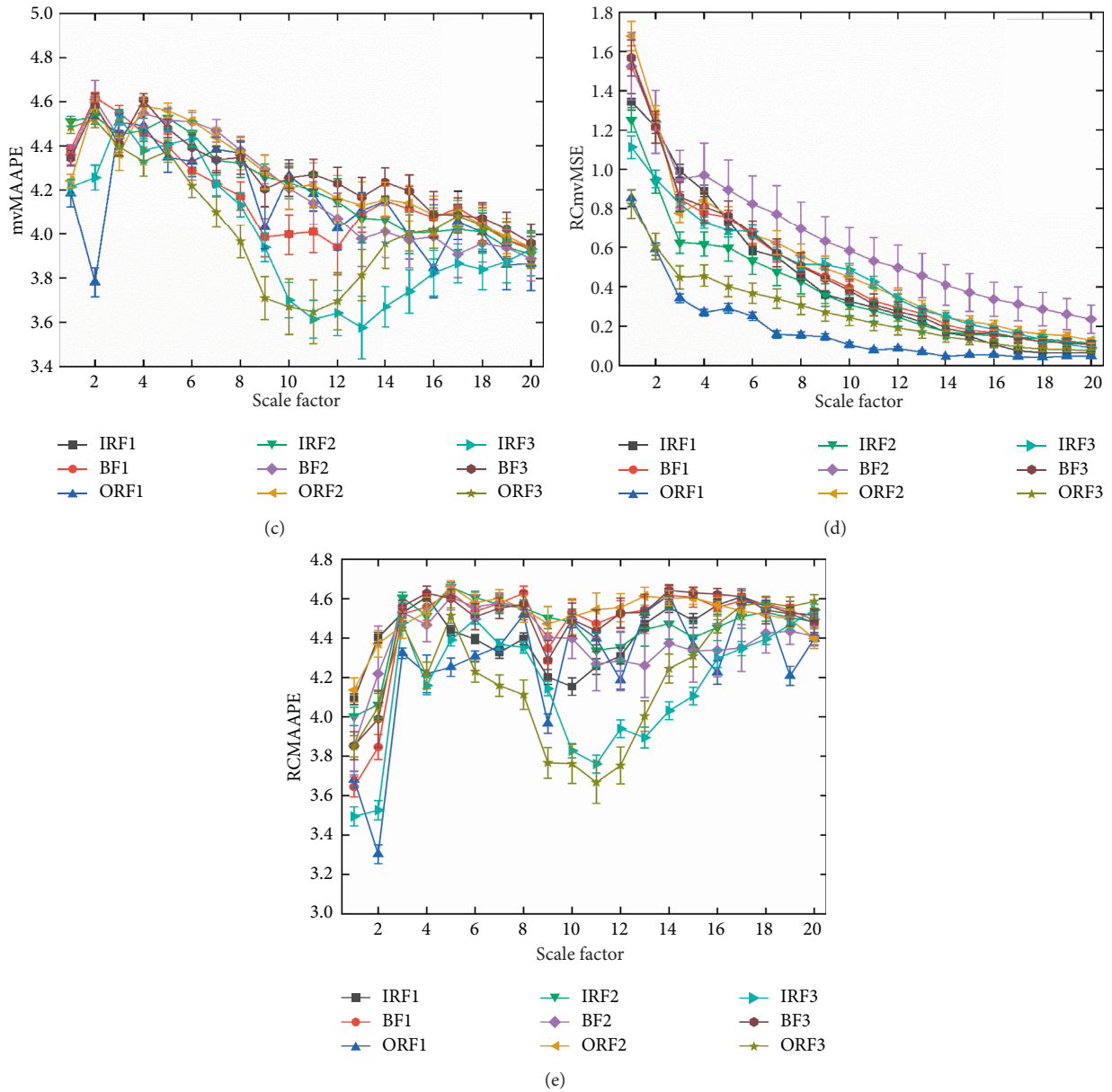


FIGURE 8: The entropy results of rolling bearing data analyzed by adopting five approaches. (a) RCmvMAAPE; (b) RCmvMPE; (c) mvMAAPE; (d) RCmvMSE; (e) RCMAAPE.

features, which not only improves the recognition accuracy but also improves the classification efficiency.

This section discusses the superiority of using WOA algorithm to optimize KELM in fault identification. For comparison, three commonly used classifiers are used for comparison, namely, support vector machine (SVM), extreme learning machine (ELM), and kernel extreme learning machine (KELM). The ratio of training samples to testing samples remains the same. The diagnostic results of the five approaches using diverse classifiers are listed in Table 4. It can be seen that when the four classifiers are combined with the five feature extraction methods, the classification accuracy of WOA-KELM is the highest, which shows that WOA-KELM is an effective classifier. In addition, it can be clearly found that when the features obtained by different

feature extraction methods are input to the four classifiers, the classification accuracy of RCmvMAAPE is the highest, which further verifies that the raised RCmvMAAPE approach has excellent performance in feature extraction.

5.2. Health Condition Detection Experiment of Gearbox

5.2.1. *Experimental Rig and Data Introduction.* The gearbox experiment data were collected from the experiment platform QPZZ-II that is built by Jiangsu Qianpeng Diagnosis Engineering Co., Ltd. The overall structure of the experimental platform is shown in Figure 13. The experimental platform is composed of gearbox, motor, iron base, capacitance, and sensors. The sensors are installed above the

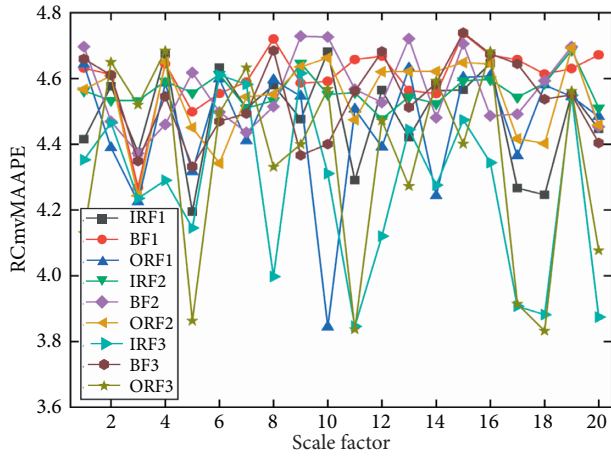


FIGURE 9: Distribution of multiscale feature after applying the mRmR approach.

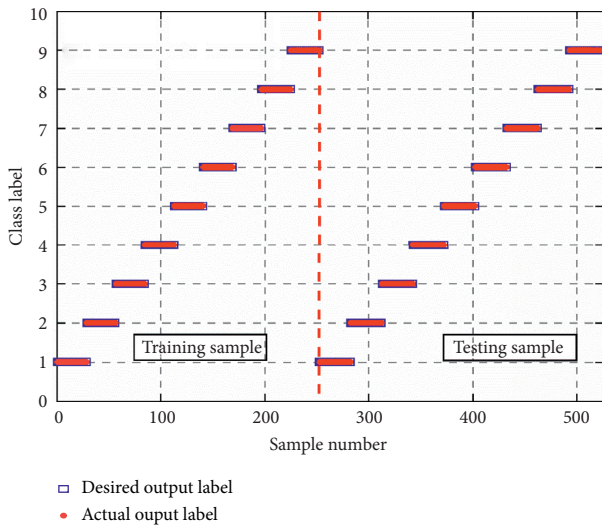


FIGURE 10: The recognition results of raised approach for rolling bearing.

TABLE 2: Identification result of five approaches for rolling bearings with mRmR feature selection.

Diverse approaches	Accuracy (%)			
	Max	Min	Mean	SD
The proposed method	100	99.26	99.96	0.1655
RCmvMPE and mRmR	100	97.41	98.54	0.8358
mvMAAPE and mRmR	91.11	86.30	88.69	1.2900
RCmvMSE and mRmR	95.56	92.22	93.92	1.0024
RCMAAPE and mRmR	90.37	85.56	87.52	1.4565

gearbox. The experimental data consist of eight channels of vibration signals and one channel of tachometer signals, in which the motor speed is 880 r/min. In the experiment, a total of four operating conditions were set up, including normal condition, gear pitting fault (pitting), gear tooth breaking (tooth breaking), pinion wear fault (wearing), and gear pitting fault coupling with pinion wear fault (pitting and wearing). The detailed introduction of gearbox

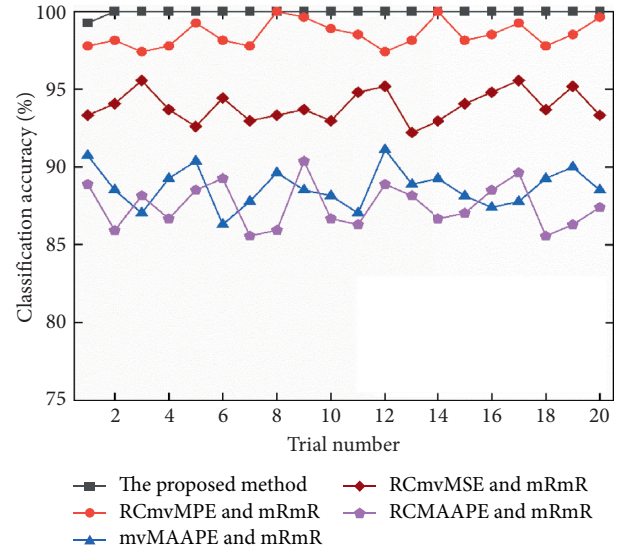


FIGURE 11: The diagnostic result of the five methods for 20 trials.

experimental data is shown in Table 5. The data acquisition equipment is QPZZ-II produced by Jiangsu Qianpeng Diagnostic Engineering Co., Ltd., with a sampling frequency of 5.12 kHz and sampling time of 6 s. Therefore, each health state contains 53248 data points. The selected channels are the acceleration signal collected by the bearing X on the motor side of the input shaft and the bearing Y on the load side of the output shaft. The collected vibration signals are divided into 26 nonoverlapping samples with length 2048. Among them, 10 samples were used for training, and the remaining 16 groups were used for testing.

5.2.2. Fault Detection. The time domain waveforms of the gearbox under four working conditions are shown in Figure 14. It is difficult to directly judge the type of gear failure based on the amplitude and frequency changes of the waveforms. According to the previous analysis, mvAAPE can be used to detect whether mechanical equipment is faulty and is successfully used to detect the health condition of rolling bearings. Due to the complicated structure of the gearbox, it is difficult to disassemble and inspect the gearbox. Therefore, it is necessary to detect the health condition of the gearbox. Figure 15 shows the mvAAPE values of all samples of the gearbox. It can be observed from the figure that all faulty samples have larger mvAAPE values, while all normal samples have smaller mvAAPE values. The value shown by the blue dashed line is defined as the mvAAPE threshold (4.2342). By comparing the mvAAPE value of the sample to be tested with the threshold, it can be judged whether the gearbox is faulty. However, the entropy values between different fault samples are relatively close, and the fault type cannot be judged intuitively. Therefore, the mvAAPE value cannot be used as a criterion for judging the fault type and further analysis is needed to obtain more obvious characteristics.

The fault samples have larger mvAAPE values, which indicates that the vibration signals of the fault samples are

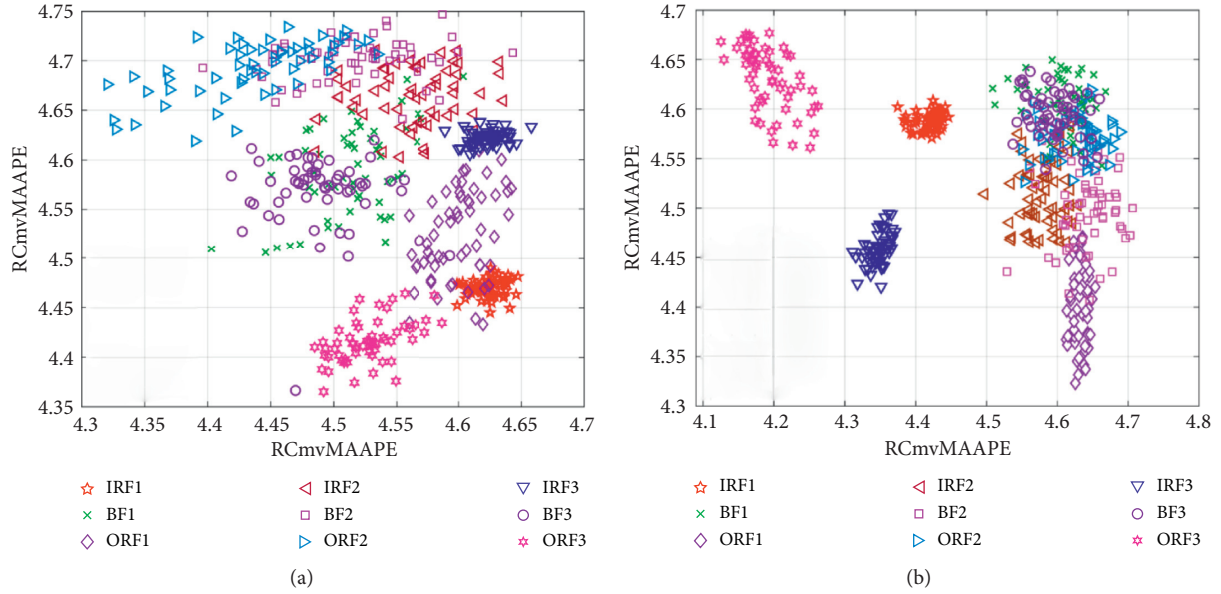


FIGURE 12: (a) Two-dimensional visualization of two random selected features without adopting mRmR. (b) Two-dimensional visualization of two new features selected utilizing mRmR.

TABLE 3: Identification result of five approaches without mRmR feature selection.

Methods	Accuracy (%)		
	Max	Min	Mean
RCmvMAAPE	94.44	89.26	92.32
RCmvMPE	91.85	88.15	90.69
mvMAAPE	87.04	84.44	85.86
RCmvMSE	90.37	87.78	89.63
RCMAAPE	86.30	81.85	84.27

more complicated than that of the normal samples. After the gearbox fails, the vibration signals enjoy obvious modulation characteristics, which are composed of multiple AM and FM signals. Compared with the vibration signals of the normal samples, the fault signals contain more impact components; meanwhile, due to the influence of random factors such as noise in the signal, the signal component is more complex, so it has a larger entropy value.

5.2.3. Fault Recognition. After detecting the gearbox failure, for the sake of identifying different fault types, the raised approach is utilized to process the fault vibration signals to obtain stronger features. Similarly, to verify the advantages of multivariate analysis, the univariate analysis method (RCMAAPE) is used for the motor side vibration signals. In addition, for the sake of studying the effectiveness of the RCmvMAAPE approach for extracting fault features, the RCmvMPE, mvMAAPE, and RCmvMSE approaches are used to analyze multichannel vibration signals. The analysis result is shown in Figures 16(a)–16(e).

It can be observed from Figure 16 that the overall trend of the RCmvMAAPE curve is consistent with that of RCmvMPE and mvMAAPE, but RCmvMAAPE has smaller entropy deviation, which indicates that the RCmvMAAPE method has

better stability. Compared with the RCmvMSE method, the RCmvMAAPE curve has more obvious fluctuation, so it can effectively highlight the earth oscillation component of gearbox fault vibration signal, so as to extract fault features more effectively. In addition, compared with the univariate analysis method RCMAAPE, the entropy deviation of RCmvMAAPE is significantly smaller, that is, its performance is better. The main reason is that the univariate analysis method only makes rough use of the fault information in the single channel vibration signal, while the rich information in other channels is not used reasonably. However, after gearbox fails, the transmission path of internal vibration is complex and has multiple directions. The vibration signals collected from each channel contain the fault information, so it is impossible to fully characterize the fault state only by performing univariate analysis. Based on the abovementioned analysis, RCmvMAAPE can effectively analyze multichannel vibration signals and has stable performance.

It can be observed from Figure 16 that the fault features extracted by RCmvMAAPE are redundant at some scales, which indicates that not all features can be used for fault classification. It is necessary to screen them to select sensitive features. In order to improve the separability of fault features, the mRmR approach is used to process the features. The distribution of multiscale features after the rearrangement is visually described in Figure 17. The dimensionality of the new multiscale fault features is selected as 9 according to the correlation with the main fault information and the importance of the features. Finally, the obtained new fault features $\tau = (19, 8, 7, 16, 5, 13, 10, 3, 2)$ are fed into the WOA-KELM classifier to determine the fault type. Figure 18 shows the fault classification results for one trial. It can be clearly observed from the figure that except two samples of pitting and wear fault are misclassified as tooth breaking fault, the other faults are accurately identified, and the

TABLE 4: The diagnostic results gained by combining diverse methods with four classifiers.

Diverse classifiers	The testing accuracy of classifiers with diverse approaches (%)					Average accuracy (%)
	RCmvMAAPE (%)	RCmvMPE (%)	mvMAAPE (%)	RCmvMSE (%)	RCMAAPE (%)	
ELM	97.04	95.93	88.52	94.44	85.19	92.22
SVM	95.56	94.07	87.04	92.96	83.33	90.59
KELM	98.52	96.30	89.26	94.81	85.92	92.96
WOA-KELM	100	99.26	91.48	95.93	88.89	95.11
Average accuracy (%)	97.78	96.39	89.07	94.54	85.83	—

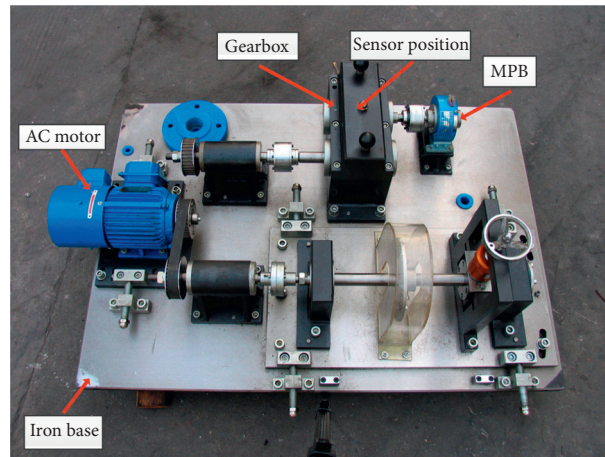


FIGURE 13: The experimental rig of the gearbox from QPZZ-II.

TABLE 5: The brief introduction of the experimental sample.

Fault type	Training sample number	Testing sample number	Class label
Normal	10	16	0
Wearing	10	16	1
Tooth breaking	10	16	2
Pitting and wearing	10	16	3

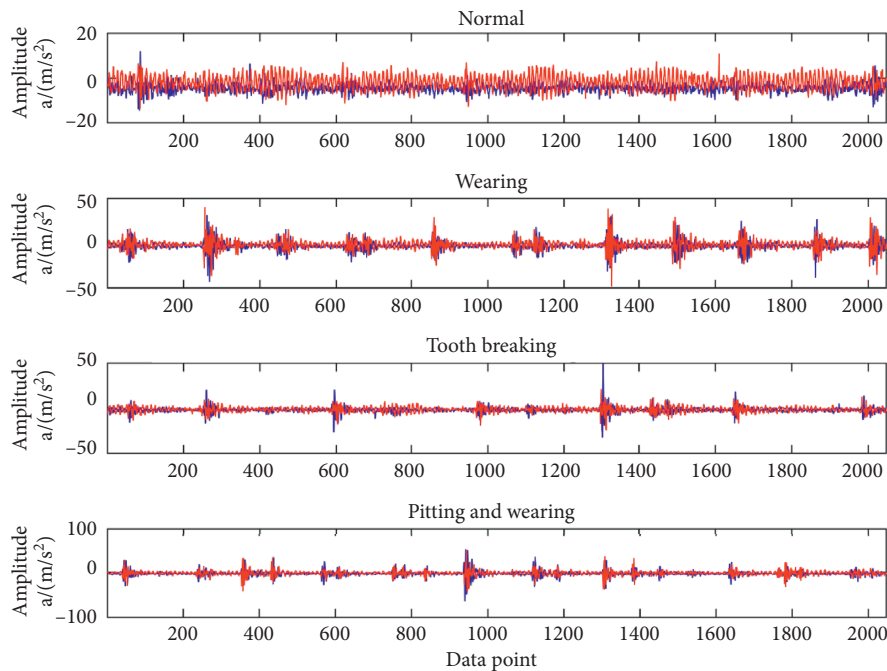


FIGURE 14: The vibration signal waveforms of the gearbox in different health conditions, where red denotes data of the motor side and blue denotes the load side.

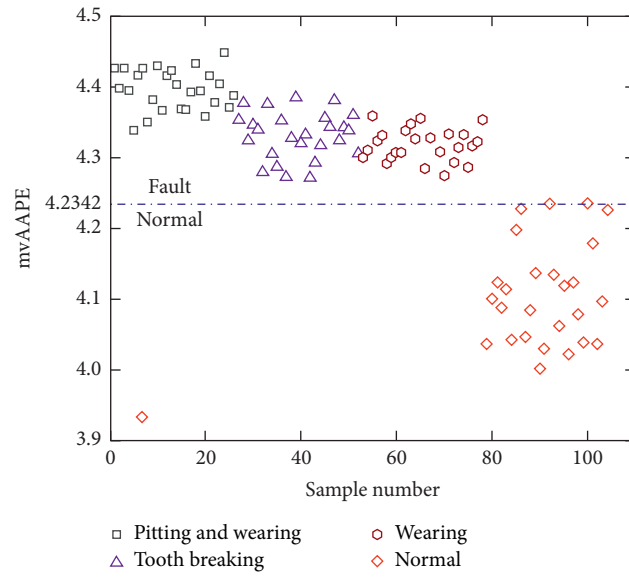


FIGURE 15: The multivariate amplitude-aware permutation entropy (mvAAPE) distribution of all samples.

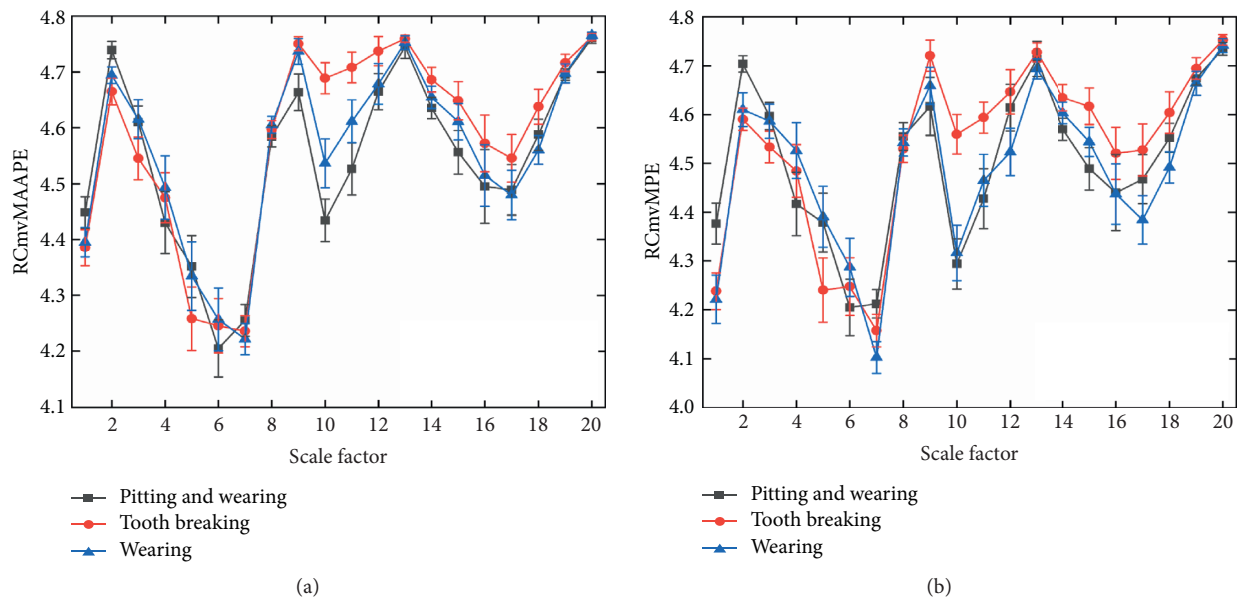


FIGURE 16: Continued.

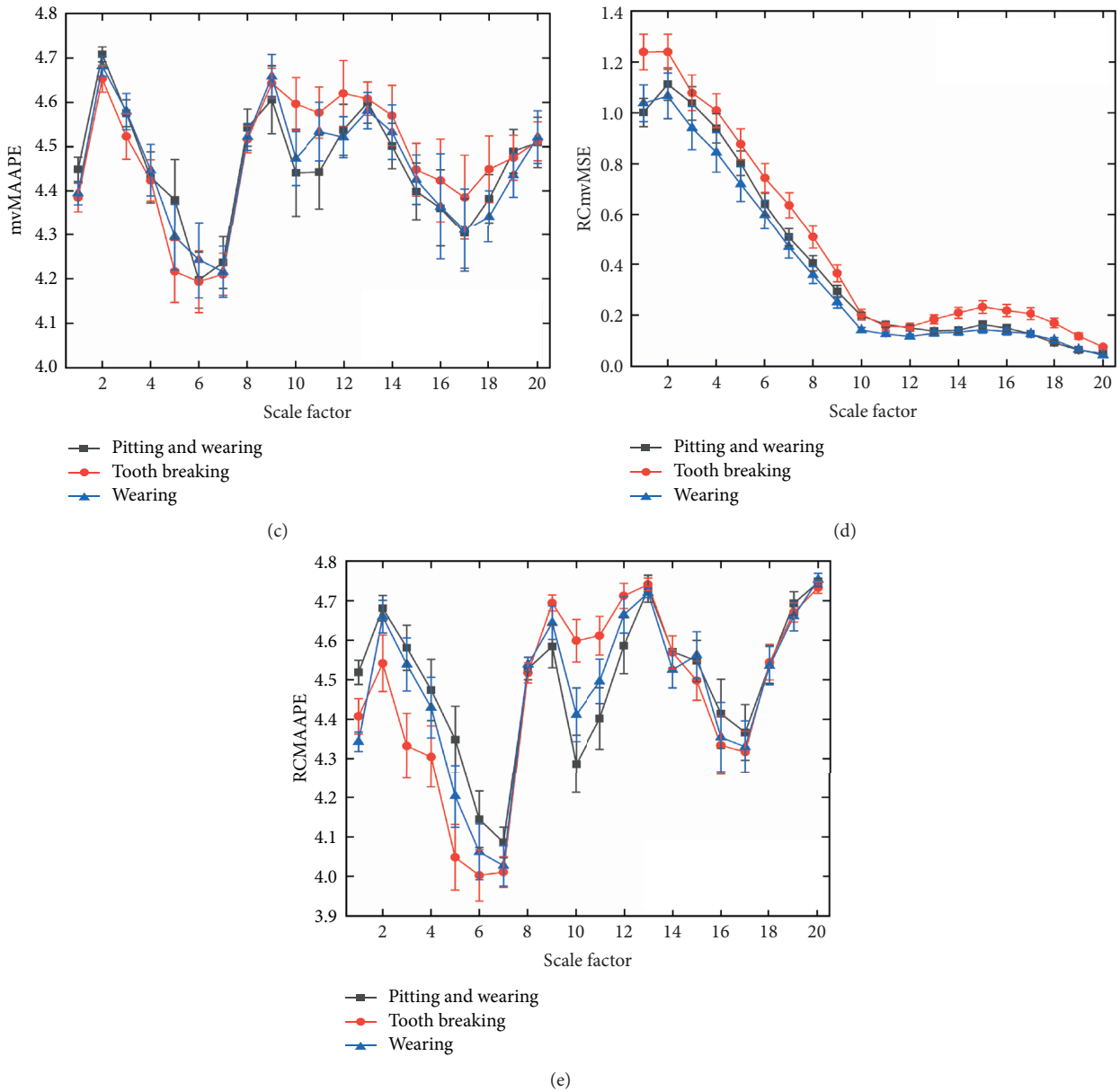


FIGURE 16: The entropy results of gearbox data analyzed by adopting five approaches. (a) RCmvMAAPE; (b) RCmvMPE; (c) mvMAAPE; (d) RCmvMSE; (e) RCMAAPE.

overall classification accuracy rate reaches 95.83%, which shows that the raised approach can available distinguish different fault types of gearbox.

Similarly, in order to reduce the large randomness of experimental results due to only performing one trial, 20 trials are repeated to obtain more reliable and accurate classification results. In addition, in order to intuitively verify the advantages of RCmvMAAPE method, four other entropy-based methods are used to diagnose gearbox faults. The detailed classification results of five approaches for 20 trials are shown in Figure 19 and Table 6. It is obvious from Table 7 that the average recognition accuracy of the presented approach is the highest and the standard deviation is the smallest, which indicates that the raised approach has

stable and excellent performance. The accuracy of RCmvMPE approach is slightly lower than that of the proposed approach, which indicates that RCmvMPE can also effectively diagnose gearbox faults. But the standard difference is large, indicating that the recognition rate is not stable. In addition, the accuracy of the multivariate analysis method is higher than that of the univariate analysis method, which verifies the necessity of multivariable analysis in gearbox fault diagnosis.

As before, for the sake of investigating the importance of mRmR feature selection, two-dimensional projections of two random features selected without adopting the mRmR method are presented in Figure 20(b), while the first two sensitive features obtained applying the mRmR approach are

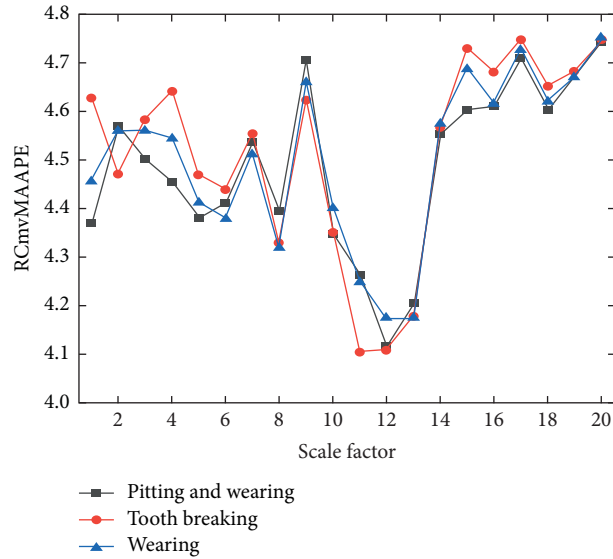


FIGURE 17: Distribution of multiscale feature after applying the mRmR approach.

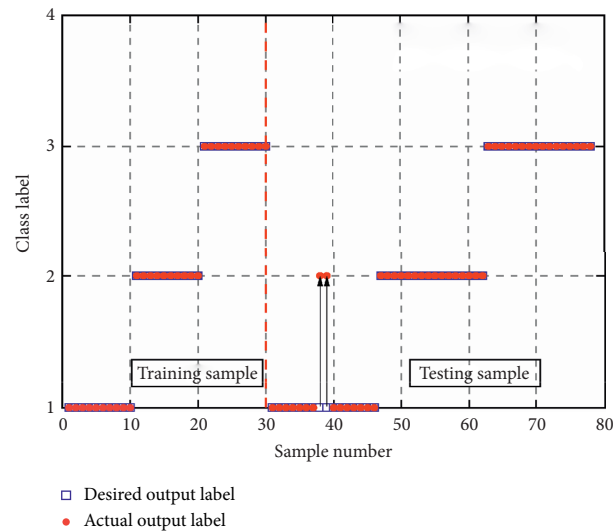


FIGURE 18: The recognition results of raised approach for gearbox.

visualized as Figure 20(a). It can be seen from the figure that the features without mRmR feature selection are disorderly and have no obvious clustering center, which indicates that the quality of features is not high and further processing is needed to obtain separable features. After mRmR feature selection, although no obvious clustering center is obtained, the separability of the three fault states becomes stronger. It can be concluded that mRmR feature selection can improve the recognition of features and has better recognition effect. Then, nine features are randomly selected and input into the WOA-KELM classifier to determine the fault type of gearbox. Similarly, each method was repeated 20 times. Table 7 shows the gearbox identification results of five methods without using mRmR feature selection for 20 trials. As can be seen from Table 7, although the highest recognition rate of the RCmvMAAPE approach is lower than that of the RCmvMPE method, the average recognition rate is

still the highest, which indicates that the performance of RCmvMAAPE is more stable. Consistent with the previous analysis, the recognition accuracy of the multivariate analysis approach is higher than that of the univariate analysis approach, which directly verifies the necessity of multivariate analysis. In a word, mRmR dimension reduction can significantly improve the fault recognition rate, that is, improve the reliability of fault identification.

To validate the necessity of utilizing WOA-KELM, three commonly used classifiers are used for comparison: SVM, ELM, and KELM. The same proportion of training and test samples is employed to train and test the classifier. Table 8 shows the classification results of five approaches using diverse classifiers. It can be seen that the RCmvMAAPE approach still has the highest fault recognition rate when using different classifiers, which is higher than that of the RCmvMPE method. Obviously, amplitude-aware

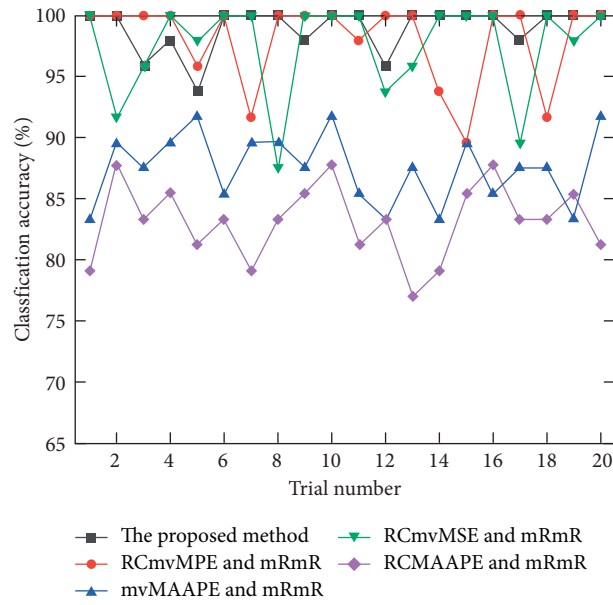


FIGURE 19: The diagnostic result of the five methods for 20 trials.

TABLE 6: Identification result of five approaches for gearbox with mRmR feature selection.

Diverse methods	Accuracy (%)			
	Max	Min	Mean	SD
The proposed method	100	93.75	98.96	1.8514
RCmvMPE and mRmR	100	89.58	98.02	3.4778
mvMAAPE and mRmR	91.67	83.33	87.50	2.8685
RCmvMSE and mRmR	100	87.50	97.5	3.9183
RCMAAPE and mRmR	87.80	77.08	83.17	3.0853

TABLE 7: Identification result of five approaches without mRmR feature selection.

Approaches	Accuracy (%)		
	Max	Min	Mean
RCmvMAAPE	89.58	85.42	87.22
RCmvMPE	91.67	83.33	86.94
mvMAAPE	81.25	72.92	78.46
RCmvMSE	87.5	81.25	84.32
RCMAAPE	75	66.67	72.53

TABLE 8: The diagnostic results gained by combining diverse methods with four classifiers.

Diverse classifiers	The testing accuracy of classifiers with diverse approaches (%)					Average accuracy (%)
	RCmvMAAPE (%)	RCmvMPE (%)	mvMAAPE (%)	RCmvMSE (%)	RCMAAPE (%)	
ELM	93.75	91.67	83.33	87.50	85.42	88.33
SVM	91.67	89.58	79.17	87.50	83.33	86.25
KELM	97.92	93.75	87.50	89.58	85.42	90.83
WOA-KELM	100	95.83	89.58	93.75	87.50	93.33
Average accuracy (%)	95.84	92.71	84.90	89.58	85.42	—

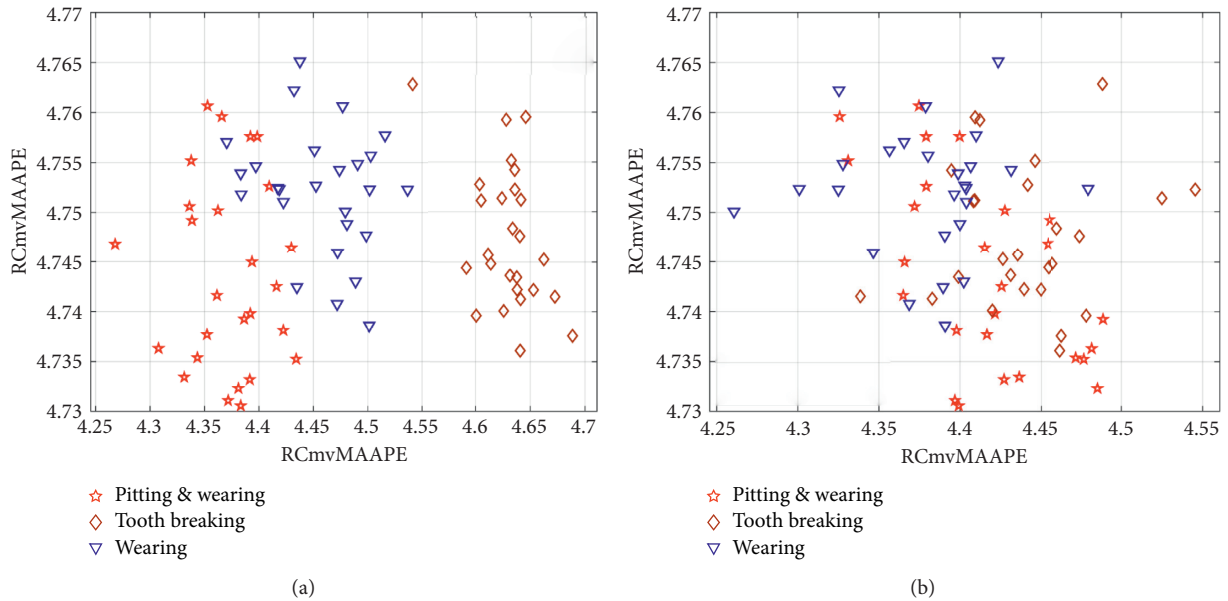


FIGURE 20: (a) Two dimensional visualization of two random selected features using mRMR. (b) two-dimensional visualization of two new features selected without using mRMR.

permutation entropy has better performance than permutation entropy by considering the amplitude and frequency information of time series. In addition, when the five methods are combined with different classifiers, the WOA-KELM classifier has the highest average recognition rate of 93.33%, which is higher than that of the KELM classifier alone. Since the performance of KELM is affected by the kernel parameters and penalty factor. The artificial setting cannot achieve the best classification effect. In conclusion, the WOA-KELM classifier has excellent performance, and the generalization performance is better than the commonly used classifiers.

6. Conclusions

In this study, a novel nonlinear analysis approach called RCmvMAAPE is raised. Various synthetic signals are analyzed and compared with RCmvMPE, mvMAAPE, and RCmvMSE. The results verify that RCmvMAAPE could effectively measure the complexity of multivariate time series and enjoys more stable performance. In the fault detection part, the mvAAPE is used to define a threshold. If the mvAAPE value of the measured sample is less than the threshold value, the equipment is normal, so as to realize the fault detection of the equipment. When a fault is detected, RCmvMAAPE is employed to extract fault features to construct initial feature vectors, and then mRmR is used to select sensitive features to form sensitive features to be classified. Finally, the sensitive feature vectors are input into the WOA-KELM classifier to determine the type and severity of the fault. The validity of the raised approach is verified by two typical examples, namely, rolling bearing and gearbox. The results demonstrate that the raised approach can not only accurately detect the fault of rotating machinery but also effectively identify the fault type. In addition,

compared with other methods, RCmvMAAPE can extract higher quality fault features from multichannel vibration signals and is superior to that of common entropy-based methods, which verifies its effectiveness in feature extraction. From the perspective of practical application, the proposed method avoids the mode classification that is full of uncertainty and improves the effectiveness and timeliness of fault diagnosis by detecting the state of rotating machinery, thereby is more in line with the actual engineering needs.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

- [1] Y. Lei, J. Lin, Z. He, and M. J. Zuo, "A review on empirical mode decomposition in fault diagnosis of rotating machinery," *Mechanical Systems and Signal Processing*, vol. 35, no. 1-2, pp. 108-126, 2013.
- [2] R. Liu, B. Yang, E. Zio, and X. Chen, "Artificial intelligence for fault diagnosis of rotating machinery: a review," *Mechanical Systems and Signal Processing*, vol. 108, pp. 33-47, 2018.
- [3] A. Rai and S. H. Upadhyay, "A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings," *Tribology International*, vol. 96, pp. 289-306, 2016.
- [4] P. Henriquez, J. B. Alonso, M. A. Ferrer, and C. M. Travieso, "Review of automatic fault diagnosis systems using audio and

- vibration signals,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 44, no. 5, pp. 642–652, 2013.
- [5] W. Deng, H. Liu, J. Xu, H. Zhao, and Y. Song, “An improved quantum-inspired differential evolution algorithm for deep belief network,” *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 10, pp. 7319–7327, 2020.
 - [6] Q. Gao, C. Duan, H. Fan, and Q. Meng, “Rotating machine fault diagnosis using empirical mode decomposition,” *Mechanical Systems and Signal Processing*, vol. 22, no. 5, pp. 1072–1081, 2008.
 - [7] J. Cheng, Y. Yang, and Y. Yang, “A rotating machinery fault diagnosis method based on local mean decomposition,” *Digital Signal Processing*, vol. 22, no. 2, pp. 356–366, 2012.
 - [8] Y. Chen, T. Zhang, Z. Luo, and K. Sun, “A novel rolling bearing fault diagnosis and severity analysis method,” *Applied Sciences*, vol. 9, no. 11, p. 2356, 2019.
 - [9] K. Dragomiretskiy and D. Zosso, “Variational mode decomposition,” *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, 2013.
 - [10] F. Zhou, X. Yang, J. Shen, and W. Liu, “Fault diagnosis of hydraulic pumps using PSO-VMD and refined composite multiscale fluctuation dispersion entropy,” *Shock and Vibration*, vol. 2020, Article ID 8840676, , 2020.
 - [11] Y. Li, X. Wang, Z. Liu, X. Liang, and S. Si, “The entropy algorithm and its variants in the fault diagnosis of rotating machinery: a review,” *IEEE Access*, vol. 6, pp. 66723–66741, 2018.
 - [12] R. Yan and R. X. Gao, “Approximate Entropy as a diagnostic tool for machine health monitoring,” *Mechanical Systems and Signal Processing*, vol. 21, no. 2, pp. 824–839, 2007.
 - [13] J. S. Richman and J. R. Moorman, “Physiological time-series analysis using approximate entropy and sample entropy,” *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 278, no. 6, pp. H2039–H2049, 2000.
 - [14] Z. Jinde, C. Junsheng, and Y. Yang, “A rolling bearing fault diagnosis approach based on LCD and fuzzy entropy,” *Mechanism and Machine Theory*, vol. 70, pp. 441–453, 2013.
 - [15] R. Yan, Y. Liu, and R. X. Gao, “Permutation entropy: a nonlinear statistical measure for status characterization of rotary machines,” *Mechanical Systems and Signal Processing*, vol. 29, pp. 474–484, 2012.
 - [16] X. Zhang, Y. Liang, J. Zhou, and Y. Zang, “A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM,” *Measurement*, vol. 69, pp. 164–179, 2015.
 - [17] K. Moshen, C. Gang, P. Yusong, and L. Yong, “Research of planetary gear fault diagnosis based on permutation entropy of CEEMDAN and ANFIS,” *Sensors*, vol. 18, no. 3, p. 782, 2018.
 - [18] H. Azami and J. Escudero, “Amplitude-aware permutation entropy: illustration in spike detection and signal segmentation,” *Computer Methods and Programs in Biomedicine*, vol. 128, pp. 40–51, 2016.
 - [19] M. Costa, A. L. Goldberger, and C.-K. Peng, “Multiscale entropy analysis of complex physiologic time series,” *Physical Review Letters*, vol. 89, no. 6, Article ID 068102, 2002.
 - [20] Y. Chen, T. Zhang, W. Zhao et al., “fault diagnosis of rolling bearing using multiscale Amplitude-aware permutation entropy and random forest,” *Algorithms*, vol. 12, no. 9, 2019.
 - [21] X. Wang, S. Si, Y. Li, and X. Du, “An integrated method based on refined composite multivariate hierarchical permutation entropy and random forest and its application in rotating machinery,” *Journal of Vibration and Control*, vol. 26, no. 3-4, pp. 146–160, 2020.
 - [22] H. Azami and J. Escudero, “Refined composite multivariate generalized multiscale fuzzy entropy: a tool for complexity analysis of multichannel signals,” *Physica A: Statistical Mechanics and its Applications*, vol. 465, pp. 261–276, 2017.
 - [23] M. U. Ahmed and D. P. Mandic, “Multivariate multiscale entropy: a tool for complexity analysis of multichannel data,” *Physical Review E*, vol. 84, no. 6, Article ID 61918, 2011.
 - [24] H. Peng, F. Long, and C. Ding, “Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1226–1238, 2005.
 - [25] G. B. Huang, “An insight into extreme learning machines: random neurons, random features and kernels,” *Cognitive Computation*, vol. 6, no. 3, pp. 1–15, 2014.
 - [26] Á. Arnaiz-González, A. Fernández-Valdivielso, A. Bustillo, and L. N. López de Lacalle, “Using artificial neural networks for the prediction of dimensional error on inclined surfaces manufactured by ball-end milling,” *The International Journal of Advanced Manufacturing Technology*, vol. 83, no. 5–8, pp. 847–859, 2016.
 - [27] X. Liu and J. Tang, “Mass classification in mammograms using selected geometry and texture features, and a new SVM-based feature selection method,” *IEEE Systems Journal*, vol. 8, no. 3, pp. 910–920, 2014.
 - [28] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, “Extreme learning machine: theory and applications,” *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, 2006.
 - [29] J. Kennedy and R. Eberhart, “Particle swarm optimization,” in *Proceedings of ICNN’95-International Conference on Neural Networks*, Perth, Australia, August 1995.
 - [30] W. Deng, J. Xu, and H. Zhao, “An improved ant colony optimization algorithm based on hybrid strategies for scheduling problem,” *IEEE Access*, vol. 7, pp. 20281–20292, 2019.
 - [31] S. Mirjalili and A. Lewis, “The whale optimization algorithm,” *Advances in Engineering Software*, vol. 95, no. 95, pp. 51–67, 2016.
 - [32] Case Western Reserve University, “Bearing data center,” 2019, <http://csegroups.case.edu/%20bearingdatacenter/pages/download-data-fifile>.