

Received February 3, 2021, accepted February 28, 2021, date of publication March 2, 2021, date of current version March 31, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3063555

Complementary Ensemble Adaptive Local Iterative Filtering and Its Application to Rolling Bearing Fault Diagnosis

YI ZHANG¹, YONG LV¹, AND MAO GE¹

Key Laboratory of Metallurgical Equipment and Control Technology, Ministry of Education, Wuhan University of Science and Technology, Wuhan 430081, China
Hubei Key Laboratory of Mechanical Transmission and Manufacturing Engineering, Wuhan University of Science and Technology, Wuhan 430081, China

Corresponding author: Yong Lv (lv Yong@wust.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 51875416, and in part by the Natural Science Foundation Innovation Group Program of Hubei Province under Grant 2020CFA033.

ABSTRACT A fault diagnosis model for rolling bearing based on complementary ensemble adaptive local iterative filtering (CEALIF), Laplacian score (LS) feature selection, and genetic algorithm-based backpropagation neural network (GA-BPNN) is proposed in this article. When the rolling bearing fails, the field-measured waveforms usually shows strong nonlinearity and non-stationary characteristics. Adaptive Local Iterative Filtering (ALIF) is an alternative novel approach to empirical mode decomposition to decompose complex signals into multiple intrinsic mode functions (IMFs), but modal aliasing will occur in actual processing. Aiming at the phenomenon of modal aliasing, a noise-assisted analysis methodology, namely complementary ensemble adaptive local iterative filtering, which could overcome the modal aliasing problem of ALIF. This article applies it to the pre-processing of rolling bearing time series. Then, the time domain (TD) statistical features of the IMFs, their Fourier frequency domain (FD) features, and the time frequency domain (TFD) energy features are extracted to capture the fault information. Meanwhile, to avoid feature redundancy and enhance the diagnostic performance, the LS is adopted to rank the features to improve the fault characteristics. Subsequently, the optimized feature vectors are entered into the GA-BPNN to automatically achieve the fault type recognition. The experimental data analysis results of rolling bearings indicate that the model can effectively diagnose the degree and type of failure.

INDEX TERMS Complementary ensemble adaptive local iterative filtering (CEALIF), Laplacian score (LS), feature selection, genetic algorithm-based BP neural network, fault diagnosis.

I. INTRODUCTION

The rolling bearing is an integral component of the rotating machinery and its failure would directly affect the functioning of the machinery as a whole [1], [2]. Therefore, it has become increasingly important to improve bearing reliability and detect bearing failure in timely and accurately [3]. When the rolling bearing fails, it will produce periodic pulse impact force, causing the system's nonlinear vibration, which makes the collected vibration signal always has nonlinear nonstationary features, so the key to its fault diagnosis is how to derive failure characteristic information from nonlinear, nonstationary signal [4], [5].

The associate editor coordinating the review of this manuscript and approving it for publication was Lei Shu¹.

The traditional signal analysis approaches are often only applicable to the analysis of smooth signals, while the rolling bearing fault signal has certain limitations [6]. For rolling bearing fault signals, the signal analysis processing technology must be adaptive and stable to ensure the effective fault signal characteristic information can be obtained in the face of complex fault signals [7]. In the past few decades, vibration-based analysis approaches have been well developed in fault diagnosis, such as short-time Fourier transform (STFT); Wavelet transform (WT); Empirical mode decomposition (EMD), and so on. STFT [8], [9] employs a moving window function to intercept non-stationary signals, which will result in STFT's time resolution and frequency resolution not being optimal at the same time, and when the frequency changes rapidly, the analysis effect is not good.

WT [10], [11] has been well employed in mechanical fault diagnosis, but how to choose a wavelet basis function with better time-frequency resolution is not yet clear, which needs to be determined manually. EMD [12] is an adaptive non-stationary signal analysis technique developed by Huang *et al.*, which achieves adaptive decomposition by seeking the envelope average curve of the time series to be handled. However, this method also has many problems, such as under-envelope and over-envelope phenomena when seeking the extreme envelope, and serious modal aliasing and end-point effects will occur in the face of certain fault signals. The decomposition results are distorted or even completely invalid, which leads to the misunderstanding of signal processing analysis, which affects the accuracy of fault diagnosis. At the same time, this method has been lacking rigorous and effective mathematical proof in theory. To solve the EMD modal aliasing problem, Wu and Huang [13] suggested ensemble EMD (EEMD) utilizing the statistical properties of white noise. Li *et al.* combined EEMD with improved frequency band entropy to derive bearing failure characteristics [14]. However, when the number of integrations is not large enough, the residual noise in the IMF will increase. Therefore, Yeh *et al.* [15] proposed complementary EEMD (CEEMD) to reduce the residual noise by adding a pair of white noise to the analysis signal. However, CEEMD cannot solve problems such as different IMF orders due to the addition of different noises. So Torres *et al.* [16] proposed complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN), which can avoid the problem of generating different IMF orders due to the addition of different noises. Cheng *et al.* successfully employed CEEMDAN for rolling bearing fault diagnosis [17]. Although CEEMDAN overcomes the aforementioned problems of EEMD, it is still essentially not free from the lack of theoretical support for EMD.

Lin *et al.* improved the EMD methodology and proposed the Iterative filtering (IF) [18] algorithm, which employs a low-pass filter function instead of EMD to solve the envelope average curve in EMD to solve the issue of lack of theoretical support for EMD and its derivatives methods, and proved the convergence of the algorithm [19]. Dash *et al.* applied IF to the feature extraction of EEG signals and achieved positive results [20]–[22]. When the signal to be processed contains strong nonlinear and non-stationary components, a fixed low-pass filter function may cause problems such as distortion of the decomposition results and poor adaptive behavior. To improve the adaptive and accurate decomposition of the method, Cicone *et al.* [23]–[25] propose a novel signal processing technology, the Adaptive Local Iterative filtering (ALIF) algorithm, by reworking traditional filters to ensure that the IF method can be accurately decomposed when applied to non-stationary signals. It is derived from the EMD method and developed from the IF method. It also uses iterative decomposition technology. The difference is that it uses an adaptive filter with a filter length driven by data to achieve the necessary adaptive decomposition. The solution

of the Planck equation constitutes this method, which avoids some defects in the EMD method. The ALIF decomposition method is not only highly adaptive, it can accurately process AM-FM signals containing multiple components and decompose it into several stationary signals [26]–[29]. At the same time, the method has a stronger ability to suppress modal aliasing and is still faced with complex noise. There is a good performance. In addition, the method already has rigorous mathematical proofs, and its decomposed components have proved to be convergent [25]. Therefore, it can be introduced into the discipline of mechanical fault diagnosis.

Since the ALIF is obtained by solving the Fokker-Planck equation and the filtering interval based on extreme-scale drive, the filtering interval of the filter is adaptively adjusted according to the extreme-scale during the iterative filtering process [23]. It was found that when the signal to be processed contains high-frequency intermittent signals, the extreme value of the signal to be decomposed will change suddenly, which affects the ability of the filter to distinguish different modes, resulting in the decomposition result distortion [13], [15]. Therefore, to improve the decomposition effect of ALIF, inspired by the idea of EMD modal aliasing, a complementary ensemble adaptive local iterative filtering (CEALIF) algorithm based on noise-assisted analysis is presented. When the ALIF method is employed to process the signal, a pair of white noise with the same amplitude and opposite sign is added, which allows the different proportional components of the original signal to fill the time-frequency space evenly over the entire time range. The noise addition results in a more uniform distribution of extreme points, which improves the performance of the adaptive filter and provides better resolution to different modes. Each result is averaged over a finite number of ensembles to remove the noise effect and obtain the final result. The CEALIF algorithm is employed to the pre-processing of the vibration signal of bearing failure, which decomposes it into the sum of several IMFs, and subsequently extract the IMF characteristic parameters as well as the time frequency domain information.

After extracting the above rolling bearing fault characteristics, to avoid information redundancy due to the high dimensionality of the feature vector and reduce the efficiency of diagnosis, this article adopts Laplace score (LS) to re-rank the features [30], and select the most relevant features as the feature vectors to characterize the failure indicators, thus reducing the dimensionality of feature vectors and improving the diagnostic efficiency.

Naturally, after obtaining the feature vectors, a multiple-fault categorizer is adopted to automatically identify the failure category and the degree of damage of the rolling bearing [3]–[33]. Some common classifiers such as support vector machines (SVM) [34], [35] and Backpropagation neural network (BPNN) have been widely used in machinery and equipment diagnosis. However, SVM requires prior selection of kernel function and its parameters, which is essentially binary classification, and multiple binary classification is required for multi-classification problems. In order to reduce

the reliance on human experience and achieve automatic diagnosis, a suitable multi-classifier needs to be selected. BPNN [36]–[38] does not require the introduction of new parameters in the training process, which relies entirely on the adjustment of initial weights and thresholds. To avoid the local optimization problem of random initialization and to enhance the performance of BPNN, this article adopts genetic algorithm (GA) [39], [40] to find the optimal initial weights and thresholds of BPNN and applies them to construct predictive models. The model developed through this approach is more stable and can converge faster than the standard BPNN.

Based on the above analysis, the paper proposes a rolling bearing fault diagnosis model via CEALIF, LS, and GA-BPNN. Firstly, CEALIF is utilized to handle the vibration signal of rolling bearing and several IMFs are obtained.; secondly, the first few IMFs containing the main fault information are selected and their TD and FD fault parameters are extracted respectively, and then the TFD feature values of the vibration signal are extracted to form the initial feature vector; thirdly, the LS is adopted for each feature of the initial feature vector according to its importance, and the first few more important features are selected as feature vectors according to the order of scores from smallest to largest; finally, the feature vectors are entered into the BPNN to automatically diagnose the type and degree of rolling bearing failures. Public datasets as well as experimental data were employed to validate the efficacy of the model.

The rest of the organizational structure is given below: Section 2 introduces the ALIF method and the basic principles of CEALIF. Section 3 uses simulated signals to test the performance of CEALIF and compares it with EEMD and CEEMDAN. Section 4 gives an introduction to the various feature quantities to be extracted and Laplace scores. Section 5 validates the proposed model and compares it with several similar methods. Conclusions are given in Section 6.

II. THEORETICAL DESCRIPTION

A. INTRINSIC MODE COMPONENTS AND THEIR FILTERING

The essence of empirical mode decomposition algorithm is actually a filtering algorithm from high frequency to low frequency, and similar algorithms such as variational mode decomposition (VMD), empirical wavelet transform (EWT) and stationary wavelet transform (SWT) have been developed. The EMD algorithm obtains IMF step-by-step and the modal components satisfy the following conditions:

$$\begin{cases} c_1(t) = \lim_{n \rightarrow \infty} \Gamma_1^n(s(t)) \\ c_k(t) = \lim_{n \rightarrow \infty} \Gamma_k^n(s(t) - \sum_{i=1}^{k-1} c_i(t)) \end{cases} \quad (1)$$

where $s(t)$ is a time series to be processed, Γ is a wave operator (also known as the fluctuant operator), and ζ is a sliding operator in $\Gamma = s(t) - \zeta(s(t))$.

Further, the following is the process of IMF:

Step 1. The sliding operator $\zeta(s(t))$ is calculated for a given signal $s(t)$.

Step 2. The time series $s(t)$ is subtracted from the sliding operator $\zeta(s(t))$ to obtain the wave operator $\Gamma = s(t) - \zeta(s(t))$.

Step 3. If the fluctuant operator Γ satisfies Equation (1), then the first IMF is $c_1(t) = \Gamma$. If the condition is not satisfied, then the next process to obtain IMF component will continue to Steps 1-2 until Equation (1) is satisfied.

Step 4. Subtract the signal $s(t)$ from the preceding m IMF components and obtain the residual components

$$r(t) = s(t) - \sum_{i=1}^m c_i(t).$$

where $r(t)$ is regarded as the previous $s(t)$, and it is looped Steps 1-3 until $r(t)$ becomes a trend component. After the loop stops, the original time series can be expressed by the following equation:

$$s(t) = \sum_{i=1}^N c_i(t) + r(t) \quad (2)$$

The sliding operator ζ represents different decomposition methods if its computational method is different. In EMD algorithm, the sliding operator is computed by taking the average value of the envelopes as follows:

$$\zeta(s(t)) = \frac{1}{2} (E_1(t) + E_2(t)) \quad (3)$$

where the upper envelope is E_1 and the lower envelope is E_2 .

B. THEORETICAL DESCRIPTION OF ADAPTIVE LOCAL ITERATIVE FILTERING (ALIF)

Under the framework of EMD algorithm, researchers obtained the sliding operator by convoluting the low-pass filter with the original signal, and thus obtained another algorithm called iterative filtering algorithm (IF) [18], [19].

In the IF algorithm, the convolution of signal $s(t)$ and the filtering function $h(t)$ is regarded as a sliding operator as follows:

$$\zeta(s(t)) = \int_{-L(n)}^{L(n)} s(t+l)h(l,t)dl \quad (4)$$

where $L(n)$ is filter length, which is calculated as:

$$L(n) = 2 \lceil \frac{\eta K}{t} \rceil \quad (5)$$

where the steady-state coefficient η is 1.6, t denotes the number of extreme points of the sequence, K is the length of signal, and $\lceil \cdot \rceil$ is rounded to zero.

Due to the limitation of operation time, it is difficult to realize the situation that n tends to infinity in Equation (1). Considering the actual situation, the features of IMF can be described as follows:

$$\begin{cases} |(e_{\min} + e_{\max}) - e_0| \leq 1 \\ \frac{e_1(t) + e_2(t)}{2} = 0 \end{cases} \quad (6)$$

where e_{\min} , e_{\max} and e_0 are the minimum, the maximum and zero values of the IMF respectively, $e_1(t)$ denotes the envelope function formed by IMF's local maximum and $e_2(t)$ denotes the envelope function formed by IMF's local minimum.

Equation (7) can be employed to determine whether the component satisfies the IMF condition:

$$SD = \frac{\|\Gamma_{i,n} - \Gamma_{i,n-1}\|_2}{\|\Gamma_{i,n-1}\|_2} \quad (7)$$

whenever the SD value reaches the set threshold, iteratively stop to obtain an IMF component.

It is worth mentioning that in the iterative filtering technique, the filter $h(t)$ is a general filter, and the filter interval $L(n)$ is also a fixed value. In order to improve the adaptability and accuracy of the decomposition of this method, the traditional iterative filtering technique is improved to ensure that it can be applied to the decomposition of non-stationary signals. In ALIF, the filter $h(t)$ is given by the solution of the Fokker-Planck equation, and the filter interval $L(n)$ is also adaptive according to the change trend of the signal to be decomposed [23]. The Fokker-Planck equation is as follows:

$$\frac{\partial p}{\partial t} = -\alpha \frac{\partial(k(x)p)}{\partial x} + \beta \frac{\partial^2(f^2(x)p)}{\partial x^2} \quad (8)$$

where α and β are constants in (0,1); and $f(x)$ and $k(x)$ are two smooth differentiable functions; and $f(a) = f(b) = 0$ is satisfied in (a, b) (where $a < 0 < b$); and $k(a) < 0 < k(b)$, $f(t) > 0$ for any $x \in (a, b)$.

Among them, $\frac{\partial^2(f^2(x)p)}{\partial x^2}$ will cause a diffusion effect to make the solution of the equation move to both ends; meanwhile, $-\alpha \frac{\partial(k(x)p)}{\partial x}$ will move the solution $p(x)$ from both ends of (a, b) to the center. When the two reach equilibrium, there are as follows:

$$-\alpha \frac{\partial(k(x)p)}{\partial x} + \beta \frac{\partial^2(f^2(x)p)}{\partial x^2} = 0 \quad (9)$$

At this point, the solution to the equation (9) satisfies the condition: $\forall x \in (a, b), p(x) > 0$ and in other case $p(x) = 0$.

This shows that the solutions $p(x)$ of the above steady-state equations are all between (a, b) , and then $p(x)$ can be regarded as a filter function. With the change of the filtering interval, the filtering function also changes, so as to realize the adaptive decomposition of the ALIF algorithm.

C. RESTRAIN OF MODE MIXING AND COMPLEMENTARY ENSEMBLE ADAPTIVE LOCAL ITERATIVE FILTERING (CEALIF)

Mode mixing is the disadvantage of most iterative screening algorithms. ALIF's mode mixing is defined as the existence of multiple scales in a single modality or the existence of the same scale information in different modalities. In this article, EMD's suppressing mode mixing strategy is employed to solve the mode mixing issue of ALIF. The study found that the modal aliasing of ALIF mainly comes from the discontinuity of the time scale, which is the influence of

signal intermittent. When there is an intermittent signal in the signal to be analyzed, there is no extreme value related to the intermittent signal in certain intervals, so the extreme points of other scales supplement the missing extreme points of the intermittent signal and some of them will remain on the scale of the intermittent signal in the end, the result will cause the performance of ALIF's filter bank to be compromised, which is commonly referred to as modal aliasing. The white noise is evenly distributed throughout the time-frequency space. The original signal will be projected to the corresponding scale when white noise is added [41]–[43]. Therefore, the ALIF filter set can be restored by adding white noise. Performance. The addition of noise will include additional noise components in the decomposition results. In this article, white noise of opposite sign is added enough times to eliminate the impacts of the noise. The flowchart of the CEALIF algorithm is depicted in Fig. 1:

Step 1: Let the signal to be decomposed be $s(t)$ and add white noise of opposite sign to $s(t)$ respectively.

$$\begin{bmatrix} s_i^+(t) \\ s_i^-(t) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} s(t) \\ n_i(t) \end{bmatrix} \quad (10)$$

where $n_i(t)$, $i = 1, 2, \dots, Ne$ represents the added white noise signal, Ne represents the number of additions.

Step 2: Use the ALIF technique to decompose the target signals $s_i^+(t)$ and $s_i^-(t)$ to obtain the component $IMF_{ij}^+(t)$ and $IMF_{ij}^-(t)$.

Step 3: Repeat steps 1-2 until $i = Ne$, where Ne is the logarithm of white noise added and the number of integrated averages.

Step 4: Integrate all the IMF components obtained above to obtain the first IMF component $s(t)$.

$$IMF_j(t) = \frac{1}{2Ne} \sum_{i=1}^{Ne} [IMF_{ij}^+(t) + IMF_{ij}^-(t)] \quad (11)$$

The amplitude of white noise is added to different signals; there is no consistent standard with the integration number Ne . The magnitude of each added noise is determined from the standard deviation of the raw decomposed signal, which is generally less than 0.5 times the difference of the raw signal standard, Ne is generally within one hundred.

III. NUMERICAL SIMULATION ANALYSIS

There are widespread high-frequency intermittent signals in mechanical signals. To demonstrate the mode mixing issue and verify the validity of CEALIF, the high-frequency intermittent signal is added to the signal composed of modulated signals and harmonic components. The expression is provided by:

$$\begin{aligned} s_1(t) &= \sin(2\pi \times 200t) \times (\delta(t - 0.1) - \delta(t - 0.3) \\ &\quad + \delta(t - 0.7) - \delta(t - 0.9)) \\ s_2(t) &= \sin(2\pi \times 100t) \times (1 + \cos(2\pi \times 15t)) \\ s_3(t) &= \sin(2\pi \times 40t) \\ s(t) &= s_1(t) + s_2(t) + s_3(t) \end{aligned} \quad (12)$$

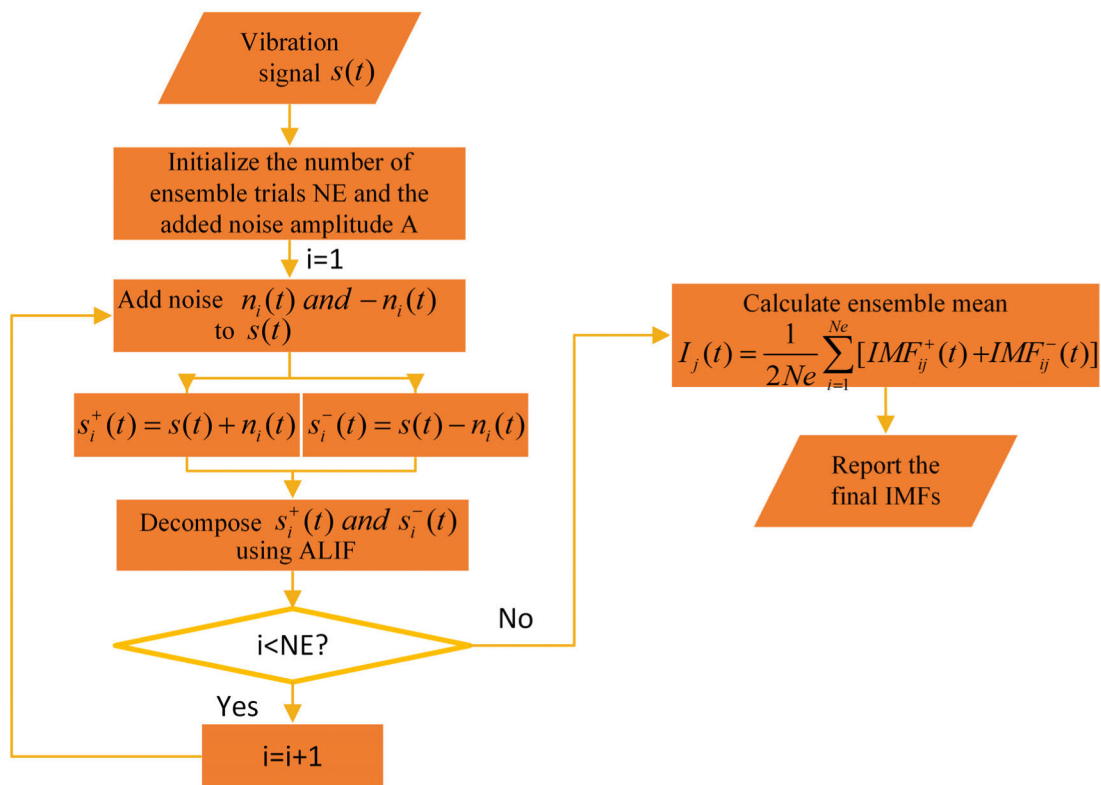


FIGURE 1. The flowchart of CEALIF algorithm.

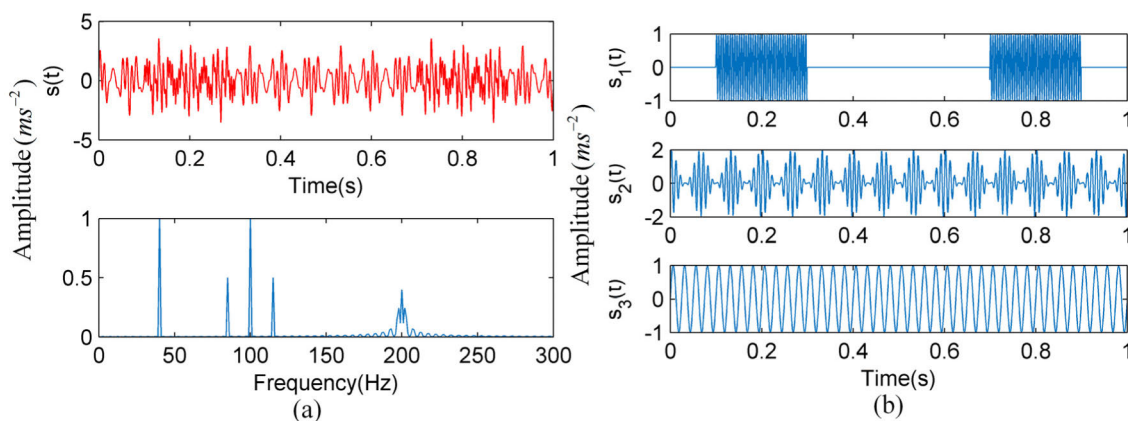


FIGURE 2. High-frequency intermittent signal and its component time-domain waveform.

where $\delta(t)$ is the step signal. $s_1(t)$ is a 2-segment high-frequency intermittent signal with a frequency of 200Hz, $s_2(t)$ and $s_3(t)$ are frequency-modulation signal and harmonic signal respectively. The mixed signal and its frequency spectrum are plotted in Fig. 2(a), and its components are presented in Fig. 2(b).

The intermittent signal is a classic example of mode mixing and the modulated signal is generally a bearing fault signal. We know that the more ensembles there are, the less noise residue there is. To achieve the decomposition effect, the Ne is set to 100. Meanwhile, to avoid the noise

amplitude is too large to swamp the original signal, the noise amplitude is set to 0.1. The mixed signal is decomposed by ALIF and CEALIF methods, respectively. Fig. 3(a) provides the ALIF processing results of the mixed-signal. According to the results, the three components are heavily contaminated by high-frequency intermittent signals and the IMFs are distorted. Fig. 3(b) presents the frequency domain diagram of each component after ALIF decomposition. Where IMF1 contains both frequency modulated signals and high-frequency intermittent signals. Among them, modulation components also appear in IMF2, while IMF2 and

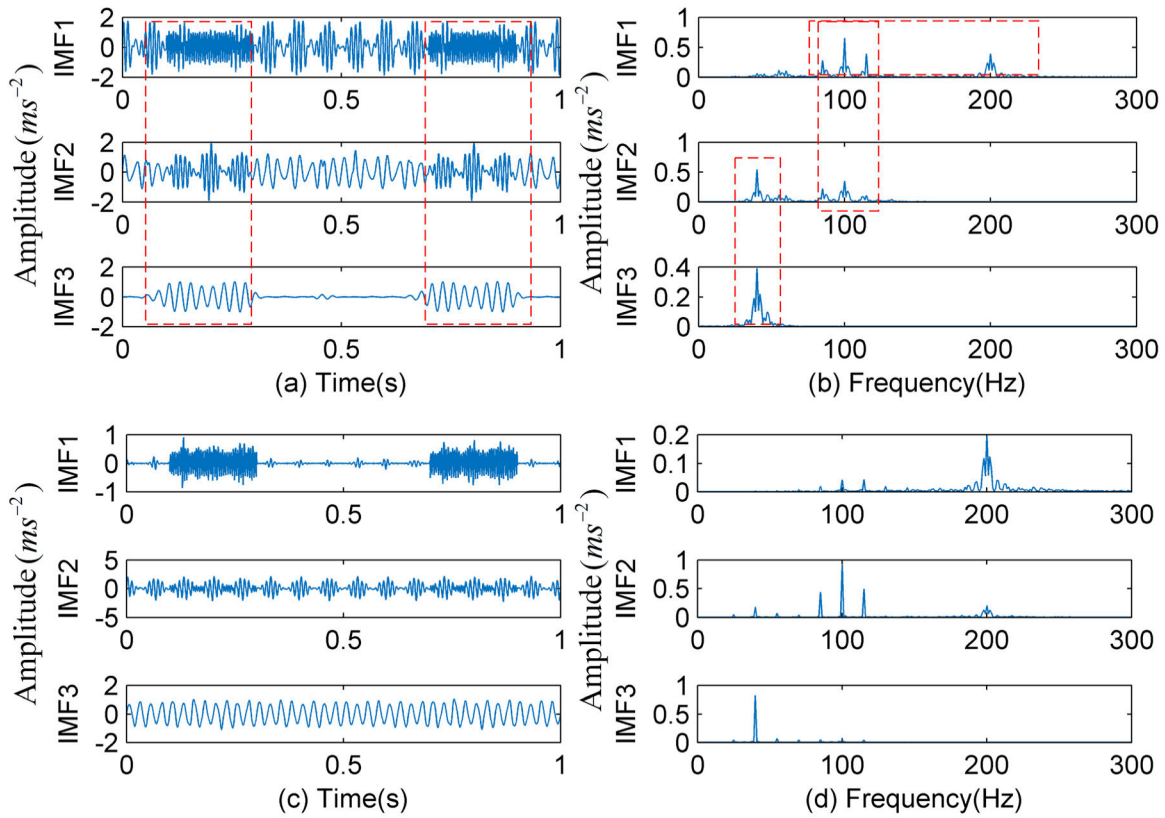


FIGURE 3. Processing results through different approaches: (a) the component obtained by ALIF and (b) its Fourier spectrum; (c) the component obtained by CEALIF and (d) its Fourier spectrum.

IMF3 both contain low-frequency harmonic components. This means that serious mode mixing has occurred in ALIF and the intermittent signals are not separated. Therefore, the decomposition capability of the ALIF algorithm needs to be improved when the signal to be processed contains high-frequency intermittent components.

The CEALIF decomposition results of the high-frequency intermittent mixed-signal and the frequency domain diagrams of its components are depicted in Fig. 3(c) and Fig. 3(d), respectively. The high-frequency intermittent signal is basically decomposed into the IMF1 component, while IMF2 and IMF3 also correspond to the original mixed-signal components. The obtained components not only correspond to the original signal components but also the decomposition results do not produce redundant components, which indicates that the CEALIF method can suppress the modal aliasing effect well and the decomposition results are basically unaffected by the additional noise.

To demonstrate the validity of the CEALIF algorithm, EMD's mode mixing improvement methods EEMD and CEEMDAN are also employed to analyze high-frequency intermittent mixed signals for comparison. EEMD and CEEMDAN methods process high-frequency intermittent mixed signals to obtain 12 and 9 IMF components, respectively. For simplicity, the first six effective IMF components are taken for analysis, and the results of EEMD and

CEEMDAN are displayed in Fig. 4 and Fig. 5, respectively. From Fig. 4, we can see that the time domain components in the high-frequency intermittent band are subject to different degrees of pollution, and the distortion is more serious. In the frequency domain, it can be found that IMF1 and IMF2 contain multiple frequency components, and low-frequency harmonic components also appear in IMF2 and IMF3 at the same time. That is to say, EEMD in processing contains high-frequency intermittent signal will be a serious mode mixing. In the CEEMDAN processing of Fig. 5, the low-frequency harmonic components are completely separated and only the high-frequency intermittent and modulated signals undergo modal mixing. That is to say, CEEMDAN will also suffer from modal aliasing when processing signals with high-frequency intermittent signals, but the aliasing will be less than EEMD and the redundant component will be less than EEMD.

The simulated signals indicate that ALIF has serious mode aliasing when the signal to be processed contains high-frequency intermittent components, while CEALIF improves the performance of adaptive filter by adding white noise based on ALIF, which makes the performance of CEALIF method better and has high recognition ability for different modal components. By comparing with EEMD and CEEMDAN, which are proposed to address the modal aliasing of EMD, CEEMDAN suffers less modal aliasing than

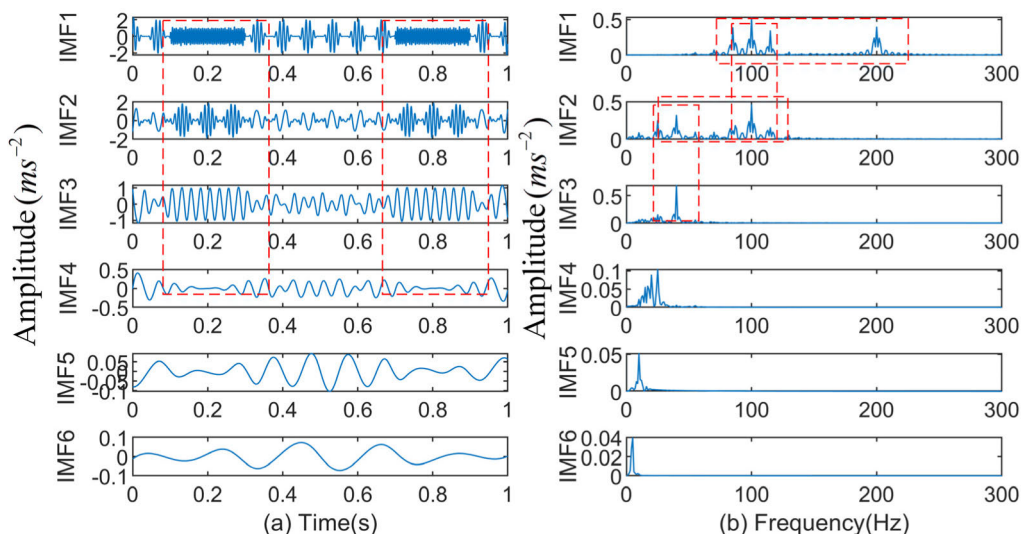


FIGURE 4. EEMD analysis results and frequency spectrum of high-frequency intermittent signals.

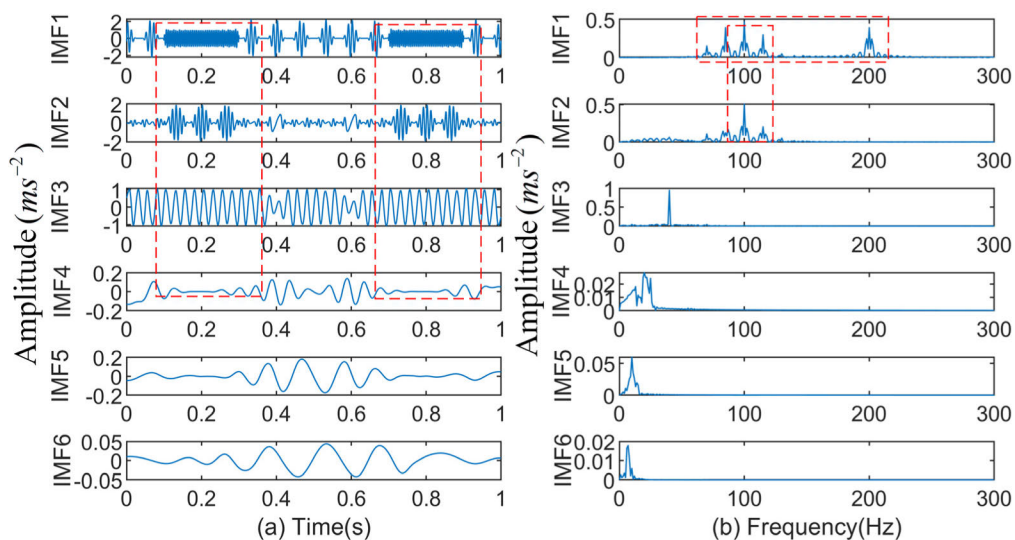


FIGURE 5. CEEMDAN analysis results and frequency spectrum of high-frequency intermittent signals.

EEMD when dealing with high-frequency intermittent signals, but there is still a certain gap compared with the CEALIF decomposition results.

IV. FEATURE EXTRACTION, SELECTION, AND RECOGNITION

A. FEATURE PARAMETER EXTRACTION

When the vibration state of the rolling bearing changes, the TD information, FD information, and TFD information of the measured signal may change accordingly [40]. As the measured rolling bearing vibration signal is more complex, the feature extraction approach based on a single TD or FD has certain limitations. To obtain more fault information, the TD, FD and TFD features of the vibration signal are extracted simultaneously in this article.

Time-domain (TD) indicators (such as kurtosis, impulse factor) and frequency domain (FD) characteristics (such as frequency variance, mean square frequency) are physical quantities based on statistical characteristics, which reflect the amplitude and energy fluctuations in the TD and FD of the vibration signal, and when the vibration signal changes, these characteristics can be accurately monitored from both the TD and FD, so they can effectively characterize the fault signal [40].

When mechanical equipment failure, not only the vibration signal of TD and FD characteristics of the change, and the time-frequency domain (TFD) energy characteristics will also change, mainly manifested in the time-frequency surface of the different time-frequency block energy distribution differences, using the time frequency entropy (TFE) to represent the energy differences.

TABLE 1. Feature parameters and their expressions.

Features	Expression	Features	Expression
Mean (M)	$f_m = \frac{1}{N} \sum_{i=1}^N x_i$	Impulse factor (IF)	$f_{IF} = \frac{f_{pk}}{\frac{1}{N} \sum_{i=1}^N x_i }$
Root mean square (RMS)	$f_{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i)^2}$	Margin factor (MF)	$f_{MF} = \frac{f_{pk}}{(\frac{1}{N} \sum_{i=1}^N x_i ^{1/2})^2}$
Peak (PK)	$f_{pk} = \max x_i $	Center frequency (FC)	$f_{FC} = \sum_{i=2}^N \dot{x}_i \dot{x}_i / (2\pi \sum_{i=1}^N x_i^2)$
Crest factor (CF)	$f_{CF} = \frac{f_{pk}}{f_{rms}}$	Root mean square frequency (RMSF)	$f_{RMSF} = \sqrt{\sum_{i=2}^N \dot{x}_i^2 / (4\pi^2 \sum_{i=1}^N x_i^2)}$
Kurtosis (KU)	$f_{ku} = \frac{N \sum_{i=1}^N (x_i - f_m)^2}{(\sum_{i=1}^N (x_i - f_m)^2)^2}$	Root variance frequency (RVF)	$f_{RVF} = \sqrt{f_{RMSF}^2 - f_{FC}^2}$
Shape factor (SF)	$f_{SF} = \frac{f_{rms}}{\frac{1}{N} \sum_{i=1}^N x_i }$	Time-frequency entropy (TFE)	$TFE(e) = -\sum_{i=1}^N e_i \ln e_i$

Noted : x_i is the signal to be analyzed and N is the number of points.

The TFE is determined as follows: the time frequency plane is divided into N blocks of equal areas, where the total energy is E and the energy of each block is E_i . Normalize the energy of each block to obtain $e_i = E_i/E$, so there is $\sum_{i=1}^N e_i = 1$. In line with the initial normalization conditions for calculating information entropy, imitating the calculation formula of information entropy, the TFE of the signal is determined as $TFE(e) = -\sum_{i=1}^N e_i \ln e_i$ [44]. In conclusion, to obtain more information about the fault status, 11 characteristics are extracted in this article, as listed in the **Table 1**.

Since the values of more feature parameters do not belong to the same order of magnitude, the feature matrix is normalized before entering the fault model. The normalization formula utilized was as follows:

$$x_{norm} = (\kappa_{max} - \kappa_{min}) \times \frac{x - x_{min}}{x_{max} - x_{min}} + \kappa_{min} \quad (13)$$

where $[\kappa_{min}, \kappa_{max}]$ is the normalized interval, x_{min} and x_{max} respectively correspond to the minimum and maximum values of the characteristic matrix respectively. This article normalizes the characteristic parameters to $[-1, 1]$.

B. FEATURE SELECTION-BASED LAPLACIAN SCORE(LS)

Although the faults in the equipment can be identified from different angles based on the above features, however, all features are not equally sensitive to faults. Some features are highly sensitive to faults, and there may also be components that are irrelevant or even redundant. Therefore, the feature

set may be further sorted and selected before it is entered into the classifier [45].

Laplacian score (LS), based on mapping and projection [46], measures feature through local retention capabilities, directly learns the feature set to extract the internal information structure of the data, transform high-dimensional features into low-dimensional ones. The selection of features with smaller scores in the low-dimensional feature space largely preserves information about the overall geometric structure contained in the failure signal character set and thus aids in rolling bearing fault recognition.

Let L_k be the LS of the k-th feature, p_{ki} be the k-th feature of the i-th sample, where $k = 1, 2, \dots, n$ represents the number of features of each sample and $i = 1, 2, \dots, m$ represents the sample. The Laplace score approach is presented below:

Step 1: Given a sample of features, the nearest neighbor graph G with m nodes is constructed, where x_i denotes the i-th node. Then the premise that x_i and x_j are connected by edges is that they belong to the same class and are k neighbors of each other.

Step 2: The distance weight of the edges in G are denoted by Equation (14).

$$S_{ij} = \exp(-\|x_i - x_j\|^2 / (2\sigma^2)) \quad (14)$$

where σ is a suitable constant; if x_i and x_j have no edges connected, $S_{ij} = 0$.

Step 3: For the k-th feature, define

$$p_k = (p_{k1} p_{k2}, \dots, p_{km})^T \quad (15)$$

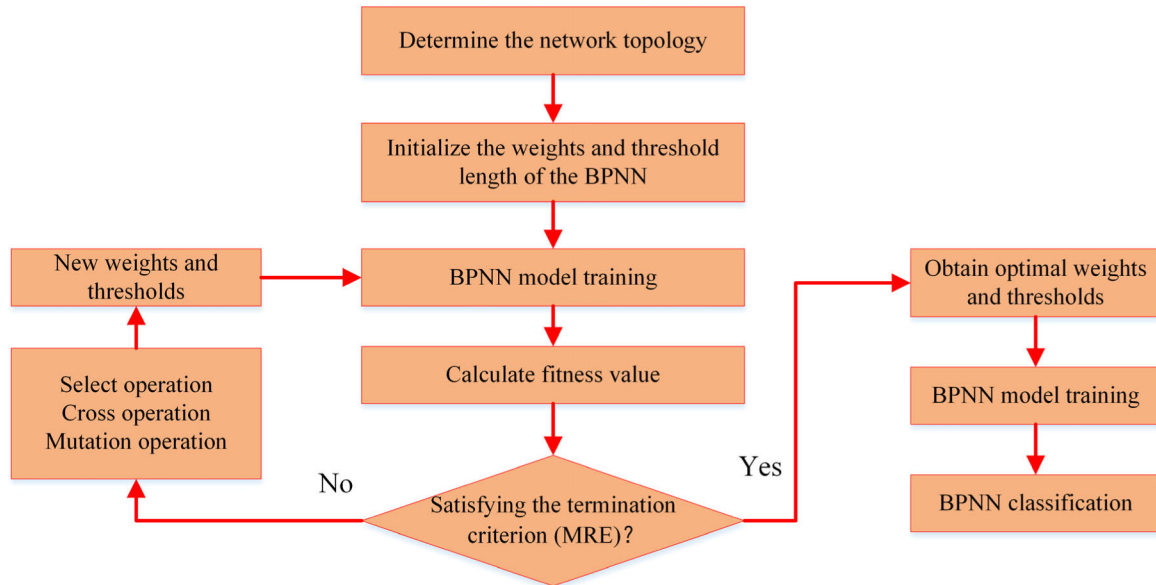


FIGURE 6. Flowchart of GA-BPNN.

$$D = \text{diag}(SI) \quad (16)$$

$$I = (1, 1, \dots, 1)^T \quad (17)$$

$$L = D - S \quad (18)$$

The matrix L is often denoted as the Laplace matrix of graph G . De-averaging p_r yields:

$$\bar{p}_k = p_k - \frac{p_k DI}{I^T DI} I \quad (19)$$

Step 4: Calculate the LS of the k -th feature.

$$L_k = \frac{\sum_{ij} (p_{ki} - p_{kj})^2 S_{ij}}{\text{Var}(p_k)} = \frac{\bar{p}_k^T L \bar{p}_k}{\bar{p}_k^T D \bar{p}_k} \quad (20)$$

where $\text{Var}()$ denotes the variance.

It is observed that for a well-defined feature, the larger S_{ij} is, the lower the corresponding Laplacian score is. Therefore, the feature with a lower LS is selected as the feature vector.

C. GENETIC ALGORITHM-BASED BPNN

BPNN have strong self-organizing learning capabilities via error backpropagation, which can realize classification through the nonlinear mapping between the layers. However, the choice of initial weights and thresholds significantly affects its nonlinear mapping performance.

Genetic algorithm (GA) [47], [48] is an intelligent optimization algorithm with powerful global nonlinear optimization function. In this article, GA was employed to optimize the initial weight and threshold to boost the nonlinear mapping capacity and convergence speed of the BPNN. In the GA, the population individuals are firstly coded; the neural network training error is then adopted as the fitness function; secondly, the best fitness individuals of the population are obtained through selection, adaptive crossover and mutation operators. Finally, the best initial weights and thresholds are

obtained by the best fitness individuals of the population until the set prediction error requirements are satisfied or the set maximum number of iterations are reached, thus obtaining a mature GA-BPNN. where the mean relative error (MRE) is set as the termination criterion, the expression of which is shown in Equation (21). Fig. 6 depicts the whole flow chart of GA-BPNN.

$$MRE = \frac{1}{N} \sum_{i=1}^N \frac{|x_i - \hat{x}|}{x_i} \times 100\% \quad (21)$$

where N denotes the sample size, x_i is the true value, and \hat{x} is the predicted value.

V. THE PROPOSED FAULT DIAGNOSIS MODEL AND EXPERIMENTAL VERIFICATION

A. THE PROPOSED MODEL BASED ON CEALIF

1) CEALIF is employed to process each sample signal and generate a sequence of IMFs for each.

2) Select the first 5 IMF components, extract their TD and FD features in section 4.1 and extract the TFE of the vibration signal to construct the following initial feature vector: $T = [M_1, RMS_1, PK_1, CF_1, KU_1, SF_1, IF_1, MF_1, FC_1, RMSF_1, RVF_1, M_2, RMS_2, PK_2, CF_2, KU_2, SF_2, IF_2, MF_2, FC_2, RMSF_2, RVF_2, M_3, RMS_3, PK_3, CF_3, KU_3, SF_3, IF_3, MF_3, FC_3, RMSF_3, RVF_3, M_4, RMS_4, PK_4, CF_4, KU_4, SF_4, IF_4, MF_4, FC_4, RMSF_4, RVF_4, M_5, RMS_5, PK_5, CF_5, KU_5, SF_5, IF_5, MF_5, FC_5, RMSF_5, RVF_5, TFE]_{1*56}$.

3) Calculating the LS of the initial feature matrix composed of the initial feature vector T , ranking the feature vectors in order of importance from lowest to highest according to their scores, and selecting the first five features as the novel failure feature vectors.

4) The obtained new failure feature vectors are input into the GA-BPNN for training the network to enable automatic

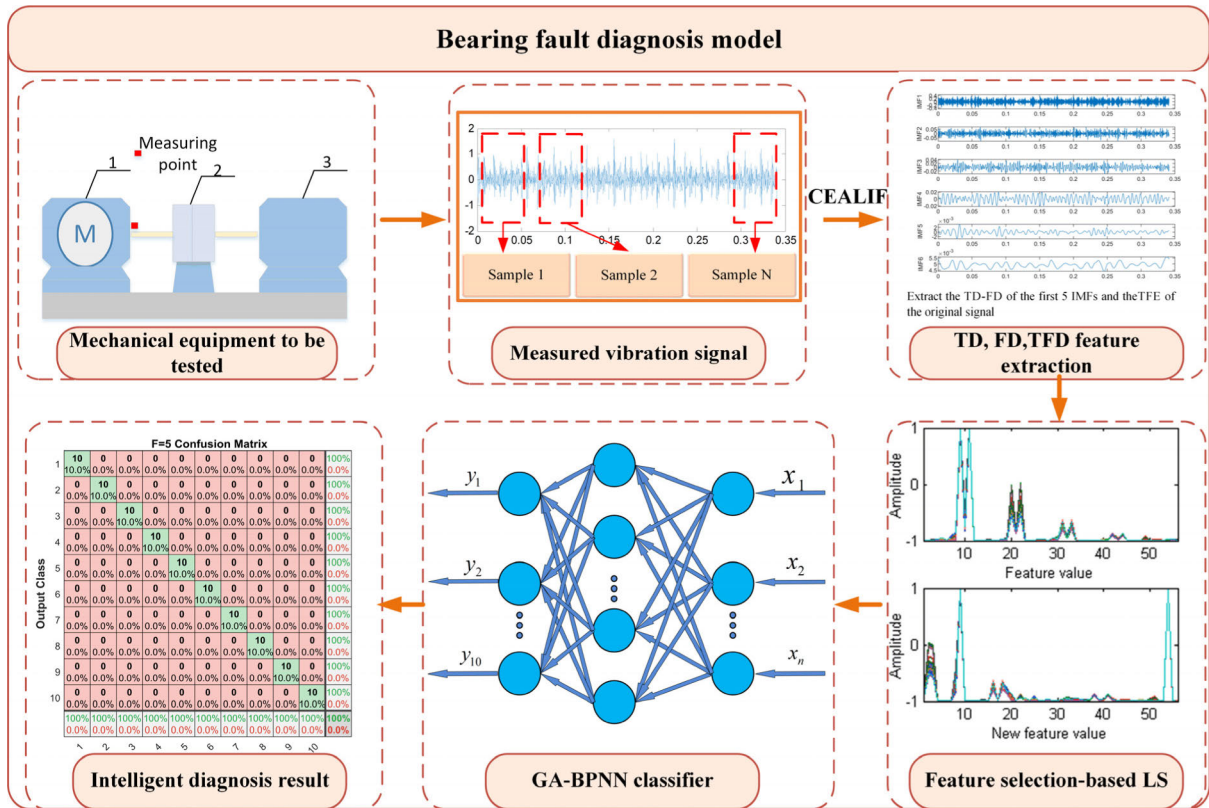


FIGURE 7. Flow chart of bearing fault diagnosis model.

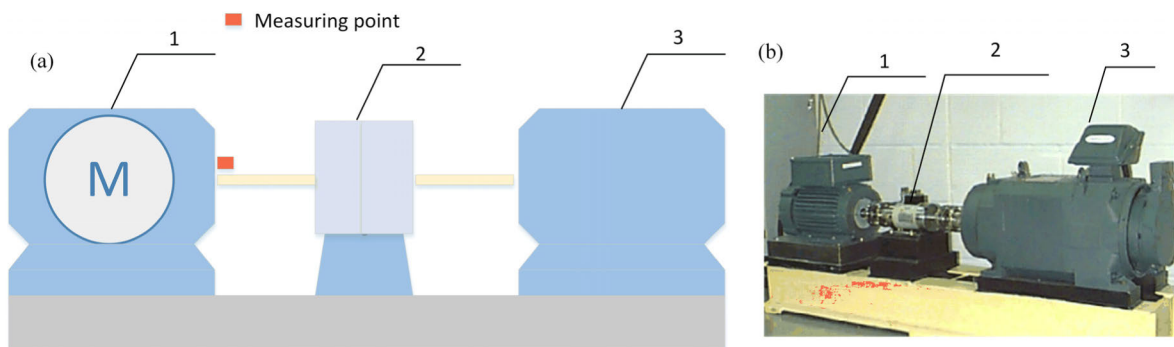


FIGURE 8. Rolling bearing experimental device and transmission diagram: (a) Transmission diagram; (b) experimental equipment, 1, 3-Hp motor, 2-Torque transducer and encoder.

diagnosis of rolling bearings. The proposed model is schematized in Fig. 7.

B. EXPERIMENTAL VERIFICATION

1) CASE 1: APPLICATION TO OPEN-SOURCE DATASETS

In order to validate the validity of the proposed rolling bearing fault diagnosis model, bearing experimental signals from Case Western Reserve University were employed to test the model performance. The transmission sketch as well as the experimental apparatus is presented in Fig. 8 and its description is given in [49]. The data collected at the drive end was utilized for analysis and truncated in 4096 points.

For the above experimental data, the types of failure are categorized as rolling element failure, inner race failure and outer race failure. The degree of failure is categorized as slight, moderate and severe, and considering the normal operating condition, a total of 10 categories need to be identified. A sample waveform for each of these categories are displayed in Fig. 9. For each category, 10 samples were randomly selected for training and the other 10 samples were employed to determine the type of failure. In total, there are 100 training samples and 100 sets of test samples. All data types are listed in Table 2. Since the measured vibration signals containing mechanical faults generally present

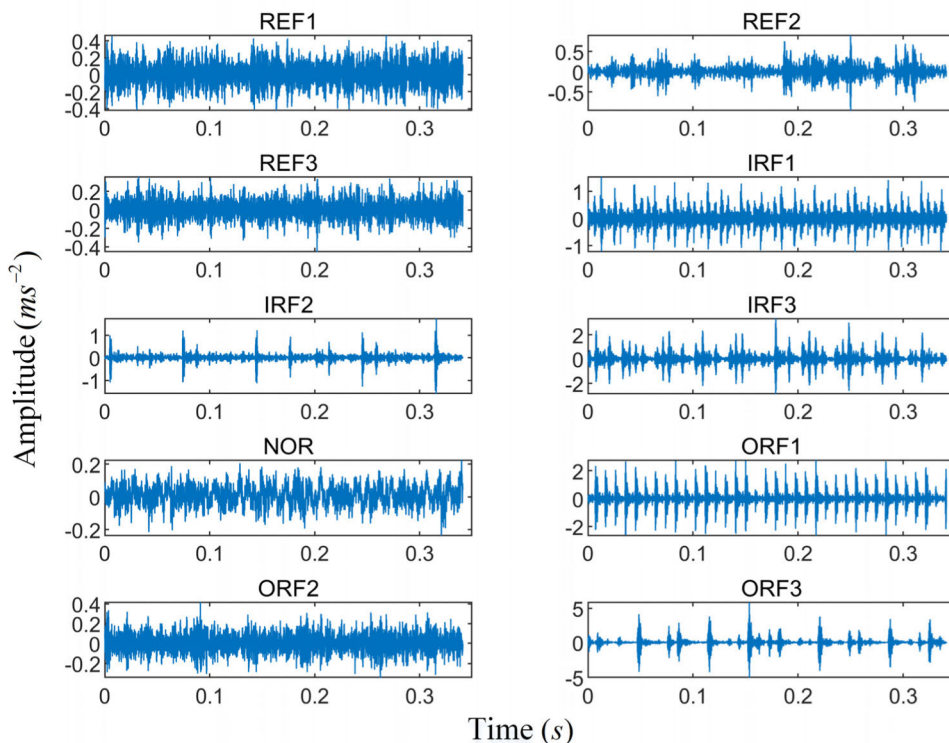


FIGURE 9. Time domain diagrams for each condition.

TABLE 2. The experimental datasets were employed for identification.

Fault Category	Degree of failure	Fault Diameter	Fault abbreviation	Label of Classification
Rolling element failure	Slight failure	0.1778mm	REF1	1
Rolling element failure	Moderate failure	0.3556mm	REF1	2
Rolling element failure	Severe failure	0.5334mm	REF1	3
Inner race failure	Slight failure	0.1778mm	IRF1	4
Inner race failure	Moderate failure	0.3556mm	IRF2	5
Inner race failure	Severe failure	0.5334mm	IRF3	6
Normal	No	0	NOR	7
Outer race failure	Slight failure	0.1778mm	ORF1	8
Outer race failure	Moderate failure	0.3556mm	ORF2	9
Outer race failure	Severe failure	0.5334mm	ORF3	10

nonlinear and non-stationary properties, it is impossible to identify the failure type only based on the waveforms in Fig. 9. Therefore, it is essential to further decompose the time-domain waveform.

Firstly, CEALIF is adopted to decompose each vibration signal of each category separately and to obtain multiple IMFs; secondly, the aforementioned characteristic parameters of the first five IMF components and the TFE of the vibration signal are extracted to construct the feature vector $T = [T_i]_{1 \times 56}$. Then, the feature vectors T of 100 training samples is formed into the initial feature matrix $\tilde{T} = [T]_{100 \times 56}$. The LS of each feature in the matrix $\tilde{T} = [T]_{100 \times 56}$ is calculated

and the 56 features are sorted according to the LS from smallest to largest, and then the top 5 features with smaller scores are considered as the optimal feature vectors (the number of optimal feature vectors is denoted by F) to form the final feature matrix $\tilde{IT}_{100 \times 5}$. For the above experimental datasets, the order of LS from low to high is: $LS(T_{56}) < LS(T_{22}) < LS(T_{20}) < LS(T_{13}) < LS(T_{17}) < LS(T_2) < LS(T_{39}) < LS(T_{10}) < LS(T_9) < LS(T_{28}) < LS(T_6) < LS(T_{35}) < LS(T_{14}) < LS(T_3) < LS(T_{46}) < LS(T_{31}) < LS(T_{50}) < LS(T_{33}) < LS(T_{21}) < LS(T_8) < LS(T_{36}) < LS(T_{32}) < LS(T_5) < LS(T_{45}) < LS(T_7) < LS(T_{24}) < LS(T_{47}) < LS(T_4) < LS(T_{25}) < LS(T_{16}) < LS(T_{27}) < LS(T_{30}) <$

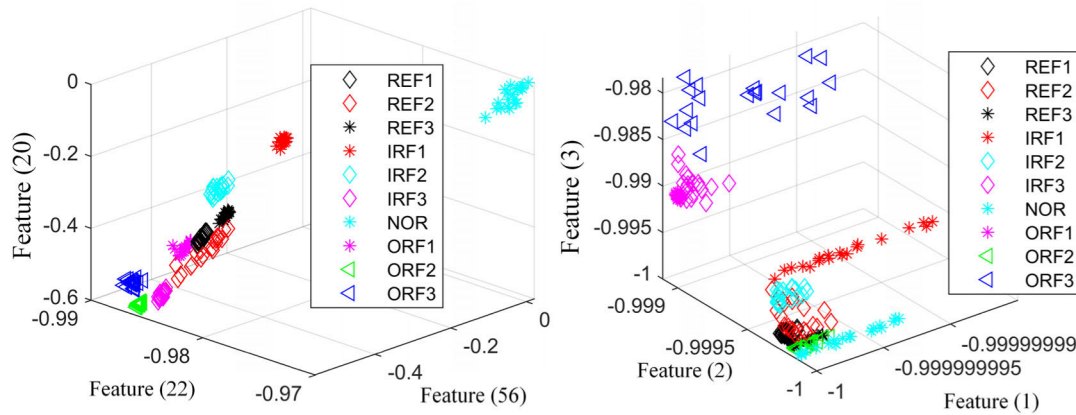


FIGURE 10. Comparison between the sorted optimized feature set and the random initial feature set (left T56, T22, T20 right T1, T2, T3).

$LS(T_{19}) < LS(T_{29}) < LS(T_{18}) < LS(T_{15}) < LS(T_{54}) <$
 $LS(T_{42}) < LS(T_{55}) < LS(T_{26}) < LS(T_{44}) < LS(T_{43}) <$
 $LS(T_{53}) < LS(T_{52}) < LS(T_{51}) < LS(T_{48}) < LS(T_{37}) <$
 $LS(T_{40}) < LS(T_{38}) < LS(T_{41}) < LS(T_{49}) < LS(T_{12}) <$
 $LS(T_1) < LS(T_{11}) < LS(T_{23}) < LS(T_{34}).$

According to LS, the first five optimal feature values can be obtained as the TFE of the original signal, the RMSF of IMF2, the FC of IMF2, the RMS of IMF2, and the SF of IMF2, which indicates that the second component of IMF2 obtained from the decomposition contains important fault information. In addition, the first five selected features completely contain the TD, FD, and TFD indicators, which make the diagnosis results more reliable. Furthermore, to demonstrate the better fault correlation of the sort-optimized feature set, we plotted the first three features of the sort-optimized feature set (T_{56}, T_{22}, T_{20}) and compared them with the random features (T_1, T_2, T_3) before the sort-optimization, as illustrated in Fig. 10. It can be seen that the samples in the sort-optimized feature set are mostly clustered around their corresponding centers and a clear demarcation line between classes can be observed. On the contrary, the unordered optimized feature set basically exhibits a dispersion trend and most of the categories have no obvious boundaries between them.

Finally, the optimal feature vector matrix $\tilde{IT}_{100 \times 5}$ is imported into the GA-BPNN categorizer for training to derive the prediction models of each category. According to the feature matrix $\tilde{IT}_{100 \times 5}$, extract the 5 optimal feature values of 100 test samples (for the following comparison, the paper still extracts 56 feature values of each vibration signal), and the feature matrix $OT_{100 \times 5}$ of the test sample is obtained, which is input into the trained GA-BPNN categorizer to generate the identification results of the test samples, as presented in Fig. 11. The results demonstrate that the proposed model has a good fault recognition effect, which not only realizes the distinction of fault categories but also the partitioning of different fault degrees, and the test samples have a high fault identification rate (100%).

In addition, the number of feature values has an important impact on the diagnosis result. If the dimension of the feature

F=5 Confusion Matrix

Output Class	1	2	3	4	5	6	7	8	9	10	Accuracy
1	10 100.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100.0%
2	0 0.0%	10 100.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100.0%
3	0 0.0%	0 0.0%	10 100.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100.0%
4	0 0.0%	0 0.0%	0 0.0%	10 100.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100.0%
5	0 0.0%	0 0.0%	0 0.0%	0 0.0%	10 100.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100.0%
6	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	10 100.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100.0%
7	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	10 100.0%	0 0.0%	0 0.0%	0 0.0%	100.0%
8	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	10 100.0%	0 0.0%	0 0.0%	100.0%
9	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	10 100.0%	0 0.0%	100.0%
10	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	10 100.0%	100.0%
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

FIGURE 11. GA-BPNN output results of the model proposed in this article.

vector is too small, that is, the number of feature values is small, it will not be able to fully reflect and distinguish the fault type and fault degree, which will affect the reliability of diagnosis; on the contrary, too many eigenvalues will cause information redundancy and increase the classifier Training time, thereby reducing diagnostic efficiency. For comparison, this article selects the first F feature values ($F = 3, 4, 5, 6, 7$) after LS sorting as the feature vector input to the GA-BPNN classifier. The number of samples from the training and the test samples remained the same. After training, a prediction model was established and the output results of the test samples were tested, as depicted in Fig. 12. The results revealed that too few features will lead to poor diagnostic accuracy, and too many features will inevitably lead to a decrease in diagnostic efficiency. It means that too many or too few features have an inappropriate effect on the

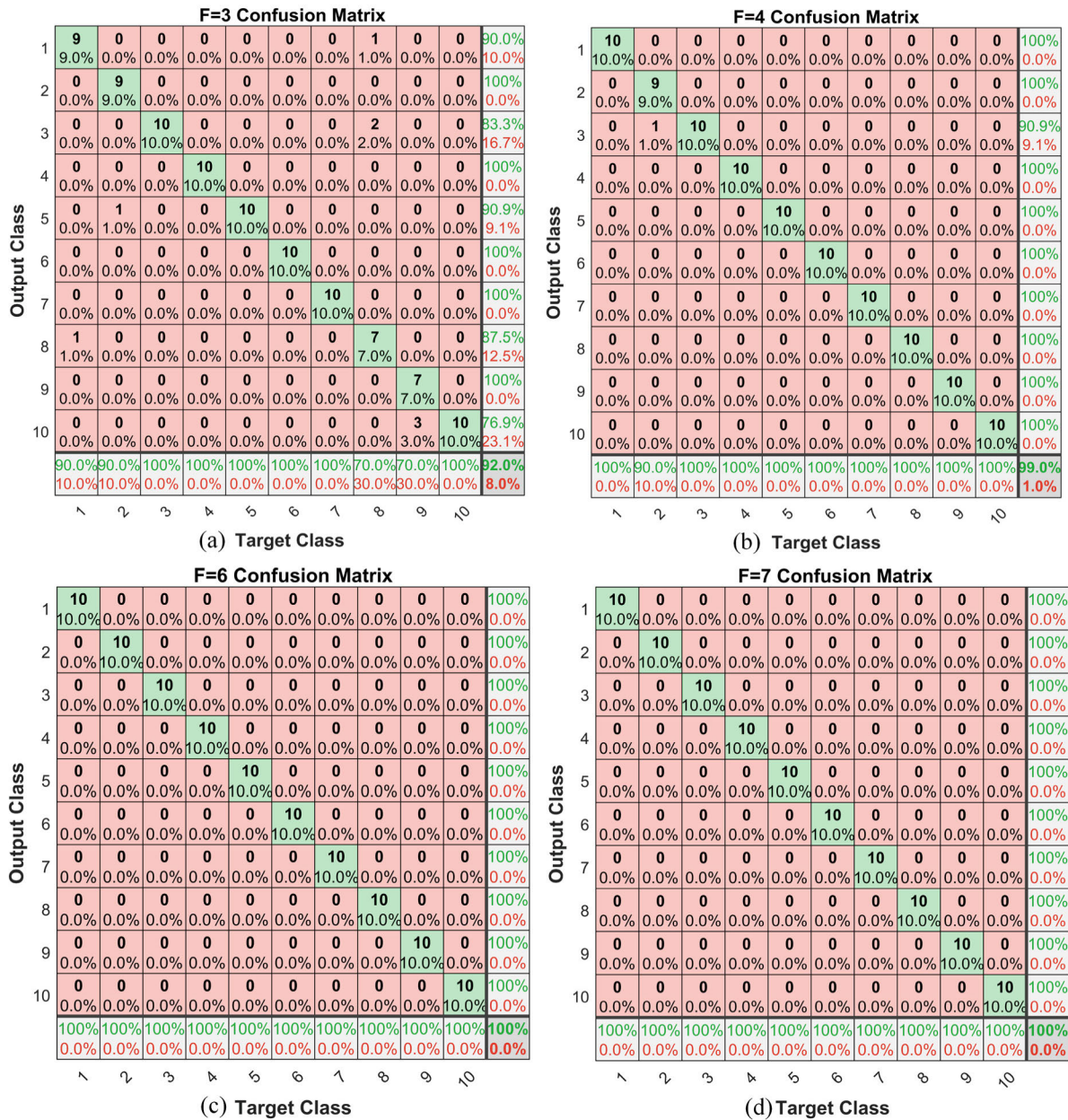


FIGURE 12. GA-BPNN output results for test samples with different feature values: (a) F = 3; (b) F = 4; (c) F = 6; (d) F = 7.

failure identification results. Meanwhile, the conclusions in Fig. 12 also verify that it is feasible and reasonable to select the first five characteristic quantities in the fault diagnosis model presented in this article.

To illustrate the necessity and superiority of the LS-optimized feature vectors, select the F (F = 3, 4, 5, 6, 7) feature values before LS sorting as feature vectors to be imported into the GA-BPNN categorizer. The number of training and test samples remains the same, the trained prediction model is created, and the test samples are obtained with the recognition rates shown in Table 3. Comparing Table 3 with Fig. 11 and Fig. 12, it can be obtained that the feature vector is not optimized by LS, and for the same GA-BPNN categorizer,

the recognition rate of the test sample is significantly lower than the results of the test sample with the LS-optimized feature vector as the input feature vector, which shows that LS optimization of feature values has certain advantages and necessities.

In addition, to illustrate the superiority of GA-BPNN classifier, an unoptimized BPNN categorizer was employed to do the same process, in which the network structure was 8 layers of hidden layer and 10 layers of output layer, the training target was 0.001, the maximum number of training was 1000, and other parameters were set by default. The feature vectors are composed of the first 3 to 7 feature values after LS optimization, and the number of samples from the training

TABLE 3. GA-BPNN output result of feature vector not sorted by LS.

Fault Category	Category	GA-BPNN(F=3)	GA-BPNN(F=4)	GA-BPNN(F=5)	GA-BPNN(F=6)	GA-BPNN(F=7)
REF1	1	1(10)	1(10)	1(10)	1(10)	1(10)
REF2	2	2(8), 3(1), 5(1)	2(8), 3(1), 5(1)	2(10)	2(9), 5(1)	2(8), 3(1), 5(1)
REF3	3	3(8), 9(2)	3(8), 9(2)	3(10)	3(8), 9(2)	3(10)
IRF1	4	4(10)	4(10)	4(10)	4(10)	4(10)
IRF2	5	5(10)	5(10)	5(10)	5(10)	5(10)
IRF3	6	6(9), 10(1)	6(9), 8(1)	6(10)	6(9), 8(1)	6(10)
NOR	7	7(10)	7(10)	7(10)	7(10)	7(10)
ORF1	8	8(10)	8(10)	8(10)	8(10)	8(10)
ORF2	9	9(7), 3(3)	9(7), 3(3)	9(8), 3(2)	9(9), 3(1)	9(10)
ORF3	10	10(10)	10(10)	10(10)	10(10)	10(9), 5(1)
Recognition accuracy		92%	92%	98%	95%	97%

Noted :9(7), 3(3) represent 10 test samples, 7 output is category 9, and 3 errors are classified as category 3 and other meanings are the same as here.

TABLE 4. The output result of the BPNN after the feature vector is selected by LS.

Fault Category	Category	BPNN(F=3)	BPNN(F=4)	BPNN(F=5)	BPNN(F=6)	BPNN(F=7)
REF1	1	1(10)	1(10)	1(4), 6(6)	1(10)	1(10)
REF2	2	2(9), 5(1)	2(7), 5(3)	2(9), 5(1)	2(10)	2(10)
REF3	3	3(3), 5(7)	3(9), 5(1)	3(10)	3(10)	3(10)
IRF1	4	4(10)	4(10)	4(10)	4(10)	4(9), 1(1)
IRF2	5	5(10)	5(10)	5(10)	5(9), 2(1)	5(10)
IRF3	6	6(10)	6(10)	6(10)	6(10)	6(10)
NOR	7	7(10)	7(10)	7(10)	7(10)	7(10)
ORF1	8	8(10)	8(10)	8(10)	8(10)	8(10)
ORF2	9	9(10)	9(10)	9(10)	9(10)	9(10)
ORF3	10	10(10)	10(10)	10(10)	10(10)	10(10)
Recognition accuracy		92%	96%	93%	99%	99%

and the test samples remained the same. The output results of the test samples are obtained by training BPNN classifier, and the recognition rate of the test samples is shown in **Table 4**. Comparing **Table 4** with **Fig. 11** and **Fig. 12**, it is found that for the same features, the recognition rate of BPNN is lower than GA-BPNN, which indicates the necessity of optimizing the neural network and verifies that GA-BPNN is an effective classification method.

2) CASE 2: APPLICATION TO ROTATING MECHANICAL FAILURE TEST BENCHES

To further validate the proposed model, different bearing damage vibration signals were collected under different operating conditions, with damage types including rolling element failure, outer race failure, and inner race failure. The experimental equipment is shown in **Fig. 13**, which consists of an AC motor, a coupling and a magnetic powder brake. The test bearing is a cylindrical roller bearing of model SKF 6202-2Z. The above damages were simulated using the EDM technique

TABLE 5. Bearing test states.

Operation state	1	2	3
Rotating speed (rpm)	1200	1320	1500
Load (N/m)	0	0.8	1.6

with a damage depth of 1.5 mm and the three faulty bearings are presented in **Fig. 14**. The data sampling frequency is 2560 Hz. Three operating conditions were designed, as shown in **Table 5**, with four bearing classes for each condition. Each of the different operating conditions contained 35 sets of data samples per operating condition, resulting in a total of 140 sets of data samples per operating condition, each consisting of 4096 data points.

For the above experimental data, four categories of bearing working conditions under different operation states are identified separately. The waveforms of the four bearing categories under the three operating conditions are depicted in **Fig. 15**.

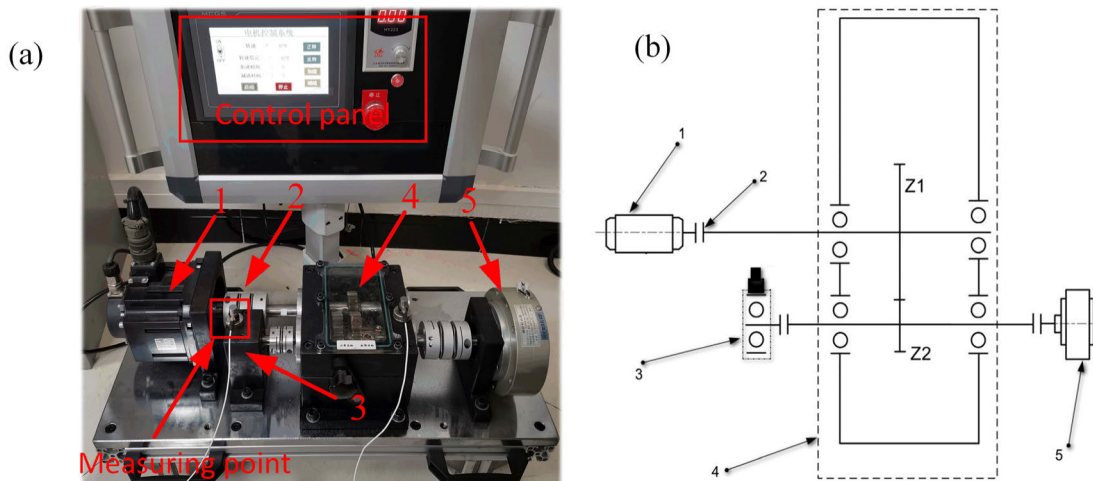


FIGURE 13. Rotating machinery experiment platform and transmission diagram: (a) Experimental equipment; (b) Transmission diagram, 1-AC motor, 2-coupling, 3-bearing housing, 4-Gearbox, 5-Magnetic powder brake.

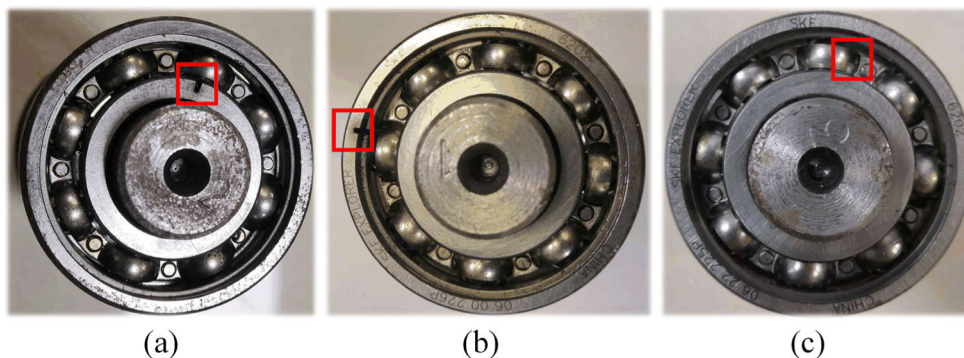


FIGURE 14. Bearing defect diagram: (a) Inner race failure; (b) Outer race failure; (c) Rolling element failure.

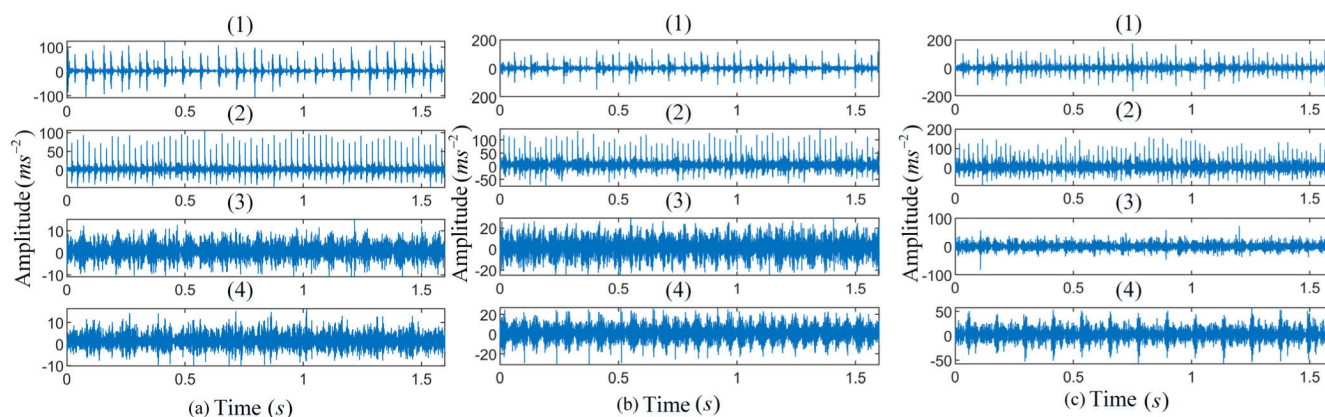


FIGURE 15. The collected vibration signals of four categories ((1) Inner race failure; (2) Outer race failure; (3) Rolling element failure; (4) Normal) under different operating state: (a) operation at 1200 rpm with a load of 0; (b) operation at 1320 rpm with a load of 0.8 N/m; (c) operation at 1500 rpm with a load of 1.6 N/m.

For each category, 20 samples were arbitrarily picked for training and the remainder were taken for testing, and the data sample labels are shown in Table 6. Hence, there were

80 training samples and 60 test samples for each condition. The above operation was implemented for all three operating states listed in Table 5.

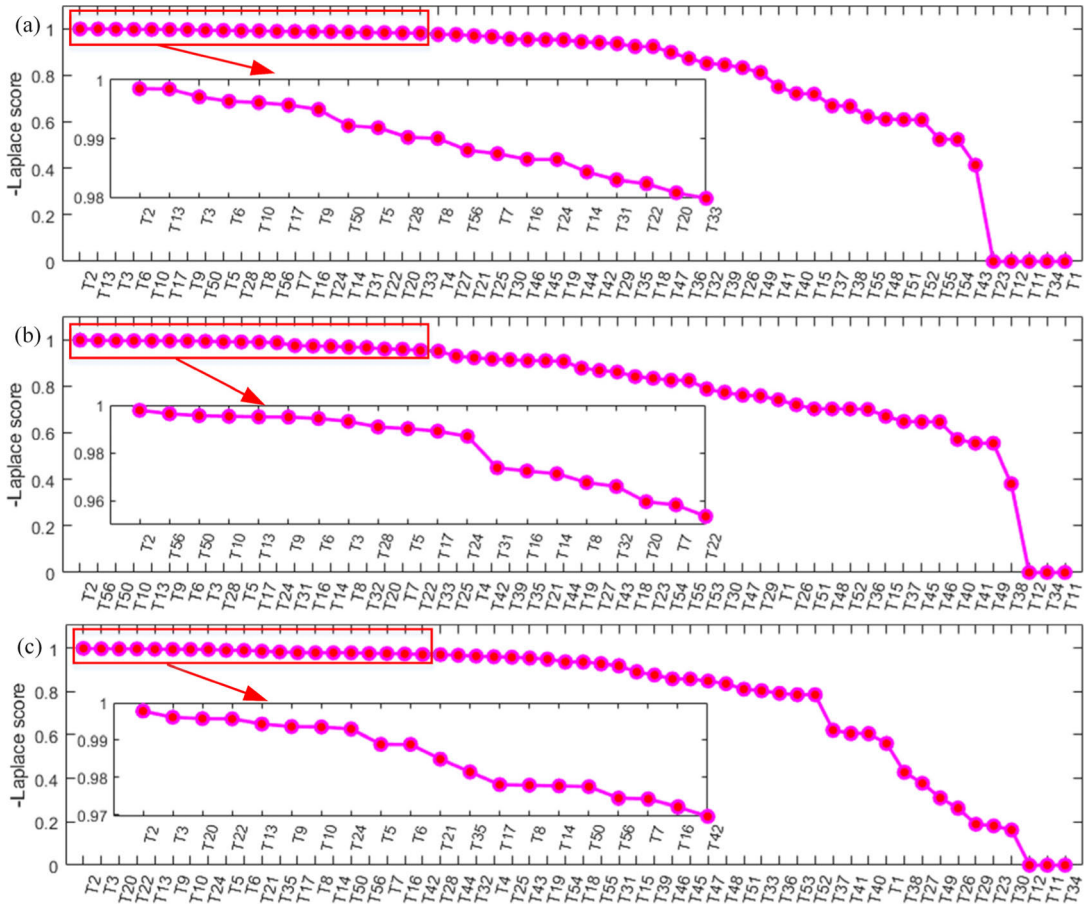


FIGURE 16. LS in 3 operating states: (a) operation at 1200 rpm with a load of 0; (b) operation at 1320 rpm with a load of 0.8 N/m; (c) operation at 1500 rpm with a load of 1.6 N/m.

TABLE 6. Experimental data for identification.

Fault Category	Train number	Test number	Label of Classification
Normal	20	15	1
Outer race failure	20	15	2
Inner race failure	20	15	3
Rolling element failure	20	15	4

The model proposed in this article is deployed for three different operating states. Firstly, CEALIF is employed to pre-process each sample and obtain a number of IMFs, and then TD and FD parameters of the first five IMFs and the TFE of the raw vibration signal are extracted in turn, which can constitute the feature vector $T = [T_i]_{1 \times 56}$. The feature vector T of the 80 training samples is then constructed into the initial feature matrix $\tilde{T} = [T]_{80 \times 56}$. Since the values of each feature parameter do not belong to the same order of magnitude, the feature matrix is normalized by Equation (13). The next step is to compute the LS of each characteristic and prioritize it from smallest to largest, select the first five features with the smallest score as the optimal features and form

the optimal feature vector matrix. For the above experimental data, the inverse of the LS was adopted as an indicator for a clearer display, as shown in Fig. 16, which shows the LS scores for the three operating conditions.

According to the LS optimization ranking in Fig. 16, the first five important indicators for operating state 1 are RMS of IMF1, the RMS of IMF2, the PK of IMF1, the SF of IMF2, the RMSF of IMF1; the first five important indicators for operating state 2 are the RMS of IMF1, TFE for IMF5, the SF of IMF1, the RMSF of IMF1, the M of IMF2; the first five important indicators for operating state 3 are the RMS of IMF1, the PK of IMF1, the FC of IMF2, the RVF of IMF2, and the RMS of IMF2. The above data shows that the vibration signal fault information is essentially found in the first two IMFs and the vibration signal time-frequency entropy. It also justifies choosing different indicators such as TD, FD and TFD as indicators of fault characteristics. In addition, to demonstrate better fault correlation of the ranked optimized feature set, for ease of observation, we plotted the first three features of the five important indicators mentioned above and compared them with the unoptimized random features, as indicated in Fig. 17. Most of the samples of the feature set that were LS ranked optimized in the three conditions in the

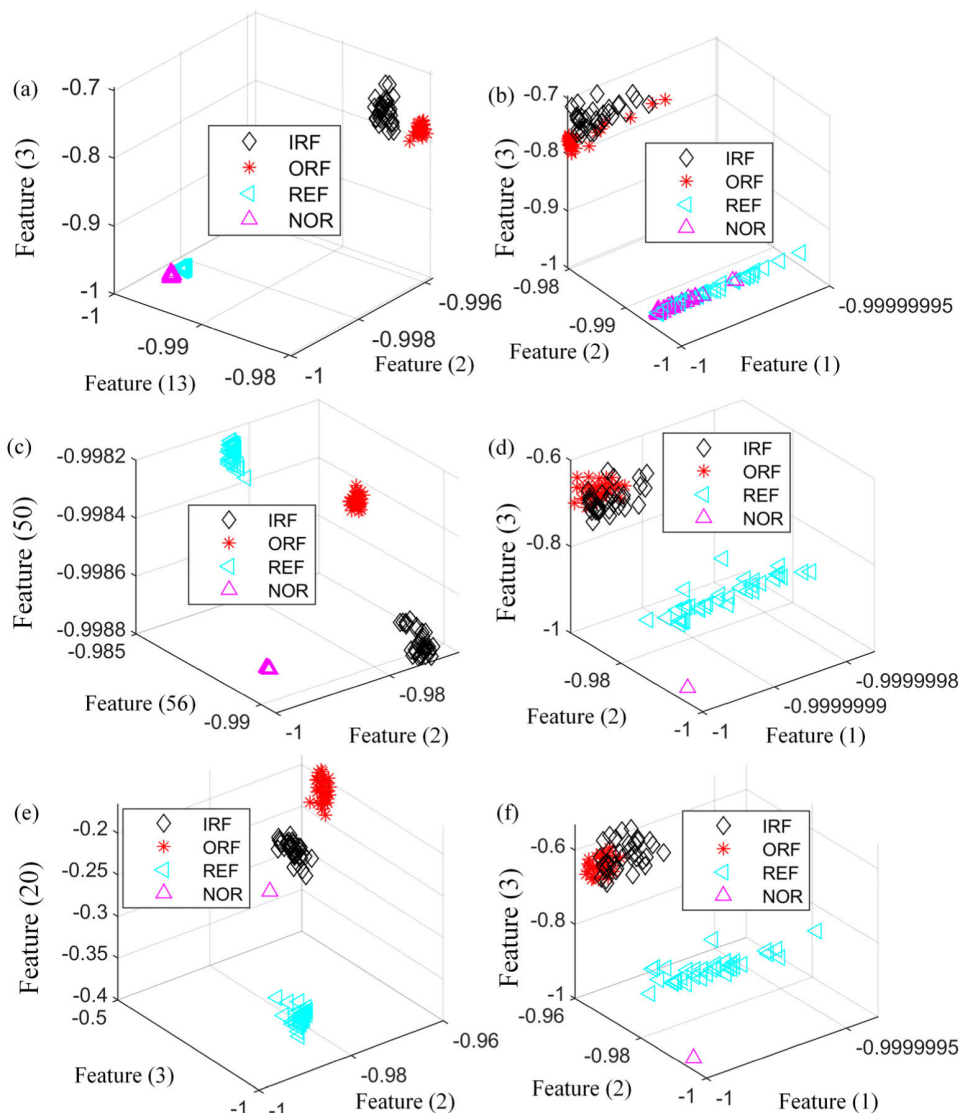


FIGURE 17. The comparison between the sorted optimized feature set and the random initial feature set: (a) Optimal feature set after sorting in operating state 1 (T2, T13, T3); (b) Random initial feature set of operating state 1 (T1, T2, T3); (c) Optimal feature set after sorting in operating state 2 (T2, T56, T50); (d) Random initial feature set of operating state 2 (T1, T2, T3); (e) Optimal feature set after sorting in operating state 3 (T2, T3, T20); (f) Random initial feature set of operating state 3 (T1, T2, T3).

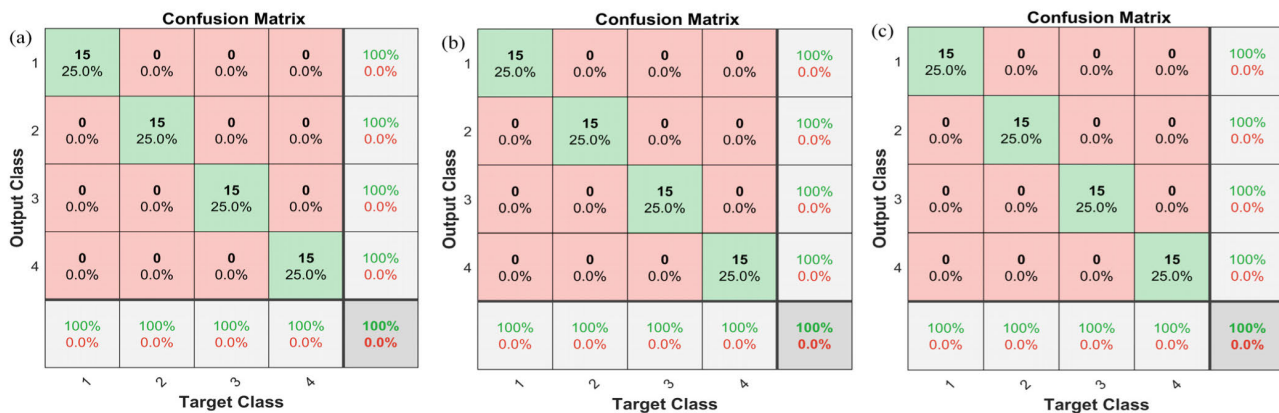


FIGURE 18. GA-BPNN recognition results under different loads: (a) operation at 1200 rpm with a load of 0; (b) operation at 1320 rpm with a load of 0.8 N/m; (c) operation at 1500 rpm with a load of 1.6 N/m.

figure are clustered around their centers and there are clear boundaries between categories. In contrast, the unoptimized initial feature set largely exhibits a dispersion trend, with no obvious clustering centers except for the normal bearing operating conditions. In particular, the inner and outer race failures are mixed together in all three operating states and cannot be effectively recognized.

Finally, the optimal feature vector matrix $\tilde{IT}_{80 \times 5}$ obtained above is input to the GA-BPNN categorizer for training to derive the prediction model for each category. According to the feature matrix $\tilde{IT}_{80 \times 5}$, extract the first five optimal feature values of 60 test samples (for consistency, the paper still extracts 56 feature values of each sample) to obtain the feature matrix $\tilde{OT}_{60 \times 5}$ of the training sample, and input $\tilde{OT}_{60 \times 5}$ into the trained GA-BPNN classifier for automatic recognition. Fig. 18 depicts the recognition rates of the test samples under three different operating conditions. The results demonstrate that the proposed model has good fault recognition and can distinguish well between different loading conditions and different rotational speed conditions, and the fault recognition rates of the test samples are 100%.

VI. CONCLUSION

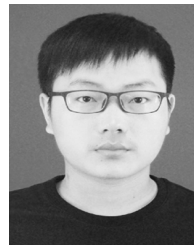
A rolling bearing fault diagnosis model integrating CEALIF, LS, and GA-BPNN is presented. In the proposed model, TD, FD, and TFD information were utilized as damage indicators, where the CEALIF method was employed for signal pre-processing. Then with the help of LS and GA-BPNN, the optimal feature vector was adopted to examine the rolling bearing failure category. The experimental results demonstrate the superior capabilities of the model in determining the different typologies of rolling bearing failures as well as the different degrees of damage. The main aspects of the work are as follows: (1) the noise-assisted CEALIF method was proposed to suppress the ALIF modal aliasing problem, and the simulation signal part showed that the CEALIF method was more stable in processing signals containing high-frequency intermittent signals, which was an effective data analysis technique; (2) the TD, FD and TFD features were extracted simultaneously, which reflected the fault characteristics in many aspects and improved the reliability of the diagnostic results; (3) the LS was introduced to rank the obtained feature parameters according to their importance for selection and reduce redundant features to improve the classification performance; (4) GA is employed to address the shortcomings of the tendency of BP neural networks to fall into local optimality as well as slow convergence to obtain the best BPNN performance. In further investigations, more alternative characteristic parameters and more advanced damage indicators will be proposed to provide enhanced classification performance, which can guide further studies in bearing failure diagnosis.

REFERENCES

[1] A. Glowacz, "Fault diagnosis of single-phase induction motor based on acoustic signals," *Mech. Syst. Signal Process.*, vol. 117, pp. 65–80, Feb. 2019.

- [2] T. Berredjem and M. Benidir, "Bearing faults diagnosis using fuzzy expert system relying on an improved range overlaps and similarity method," *Expert Syst. Appl.*, vol. 108, pp. 134–142, Oct. 2018.
- [3] Y. Lv, R. Yuan, and G. Song, "Multivariate empirical mode decomposition and its application to fault diagnosis of rolling bearing," *Mech. Syst. Signal Process.*, vol. 81, pp. 219–234, Dec. 2016.
- [4] K. Zhu, L. Chen, and X. Hu, "A multi-scale fuzzy measure entropy and infinite feature selection based approach for rolling bearing fault diagnosis," *J. Nondestruct. Eval.*, vol. 38, no. 4, pp. 1–13, Dec. 2019, doi: 10.1007/s10921-019-0623-4.
- [5] J. Li, J. Zhang, M. Li, and Y. Zhang, "A novel adaptive stochastic resonance method based on coupled bistable systems and its application in rolling bearing fault diagnosis," *Mech. Syst. Signal Process.*, vol. 114, pp. 128–145, Jan. 2019.
- [6] X. Li, Y. Yang, H. Pan, J. Cheng, and J. Cheng, "Non-parallel least squares support matrix machine for rolling bearing fault diagnosis," *Mechanism Mach. Theory*, vol. 145, Mar. 2020, Art. no. 103676.
- [7] W. Huang, J. Cheng, and Y. Yang, "Rolling bearing fault diagnosis and performance degradation assessment under variable operation conditions based on nuisance attribute projection," *Mech. Syst. Signal Process.*, vol. 114, pp. 165–188, Jan. 2019.
- [8] K. Satpathi, Y. M. Yeap, A. Ukil, and N. Geddada, "Short-time Fourier transform based transient analysis of VSC interfaced point-to-point DC system," *IEEE Trans. Ind. Electron.*, vol. 65, no. 5, pp. 4080–4091, May 2018.
- [9] Y. Zhang, K. Xing, R. Bai, D. Sun, and Z. Meng, "An enhanced convolutional neural network for bearing fault diagnosis based on time–frequency image," *Measurement*, vol. 157, Jun. 2020, Art. no. 107667.
- [10] W. Qiao and Z. Yang, "Forecast the electricity price of U.S. Using a wavelet transform-based hybrid model," *Energy*, vol. 193, Feb. 2020, Art. no. 116704.
- [11] W. Deng, S. Zhang, H. Zhao, and X. Yang, "A novel fault diagnosis method based on integrating empirical wavelet transform and fuzzy entropy for motor bearing," *IEEE Access*, vol. 6, pp. 35042–35056, 2018.
- [12] N. E. Huang and Z. Wu, "A review on Hilbert–Huang transform: Method and its applications to geophysical studies," *Rev. Geophys.*, vol. 46, no. 2, p. RG2006, 2008.
- [13] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," *Adv. Adapt. Data Anal.*, vol. 1, no. 1, pp. 1–41, Jan. 2009.
- [14] H. Li, T. Liu, X. Wu, and Q. Chen, "Application of EEMD and improved frequency band entropy in bearing fault feature extraction," *ISA Trans.*, vol. 88, pp. 170–185, May 2019.
- [15] J.-R. Yeh, J.-S. Shieh, and N. E. Huang, "Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method," *Adv. Adapt. Data Anal.*, vol. 2, no. 2, pp. 135–156, Apr. 2010.
- [16] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2011, pp. 4144–4147.
- [17] Y. Cheng, Z. Wang, B. Chen, W. Zhang, and G. Huang, "An improved complementary ensemble empirical mode decomposition with adaptive noise and its application to rolling element bearing fault diagnosis," *ISA Trans.*, vol. 91, pp. 218–234, Aug. 2019.
- [18] L. Lin, Y. Wang, and H. Zhou, "Iterative filtering as an alternative algorithm for empirical mode decomposition," *Adv. Adapt. Data Anal.*, vol. 01, no. 04, pp. 543–560, Oct. 2009.
- [19] A. Cicone and P. Dell'Acqua, "Study of boundary conditions in the iterative filtering method for the decomposition of nonstationary signals," *J. Comput. Appl. Math.*, vol. 373, Aug. 2020, Art. no. 112248.
- [20] C.-C. Hsu, C.-L. Yeh, W.-K. Lee, H.-T. Hsu, K.-K. Shyu, L. P.-H. Li, T.-Y. Wu, and P.-L. Lee, "Extraction of high-frequency SSVEP for BCI control using iterative filtering based empirical mode decomposition," *Biomed. Signal Process. Control*, vol. 61, Aug. 2020, Art. no. 102022.
- [21] D. P. Dash, M. H. Kolekar, and K. Jha, "Multi-channel EEG based automatic epileptic seizure detection using iterative filtering decomposition and hidden Markov model," *Comput. Biol. Med.*, vol. 116, Jan. 2020, Art. no. 103571.
- [22] R. R. Sharma, P. Varshney, R. B. Pachori, and S. K. Vishvakarma, "Automated system for epileptic EEG detection using iterative filtering," *IEEE Sensors Lett.*, vol. 2, no. 4, pp. 1–4, Dec. 2018.
- [23] A. Cicone, J. Liu, and H. Zhou, "Adaptive local iterative filtering for signal decomposition and instantaneous frequency analysis," *Appl. Comput. Harmon. Anal.*, vol. 41, no. 2, pp. 384–411, Sep. 2016.

- [24] A. Cicone, J. Liu, and H. Zhou, "Hyperspectral chemical plume detection algorithms based on multidimensional iterative filtering decomposition," *Phil. Trans. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 374, no. 2065, Apr. 2016, Art. no. 20150196.
- [25] A. Cicone, C. Garoni, and S. Serra-Capizzano, "Spectral and convergence analysis of the discrete ALIF method," *Linear Algebra Appl.*, vol. 580, pp. 62–95, Nov. 2019.
- [26] X. An, H. Zeng, and C. Li, "Demodulation analysis based on adaptive local iterative filtering for bearing fault diagnosis," *Measurement*, vol. 94, pp. 554–560, Dec. 2016.
- [27] I. Mitiche, G. Morison, A. Nesbitt, M. Hughes-Narborough, B. Stewart, and P. Boreham, "Classification of partial discharge signals by combining adaptive local iterative filtering and entropy features," *Sensors*, vol. 18, no. 2, p. 406, Jan. 2018.
- [28] X. An, H. Zeng, W. Yang, and X. An, "Fault diagnosis of a wind turbine rolling bearing using adaptive local iterative filtering and singular value decomposition," *Trans. Inst. Meas. Control*, vol. 39, no. 11, pp. 1643–1648, Nov. 2017.
- [29] D. Yang, B. Wang, G. Cai, and J. Wen, "Oscillation mode analysis for power grids using adaptive local iterative filter decomposition," *Int. J. Electr. Power Energy Syst.*, vol. 92, pp. 25–33, Nov. 2017.
- [30] X. He, D. Cai, and P. Niyogi, "Laplacian score for feature selection," in *Proc. Adv. Neural Inf. Process. Syst.*, 2006, pp. 507–514.
- [31] X. Li, Y. Yang, H. Shao, X. Zhong, J. Cheng, and J. Cheng, "Symplectic weighted sparse support matrix machine for gear fault diagnosis," *Measurement*, vol. 168, Jan. 2021, Art. no. 108392.
- [32] H. Zhiyi, S. Haidong, J. Lin, C. Junsheng, and Y. Yu, "Transfer fault diagnosis of bearing installed in different machines using enhanced deep auto-encoder," *Measurement*, vol. 152, Feb. 2020, Art. no. 107393.
- [33] P. Kundu, A. K. Darpe, and M. S. Kulkarni, "An ensemble decision tree methodology for remaining useful life prediction of spur gears under natural pitting progression," *Struct. Health Monitor.*, vol. 19, no. 3, pp. 854–872, May 2020.
- [34] K. Zhu, L. Chen, and X. Hu, "Rolling element bearing fault diagnosis by combining adaptive local iterative filtering, modified fuzzy entropy and support vector machine," *Entropy*, vol. 20, no. 12, p. 926, Dec. 2018.
- [35] F. Wang, Z. Chen, and G. Song, "Monitoring of multi-bolt connection looseness using entropy-based active sensing and genetic algorithm-based least square support vector machine," *Mech. Syst. Signal Process.*, vol. 136, Feb. 2020, Art. no. 106507.
- [36] R. Yuan, Y. Lv, H. Li, and G. Song, "Robust fault diagnosis of rolling bearings using multivariate intrinsic multiscale entropy analysis and neural network under varying operating conditions," *IEEE Access*, vol. 7, pp. 130804–130819, 2019.
- [37] S. Kumar, A. Kumar, and R. K. Jha, "A novel noise-enhanced back-propagation technique for weak signal detection in Neyman–Pearson framework," *Neural Process. Lett.*, vol. 50, no. 3, pp. 2389–2406, Dec. 2019.
- [38] S. D. Desai, S. Giraddi, P. Narayankar, N. R. Pudakalakatti, and S. Sulegaon, "Back-propagation neural network versus logistic regression in heart disease classification," in *Advanced Computing and Communication Technologies*. Cham, Switzerland: Springer, 2019, pp. 133–144.
- [39] D. S. Weile and E. Michielssen, "Genetic algorithm optimization applied to electromagnetics: A review," *IEEE Trans. Antennas Propag.*, vol. 45, no. 3, pp. 343–353, Mar. 1997.
- [40] Y. Lei, Z. He, and Y. Zi, "EEMD method and WNN for fault diagnosis of locomotive roller bearings," *Expert Syst. Appl.*, vol. 38, no. 6, pp. 7334–7341, Jun. 2011.
- [41] J. Zheng, J. Cheng, and Y. Yang, "Partly ensemble empirical mode decomposition: An improved noise-assisted method for eliminating mode mixing," *Signal Process.*, vol. 96, pp. 362–374, Mar. 2014.
- [42] A. Karagiannis and P. Constantinou, "Noise-assisted data processing with empirical mode decomposition in biomedical signals," *IEEE Trans. Inf. Technol. Biomed.*, vol. 15, no. 1, pp. 11–18, Jan. 2011.
- [43] D. Cho, B. Min, J. Kim, and B. Lee, "EEG-based prediction of epileptic seizures using phase synchronization elicited from noise-assisted multivariate empirical mode decomposition," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 8, pp. 1309–1318, Aug. 2017.
- [44] D. Yu, Y. Yang, and J. Cheng, "Application of time–frequency entropy method based on Hilbert–Huang transform to gear fault diagnosis," *Measurement*, vol. 40, nos. 9–10, pp. 823–830, Nov. 2007.
- [45] B. Pang, T. Tian, and G.-J. Tang, "Fault state recognition of wind turbine gearbox based on generalized multi-scale dynamic time warping," *Struct. Health Monitor.*, vol. 2020, Dec. 2020, Art. no. 1475921720978622.
- [46] X.-Y. Jing, S. Li, D. Zhang, J. Yang, and J.-Y. Yang, "Supervised and unsupervised parallel subspace learning for large-scale image recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 10, pp. 1497–1511, Oct. 2012.
- [47] L. Li, G. Song, and J. Ou, "A genetic algorithm-based two-phase design for optimal placement of semi-active dampers for nonlinear benchmark structure," *J. Vib. Control*, vol. 16, no. 9, pp. 1379–1392, Aug. 2010.
- [48] K. Nara, A. Shiose, M. Kitagawa, and T. Ishihara, "Implementation of genetic algorithm for distribution systems loss minimum re-configuration," *IEEE Trans. Power Syst.*, vol. 7, no. 3, pp. 1044–1051, Aug. 1992.
- [49] Y. Lei, Z. He, and Y. Zi, "A new approach to intelligent fault diagnosis of rotating machinery," *Expert Syst. Appl.*, vol. 35, no. 4, pp. 1593–1600, Nov. 2008.



YI ZHANG received the B.S. degree in mechanical engineering from the Department of Mechanical Engineering, Wuhan University of Science and Technology, Wuhan, China, in 2015, where he is currently pursuing the Ph.D. degree in mechanical engineering. His current research interests include fault detection and classification, fault diagnosis, and structural health monitoring. He received the National Award for Distinguished Doctorates from the Wuhan University of Science and Technology.



YONG LV received the B.S. and M.S. degrees from the Department of Mechanical Engineering, Wuhan University of Science and Technology, Wuhan, China, in 1998 and 2000, respectively, and the Ph.D. degree from the University of Science and Technology Beijing, Beijing, China, in 2004. He is currently a Full Professor with the School of Machinery and Automation, Wuhan University of Science and Technology. He has hosted three National Natural Science Foundations of China.

He has hosted more than 30 projects in the field of machinery dynamics and condition monitoring. His current research interests include fault detection and isolation, prognostics and health management, structural health monitoring, and nonlinear dynamics.



MAO GE received the B.S. degree in mechanical engineering and automation from the Wuhan University of Science and Technology, Wuhan, China, in 2015, where he is currently pursuing the Ph.D. degree with the Faculty of Mechanical Engineering. His research interests include mechanical equipment fault diagnosis, high-dimensional signal processing, and nonlinear signal processing.