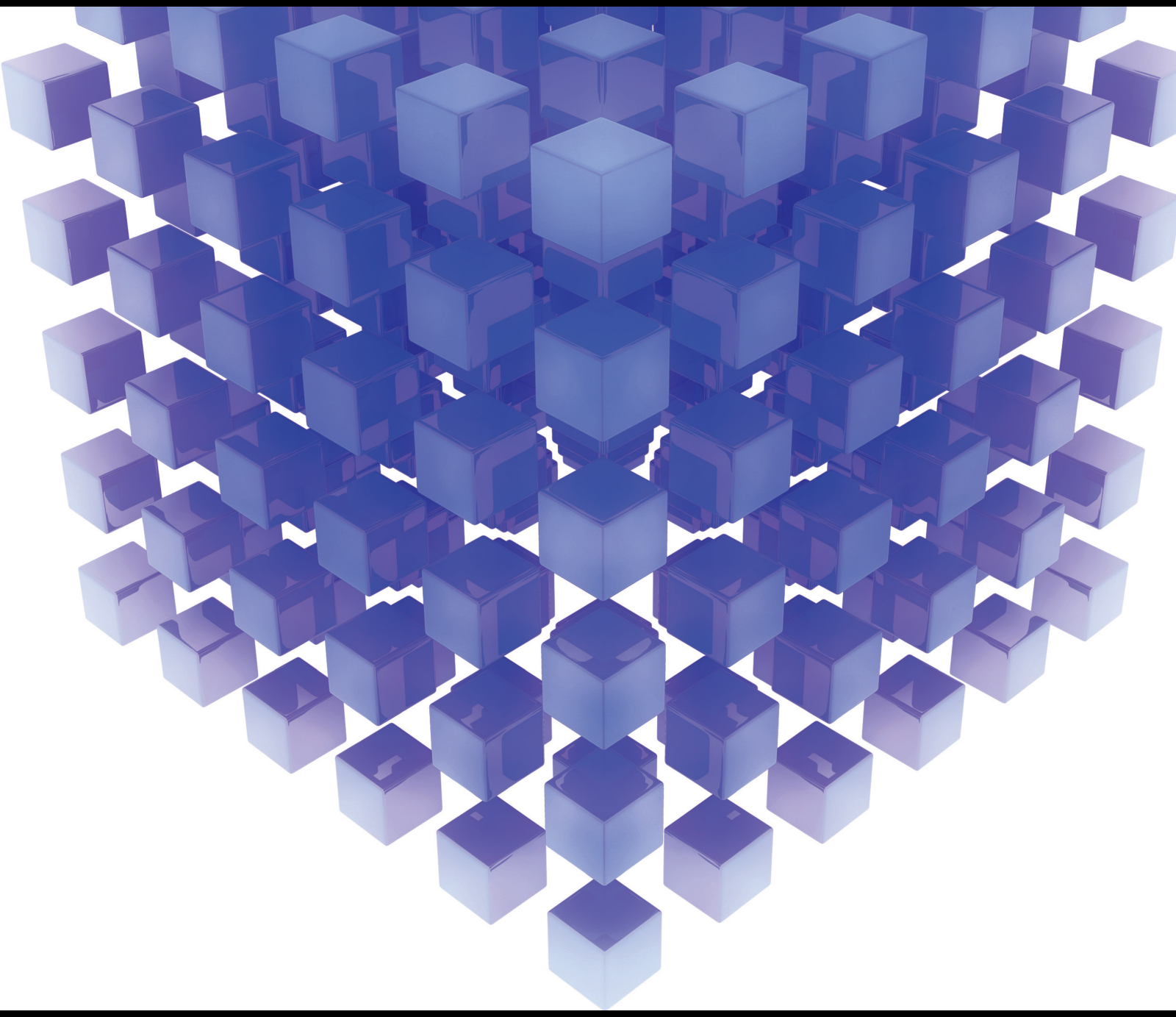
Mathematical Problems in Engineering

Condition Monitoring of Bearing and Gearbox Transmission Systems

Lead Guest Editor: Kesheng Wang
Guest Editors: Ke Feng and Qing Ni



Condition Monitoring of Bearing and Gearbox Transmission Systems

Mathematical Problems in Engineering

Condition Monitoring of Bearing and Gearbox Transmission Systems

Lead Guest Editor: Kesheng Wang
Guest Editors: Ke Feng and Qing Ni

Contents

Hindawi
Mathematical Problems in Engineering
Volume 2022, Article ID 4034477, 17 pages
https://doi.org/10.1155/2022/4034477

Research Article

Rolling Bearing Fault Diagnosis Based on MFDFA-SPS and ELM

Yunfan Yang [1] **and Caiping Xi** [2]

[1] *Ocean College, Jiangsu University of Science and Technology, Zhenjiang 212003, China*
[2] *School of Automation, Jiangsu University of Science and Technology, Zhenjiang 212003, China*

Correspondence should be addressed to Caiping Xi; xicaiping@just.edu.cn

Received 2 May 2022; Revised 14 June 2022; Accepted 15 June 2022; Published 18 July 2022

Academic Editor: Ke Feng

Rolling bearings, as important parts on supporting rotating shafts, frequently suffer from fatigue failures. If these rolling bearing failures are not found in time, it will have a huge impact on the whole mechanical system's operating safety and operating life. To improve the diagnosis of different faults as well as different degrees of faults, a fault diagnosis method based on the multifractal detrended fluctuation analysis (MFDFA) method-singularity power spectrum (SPS) with extreme learning machine (ELM) is proposed. First, MFDFA and SPS analyses are performed on vibration acceleration signals with different faults and different degrees of damage under the same operating conditions, the spectral parameters of stability and quantitative description of differentiation are selected for feature extraction, and then the selected six feature parameters are put into the extreme learning machine for fault classification. The effectiveness of the MFDFA-SPS feature extraction method is demonstrated by analyzing and testing the measured bearing signals. The fault diagnosis accuracy of the bearing fault signals can reach 99.2% based on the MFDFA-SPS with ELM method by using the Case Western Reserve database. The improvements are 6.79% and 18.42% compared to the fault diagnosis methods based on MFDFA with ELM and SPS with ELM. Compared with the methods based on MFDFA-SPS with LSSVM classifier and SVM classifier, the accuracy improvements are 3.54% and 4.25%, respectively. The results show that the method proposed in this paper can achieve the diagnosis of bearing faults and the method based on MFDFA-SPS with ELM is more efficient than the methods based on MFDFA-SPS with LSSVM and SVM classifiers, which is suitable for practical engineering problem-solving.

1. Introduction

Rolling bearing is one of the important components of the mechanical transmission system, due to corrosion and fatigue factors, and easy to cause bearing abrasion, pitting, and cracking failures [1]. Different faults cause different vibration noise and increased rotational resistance, making the whole transmission system life decay or even fail [2]. Therefore, it is important to ensure the normal operation of the bearings in the gearbox or drive shaft and the timely diagnosis and analysis of the structural components after a failure occurs for the normal operation of the whole unit.

Since the concept of "fractal" was first introduced by Mandelbrot in the 1980s [3], this theory of studying nonlinear systems has been widely applied in many fields and has achieved fruitful results. The object of study of fractal theory is not smooth and irregular geometry in nature, and the bearing fault diagnosis method based on fractal theory is one of the more suitable and effective methods for bearing fault identification. Commonly used fractal theory basis includes iterative function system (IFS), fractal Brownian random field, single fractal, extended fractal, multifractal, and multifractal correlation [4, 5]. The difference in internal signal characteristics between faulty vibration signals and normal operation vibration signals is usually used in research for fault diagnosis by extracting fractal geometric change features.

For decades, scholars worldwide have proposed many methods for fault diagnosis of bearings, and the current mainstream diagnosis methods are the temperature analysis method [6] and vibration analysis method [7]. The enriched

fault features such as cyclic power spectra [8] and vibration cyclostationarity [9] parameters can be extracted when using bearing vibration signals for fault monitoring, which enables excellent diagnostic performance. The vibration analysis method mainly performs digital signal acquisition of the faulty bearing vibration signal, and its acquisition method is low cost and contains rich fault information in the vibration signal, which is convenient for real-time diagnosis and analysis.

For the analysis of nonlinear signal characteristics from the perspective of multifractals, Shao et al. [10] used the fractal box dimension method to quantitatively characterize the vibration signals of faulty bearings and found that the fractal dimensions of the faults differed from each other in their manifestations. Liu et al. [11] used detrended fluctuation analysis (DFA) to extract features of wind turbine bearing faults and proved that the fault characteristic parameters can be obtained by multifractal detrended fluctuation analysis (MFDFA). Lin and Chen [12] used the shape and location feature parameters of the multifractal spectrum obtained by the detrended fluctuation analysis and combined with the martingale repulsion to achieve the identification of the bearing damage degree. Chen et al. [13] proposed a fault prediction method based on fractal dimension using an extreme learning machine (ELM). Xiong et al. [14] used multifractal detrended fluctuation analysis methods with PSO optimized LSSVM classification method to achieve fault diagnosis of bearings. Numerous worldwide research papers have demonstrated that fractal theory can be applied to faulty bearing vibration signal processing, and the difference lies in the different effects of using different feature parameters' extraction criteria and classifiers. At present, the classifier based on combining multifractal theory and deep learning is the current niche use diagnosis method, and the diagnosis effect is fast and has a good diagnosis rate with further development trends. ELM and deep learning are complementary, and some applications combine the two to get promising results, such as using CNNs for feature extraction [15] and ELMs for classifiers.

In summary, to further realize the applicability of fractal theory to bearing fault analysis and diagnosis, this paper proposes a rolling bearing composite fault method based on MFDFA-singularity power spectrum (SPS) and ELM. The method is a quantitative description of nonlinear signals using multifractals for feature extraction. ELM is applied for further exploration of multiple classifications, and the validity of the method is verified by the actual measurement of many different types of faulty bearing signals.

This paper is organized as follows: Section 2 describes the specific model of the MFDFA algorithm; Section 3 describes the specific steps of the SPS algorithm; Section 4 presents a brief description of the ELM tool and the way to extract features using the algorithms in Sections 2 and 3; Section 5 describes the source of data, the specific feature parameter distribution, and the results of the experiments; the main findings and conclusions of the article are summarized in Section 6.

2. The Multifractal Detrended Fluctuation Analysis

The MFDFA algorithm implementation steps are described as follows [12]:

Step 1: for a times series $x(t)$ of length N, $t = 1, 2, \ldots, N$, construct a new sequence $y(t)$:

$$y(t) = \sum_{i=1}^{t} \left\{ x(i) - \frac{1}{N} \sum_{j=1}^{N} x(j) \right\}, \quad t = 1, 2, \ldots, N. \tag{1}$$

Step 2: partition the sequence $y(t)$ into $N_S = [N/S]$ mutually disjoint equal-length subintervals. There are s data points in each interval, $\max\{m+2, 10\} \le s \le N/10$. Since the sequence length N is not necessarily an integer multiple of s, then $y(t)$ will have segments remaining, and to make the sequence of the last small segment processed, the partitioning process can be repeated from the tail of the sequence $y(t)$, thus obtaining $2N_s$ equal-length subintervals. m represents the order of polynomial fit, generally taken as $m = 1, 2, 3, \ldots$. Here, MFDFA can be denoted as MFDFA-m, and when m takes the value of 1, the method is denoted as MFDFA-1.

The mth-order polynomial $y_v^{(m)}(t)$ in the vth subinterval $(v = 1, 2, \ldots, 2N_s)$ is fitted by the least-squares method, and the residual corresponding to the label t in this interval is calculated as follows:

$$\tilde{y}_v(t) = y(t) - y_v^{(m)}(t), t = (v-1)s + i, \quad i = 1, 2, \ldots, s. \tag{2}$$

Step 3: calculate the mean square error separately:

$$F_v(s) = \left[\frac{1}{s} \sum_{i=1}^{s} \tilde{y}_v^2[(v-1)s + i] \right]^{1/2}, \quad v = 1, 2, 3, \ldots, N_s,$$

$$F_v(s) = \left[\frac{1}{s} \sum_{i=1}^{s} \tilde{y}_v^2[N - (v - N_s)s + i] \right]^{1/2}, \quad v = N_s + 1, \ldots, 2N_s. \tag{3}$$

Step 4: calculate the global root mean square value of order q, when $q \neq 0$, with

$$F_q(s) = \left\{ \frac{1}{2N_s} \sum_{v=1}^{2N_q} [F_v(s)]^q \right\}^{1/q}. \tag{4}$$

When $q = 0$, the calculation formula is

$$\ln\left[F_{q=0}(s)\right] = \frac{1}{2N_s} \sum_{v=1}^{2N_q} \ln[F_v(s)]. \tag{5}$$

Step 5: change the value of scale s to obtain the corresponding $F_q(s)$; if the signal sequence has a power-law relationship scale-invariant signal, then

$$F_q(s) \sim s^{H(q)}, \tag{6}$$

where $H(q)$ is the q-order Hurst exponent.

Find the scaling range with a linear relationship between $\log_{10}[F_q(s)]$ and $\log_{10}(s)$. Find the slope of its linear regression line, denoted as $H(q)$, by calculating the first-order polynomial fit coefficient between the scale s and the logarithmic value of the function $F_q(s)$ in the scaling range. If the relationship between $\log_{10}[F_q(s)]$ and $\log_{10}(s)$ is S-shaped curve, the value of $H(q)$ cannot be estimated.

Step 6: the correlation between $H(q)$ and the multifractal mass exponent $\tau(q)$ can be expressed as

$$\tau(q) = qH(q) - D_f, \tag{7}$$

where D_f denotes the topological dimension of the multifractal signal, and for a one-dimensional time series of signals, $D_f = 1$.

Step 7: the singularity intensity function $\alpha(q)$ and the singularity dimension $f(\alpha)$ can be obtained by the Legendre transformation:

$$\begin{aligned} \alpha(q) &= \frac{\mathrm{d}\tau(q)}{\mathrm{d}q}, \\ f(\alpha) &= \inf_q[q\alpha(q) - \tau(9)]. \end{aligned} \tag{8}$$

In the process of changing q value from negative to positive, if $H(q)$ remains almost constant, the relationship between q-order mass exponent $\tau(q)$ and q is linear; then, the signal has single fractal characteristics and the width of the singularity spectrum is small and tends to 0. On the contrary, if $H(q)$ changes to a large extent, the relationship between q-order mass exponent $\tau(q)$ and q is not linear; then, the signal has multifractal characteristics and the width of the singularity spectrum is large. In this paper, we use MFDFA-1 method with one-direction partition process, just do the partitioning process from head to tail of the sequence $y(t)$, thus obtaining N_s equal-length subintervals.

3. The Multifractal Singularity Power Spectrum Algorithm (SPS)

3.1. SPS Algorithm. Suppose a set of signals or time series is a continuous signal.

Defining $x_\alpha(t) = \{x(t)\colon \alpha(x(t)) = \alpha\}$ as a fractal subband signal or a fractal subset of signal $x(t)$, the time-indexed set corresponding to $x_\alpha(t)$ constitutes the set $t(\alpha) = \{t, \alpha(x(t)) = \alpha\}$; then, $x_\alpha(t)$ is denoted as

$$x_\alpha(t) = \begin{cases} x(t), & t \in t(\alpha), \\ 0, & t \in \dfrac{[0,T]}{t(\alpha)}, \end{cases} \tag{9}$$

where $\alpha(x(t))$ is the singularity exponent of $x(t)$ at the time t. According to the theory of multifractals, $x(t)$ is a thick subset supported by a tight set $\cup_\alpha x_\alpha(t)$. Meanwhile, $x_{\alpha 1}(t), x_{\alpha 2}(t), x_{\alpha 3}(t), \ldots$ form the segments of $x(t)$, where $x_{\alpha i}(t)$ are mutually disjoint, and we obtain $x_{\alpha i}(t) \cap x_{\alpha j}(t) = \varphi, i \neq j$.

All measurable subsets constitute the set system $I = \{x_{\alpha 1}(t), x_{\alpha 2}(t), x_{\alpha 3}(t), \ldots\}$, and I is a topological measurable set. There exists a measurable function $f\colon I \longrightarrow \mathfrak{R}^+$, and (X, I, f) constitutes a measurable space.

Based on the singularity decomposition, we can obtain the singular subset $x_T^\alpha(t)$ of $x_T(t)$ under the fractal truncation signal. The fractal energy measure of $x_T^\alpha(t)$ can be expressed as

$$P_x^T = \int_\alpha P_x^T(\alpha)\mathrm{d}\alpha, \tag{10}$$

where $P_x^T(\alpha)$ is the singularity measure of each fractal subset $x_T^\alpha(t)$ and the singularity power spectral density of $x_T(\alpha)$. It reflects the distribution of the fractal power of $x_T(t)$ with the singularity exponent. The fractal power analysis of the singularity fractal subset $x_T^\alpha(t)$ gives the singularity power spectrum of $x_T(t)$ as

$$p_x^T(\alpha) = \frac{1}{T(x_T(t))} \int_{-T/2}^{+} \frac{\left|x_T^\alpha(t)\right|^2}{\sqrt{1 + tg^2\theta_t}} \mathrm{d}H(x_T^\alpha(t)). \tag{11}$$

The fractal power measurement of the signal sequence $x_T(t)$ and the singularity power spectrum distribution satisfy the following relationship:

$$\begin{aligned} p_x^T &= \int_\alpha p_x^T(\alpha)\mathrm{d}\alpha \\ &= \int_\alpha \int_{-T/2}^{+T/2} \frac{1}{T(x_T(t))} \frac{\left|x_T^\alpha(t)\right|^2}{\sqrt{1 + tg^2\theta_t}} \mathrm{d}H(x_T^\alpha(t))\mathrm{d}\alpha. \end{aligned} \tag{12}$$

Therefore, the SPS distribution function of the signal $x(t)$ can be defined as a limit form:

$$\begin{aligned} P_x(\alpha) &= \lim_{T \longrightarrow \infty} p_x^T \\ &= \int_{-\infty}^{+\infty} \lim_{T \longrightarrow \infty} \frac{\left|x_T^\alpha(t)\right|^2}{T(x_T(t))\sqrt{1 + tg^2\ \theta t}} H(x_T^\alpha(\mathrm{d}t)), \end{aligned} \tag{13}$$

where θ_i is the local orientation angle of the truncated signal $x_T(t)$ (equivalent to the tangent direction of a conventional curve) and $dH(x_T^\alpha(t))$ denotes the differential of the Hausdorff measure of $x_T(t)$. Based on the modulus squared of the signal, the local azimuth of the fractal time element, and the Hausdorff measure, equation (15) represents the

power distribution of the fractal signal concerning the singularity exponent. Thus, the fractal energy measurement and the singularity power spectrum satisfy the relationship

$$P_x = \int_\alpha P_x(\alpha)\mathrm{d}\alpha. \tag{14}$$

For fractal signals with a period T or finite time support $x(t), t \in [-T/2, T/2]$. Singularity energy spectral density and singularity power spectral density functions can be defined as

$$\begin{cases} E_x(\alpha) = \displaystyle\int_{-T/2}^{T/2} \frac{|x_\alpha(t)|^2}{\sqrt{1+tg^2\theta_t^\alpha}} H(x(\mathrm{d}t)), \\ P_x(\alpha) = \dfrac{1}{T(x(t))} \displaystyle\int_{-T/2}^{T/2} \frac{|x_\alpha(t)|^2}{\sqrt{1+tg^2\theta_t^\alpha}} H(x(\mathrm{d}t)). \end{cases} \tag{15}$$

It can be obtained that the singularity power spectrum reflects the fractal energy per unit singularity exponent interval, and it reflects how the fractal energy is distributed with the variation of the singularity exponent.

3.2. Specific Steps of the SPS Algorithm. The multifractal singularity power spectrum (SPS) algorithm is a quantitative analysis method to study the correlation of a sequence [16]. The steps of the SPS algorithm are as follows:

Step 1: calculate the instantaneous singularity exponent: for a given time series $x(t)$, estimate the instantaneous singularity exponent $\alpha_x(k_n)$ of the series.

Step 2: construct the singularity subset signal: according to the instantaneous singularity exponent $\alpha_x(k_n)$, set the singularity variable change interval $[\alpha_{\min}, \alpha_{\max}]$, where $\alpha_{\min} \le \alpha_x(k_n) \le \alpha_{\max}$, and divide the interval into N_α discrete singularity exponent sets $\{\alpha_q\}, q = 1, 2, \ldots, N_\alpha + 1$ according to the singularity exponent interval equivalence, where $\alpha_1 = \alpha_{\min}$ and $\alpha_{N_\alpha+1} = \alpha_{\max}$. N_α is the number of discretized singularity indices.

Step 3: according to the interval segment of the singularity exponent after discretization, the singularity decomposition of this time series is performed; the data points whose singularity indices fall in the interval $(\alpha_i, \alpha_{i+1}], i = 1, 2, \ldots, N_\alpha - 1$, are specified as subsets belonging to the same singularity exponent α_i, and the data points whose singularity indices fall in the interval $[\alpha_i, \alpha_{i+1}], i = N_\alpha$, are specified as subsets belonging to the same singularity exponent α_i; thus, it is possible to obtain the fractal subsets with the same singularity exponents. We denote the subset of the same singularity exponent α_q by $x(\alpha_q) = \{x(i), \alpha_x(i) = \alpha_q\}$ and the total sequence by $x = \{x(i)\} = \cup_{q=1}^{N_\alpha}\{x(\alpha_q)\}$, $q = 1, 2, \ldots, N_\alpha$, where $x(\alpha_q)$ denotes the set of all moments in the time series and $\{x(i)\}$ has the instantaneous singularity exponent value α_q. The number of points in the subset $x(\alpha_q)$ is related to the specific calculation, and the subset corresponding to each singularity exponent α_q is denoted as $x_{\alpha_q}(t)$.

Step 4: calculate the average power $P_x(\alpha_q)$ of the singularity subset with the same singularity exponent $\alpha = \alpha_q$: based on doing the singularity subset decomposition on $x(t)$, for the singularity subset $x_\alpha(t)$, for a given singularity exponent $\alpha = \alpha_q$, calculate the power $P_x(\alpha_q)$ of the subset (time series) with the same singularity exponent; for the convenience of presentation, denote this time series with its corresponding singularity subset $\{x(\alpha_q)\}$ as $x_{\alpha_q}(t)$ and its discrete form as $\{x_{\alpha_q}(k)\}, k = 1, 2, \ldots, N_{\alpha_q}$, where N_{α_q} is, respectively, the number of discrete point sets contained in the time series $x(t)$, as the average power $P(\alpha_q) = 1/N_{\alpha_q}\sum_k \|x_{\alpha_q}(k)\|^2$.

Step 5: calculate the singular power spectrum: iterate over all the singularity indices to obtain the singularity power spectrum $P_x(\alpha)$.

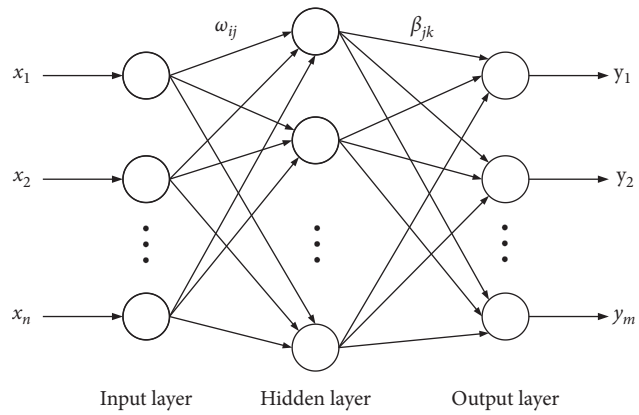


Figure 1: Structure of basic ELM.

4. Extreme Learning Machine

The ELM algorithm [17] has the advantages of strong generalization ability and fast learning speed, which overcomes the disadvantages of a traditional neural network learning algorithm (such as BP neural network algorithm) that requires continuous iteration, many training times, low learning efficiency, and slow convergence speed, and is a new type of single implicit layer feedforward neural network. The ELM consists of three parts: input layer, hidden layer, and output layer, and the specific structure is shown in Figure 1.

ω_{ij} is the value of the connection between the input layer and the hidden layer, which is denoted as

$$\omega_{ij} = \begin{bmatrix} \omega_{11} & \omega_{12} & \cdots & \omega_{1n} \\ \omega_{21} & \omega_{22} & \cdots & \omega_{2n} \\ \vdots & \vdots & & \vdots \\ \omega_{l1} & \omega_{l2} & \cdots & \omega_{ln} \end{bmatrix}. \tag{16}$$

β_{jk} is the value of the connection between the hidden layer and the output layer, which is denoted as

$$\beta_{jk} = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1m} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2m} \\ \vdots & \vdots & & \vdots \\ \beta_{l1} & \beta_{l2} & \cdots & \beta_{lm} \end{bmatrix}, \tag{17}$$

where n is the number of neurons in the input layer, l is the number of neurons in the hidden layer, and m is the number of neurons in the output layer.

The threshold of the implicit layer neuron b is expressed as

$$b = [\, b_1 \;\; b_2 \;\; \cdots \;\; b_l \,]^{\mathrm{T}}. \tag{18}$$

Let $g(x)$ be the hidden layer neuron excitation function, and the output T can be expressed as the following equation:

$$T = [\, t_1 \;\; t_2 \;\; \cdots \;\; t_m \,]^T, \tag{19}$$

$$t_j = \begin{bmatrix} t_{1j} \\ t_{2j} \\ \vdots \\ t_{mj} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{l} \beta_{i1}(\omega_i x_j + b_i) \\ \sum_{i=1}^{l} \beta_{i2}(\omega_i x_j + b_i) \\ \vdots \\ \sum_{i=1}^{l} \beta_{im}(\omega_i x_j + b_i) \end{bmatrix}, \tag{20}$$

$$\omega_i = [\, \omega_{i1} \;\; \omega_{i2} \;\; \cdots \;\; \omega_{in} \,], x_j = [\, x_{1j} \;\; x_{2j} \;\; \cdots \;\; x_{nj} \,]^T. \tag{21}$$

Equation (22) can also be expressed as $H\beta = T'$, where T' is the transpose matrix of T and H is the hidden output array, and the expression is

$$H = \begin{bmatrix} g(\omega_1 x_1 + b_1) & g(\omega_2 x_1 + b_2) & \cdots g(\omega_l x_1 + b_l) \\ g(\omega_1 x_2 + b_1) & g(\omega_2 x_2 + b_2) & \cdots g(\omega_l x_2 + b_l) \\ \vdots & \vdots & \vdots \\ g(\omega_1 x_Q + b_1) & g(\omega_2 x_Q + b_2) & \cdots g(\omega_l x_Q + b_l) \end{bmatrix}. \tag{22}$$

The system of equations' least squares β is

$$\min_\beta \| H\beta - T' \|. \tag{23}$$

The solution is $\beta' = H^\dagger T'$, where $H^\dagger$ is the Moore–Penrose generalized inverse of the output matrix of the hidden layer.

4.1. Extraction of Fractal Characteristic Features in Bearing Signals. Statistical physics can be used to obtain a partition function for the set of probabilities of regular and irregular fractals, which uses the $q-$th power of the probability as a weight to distinguish subsets with different sizes of probabilities. In some cases, the collocation function has scale invariance of the form $\varepsilon^{\tau(q)}$ over a certain range (and over an infinite range in the case of regular fractals). Figure 2 shows the sketch of the multifractal spectrum used to extract features, Figure 3 shows the sketch of the Hurst exponents, and Figure 4 shows the sketch of the singularity power spectrum.

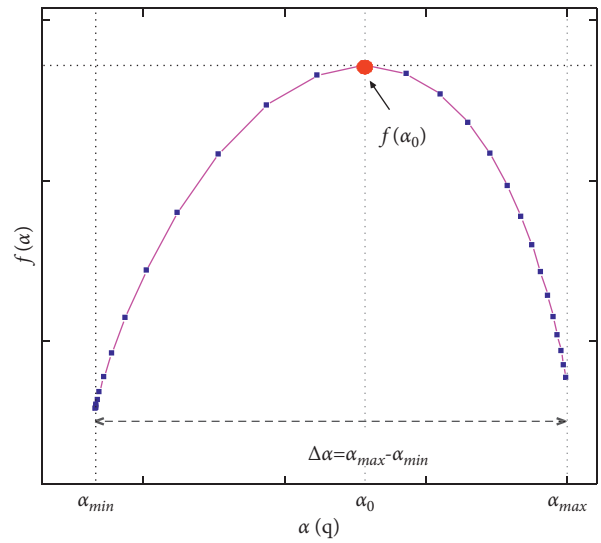


Figure 2: Multifractal spectrum.

The generalized fractal dimension D_f can be obtained from $\tau(q)$, and the multifractal spectrum $f(\alpha)$ is obtained by the Legendre transform. $\alpha_{\max}$ and $f(\alpha_{\max})$ of the multifractal spectrum reflect the properties of the smallest subset of probabilities. $\alpha_{\min}$ and $f(\alpha_{\min})$ reflect the properties of the largest subset of probabilities. $f(\alpha_0)$ reflects the properties of the most contingent subset. The width $\Delta\alpha = \alpha_{\max} - \alpha_{\min}$ of the multifractal spectrum reflects the size of the probability distribution range.

The more inhomogeneous the probability distribution is, the wider the corresponding $f(\alpha)$ curve is and therefore the larger the $\Delta\alpha$ value becomes. In the practical problem of bearing fault signals such as having irregular distribution in space, the quantitative characterization of the inhomogeneous distribution of physical quantities can be achieved with the specific parameters of the multiple fractal spectrum, so it is suitable for the feature extraction of different kinds of bearing faults and different degrees of damage to achieve the diagnosis and identification of faults.

The value of the weighting factor q is also worth discussing when calculating the multifractal spectrum of the measured signal. When calculating the multifractal spectrum according to the multifractal detrended fluctuation analysis, different values of q are used to divide the signal into regions with different levels for the study. Theoretically, the larger the range of q is $(-\infty < q < \infty)$, the more complete the representation of the multifractal spectrum, but in practice, as $|q|$ increases, the computational effort increases exponentially, and the increase of q to a certain level will inevitably cause computer overflow errors. However, if the range of q is too small, the change of $f(\alpha)$ is still large for every increasement of $\Delta|q| = 1$. The obtained is only a part of the multifractal spectrum, which cannot fully reflect the probability distribution of the object of study. In general,

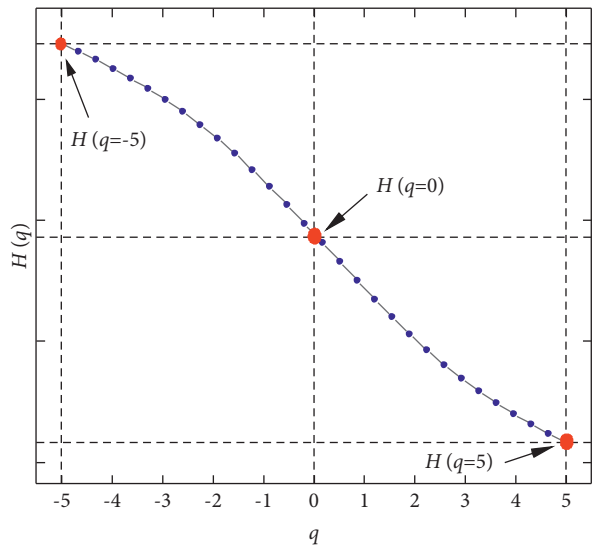


Figure 3: Hurst exponents.

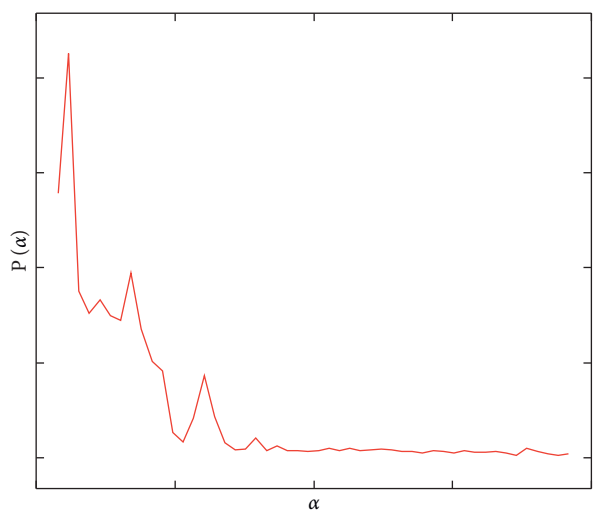


Figure 4: Singularity power spectrum.

when q increases to a certain degree, regardless of the degree of a signal probability distribution, $\alpha(q)$ and $f(\alpha)$ basically do not change with the increase of q, and the range of q can be cut off. In the calculation of the stochastic fractal, a criterion for determining $|q|_{\max}$ can be chosen based on this property, and a coefficient η for the variation of f with the value of q can be defined as follows:

$$\eta = \frac{f_q - f_{q-1}}{f_q - f_{\max}}. \tag{24}$$

When q is infinite, $f_q = f_{q-1}$, $\eta = 0$. In the actual calculation, a value slightly larger than 0 is chosen as a criterion, for example, $\eta < 0.1\%$.

In the fractal analysis of the actual faulty bearing signal, the data are divided into segments of equal length, and each segment contains 2048 points. The efficiency can be ensured when $|q|_{\max} = 5$, and without losing the multiple fractal spectrum features, the spectral parameters are characterized comprehensively.

Mandelbrot in his book "Fractals: Form, Chance, and Dimension" [15] gives a portrayal of the Hurst phenomenon, that is, river flow in which a larger-than-average flood may follow a larger flood and then, suddenly, a below-average flood, followed by more below-average floods; the whole process like there is a cycle, but not for the periodicity, and the standard statistical analysis also reveals significant correlations between observations. To separate the fractal time series from other time series, the Hurst exponent was therefore constructed to estimate the fractal dimension of the time series and the degree of series persistence or inverse persistence. When calculating the Hurst indices for measured bearing failure signals, the Hurst indices for different failures and signals with different fault levels also reflect different distributions, and the Hurst indices corresponding to different q values are taken for feature extraction when the value of $|q|_{\max}$ is given. The features taken in this paper are included in Table 1.

5. Applications

5.1. Dataset Introduction. The bearing signals used in the experiments were obtained from the Bearing Data Center at the Electrical Laboratory of Case Western Reserve University, USA [18, 19]. The experimental platform consists of a 1.5 kW (2 hp) electric motor (electric motor in the left side of Figure 5), a torque sensor/translator (torque transducer and encoder in the middle of Figure 5), and a power test meter (the dynamometer in Figure 5). The rotational speed is 1772 r/min, the sampling frequency is 12 kHz, and the sampling time is 10 s. Single point faults were introduced to the test bearings using electrodischarge machining with fault diameters of 7 mils, 14 mils, 21 mils, 28 mils, and 40 mils (1 mil = 0.001 inches). SKF bearings were used for the 7, 14, and 21 mils diameter faults, and NTN equivalent bearings were used for the 28 mils and 40 mils faults. All experimental data use drive-end acceleration data, which are called vibration acceleration signals. The validity of the proposed method in this paper is verified by selecting the data of the normal state shown in Figure 6(a) and 6 fault states. The fault state data are inner-race 7 mils (slight inner-race fault) shown in Figure 6(b), inner-race 28 mils (serious inner-race fault) shown in Figure 6(c), outer-race 7 mils (slight outer-race fault) shown in Figure 6(d), outer-race 28 mils (serious outer-race fault) shown in Figure 6(e), rolling elements 7 mils (slight ball fault) shown in Figure 6(f), and rolling elements 28 mils fault (serious ball fault) shown in Figure 6(g). Because of the inconsistent length of the bearing signals, each of the data shown in Figure 6 is only for a sampling time of 5 seconds. Each type of data is divided into 50 equal segments with 2048 points per segment. We have 7 kinds of signals in different states. The validation acceleration data of each state are divided into 50 sets, and we can get 350 sets. During the analysis of the MFDFA-1 algorithm for all the selected bearing signals, the subinterval segments are taken as 12 values equally spaced and rounded between 10 and 128, and the values of q are taken as 31 values equally spaced between $[-5, 5]$ according to the rule of equation (24). The data storage format, the feature extraction

Table 1: The six characteristic parameters extracted by MFDFA and SPS.

MFDFA	α_0	The abscissa of the multifractal spectrum vertices
	$\Delta\alpha$	The width of the multifractal spectrum
	$H(q=-5)$	The value of the hurst exponent when $q=-5$
	$H(q=0)$	The value of the hurst exponent when $q=0$
	$H(q=5)$	The value of the hurst exponent when $q=5$
SPS	$\sum P(\alpha)$	Summation of the singularity power spectrum

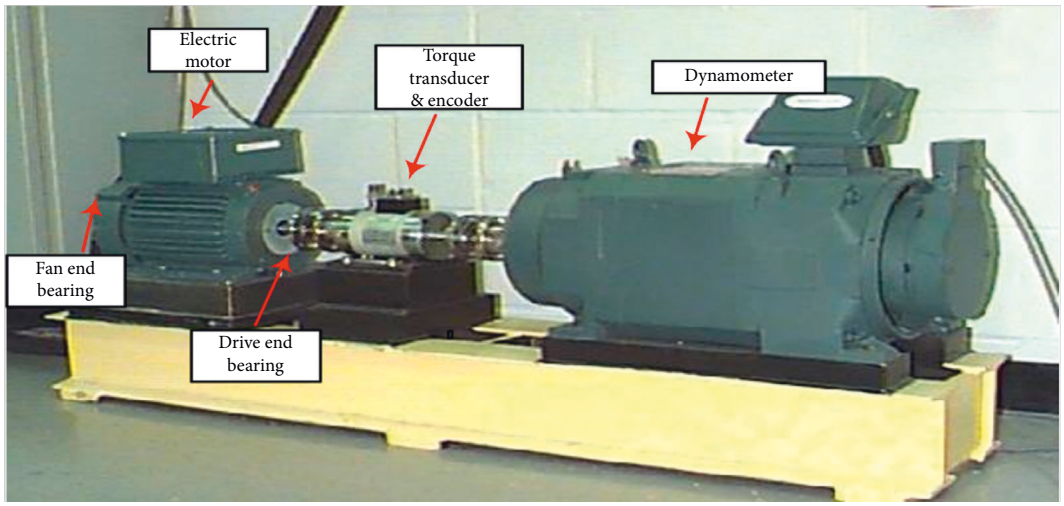


Figure 5: Experimental setup.

algorithms utilized, and the classifier application process all run on the MATLAB2020a platform. In the singularity power spectrum algorithm analysis process, the selection of subintervals is to distinguish the singularity exponent α into 50 equally spaced segments. The time-domain waveform of the vibration signal of each state bearing is shown in Figure 6.

5.2. Steps of Bearing Fault Diagnosis. To ensure the reliability of the experimental results, the bearing signals in the database are tested with the same type of bearings and the same operating conditions except for the degree of damage. The flowchart of the proposed algorithm is shown in Figure 7.

Step 1: the MFDFA and SPS algorithms are used to process the normal bearing signals and the bearing fault signals with different damage faults and different damage degrees to obtain the corresponding multiple fractal spectrum and SPS spectrum

Step 2: feature extraction using the six specific features mentioned in the above section

Step 3: the ELM classifier is used to classify the bearings in seven states, and multiple classification tools are taken to compare the classification effects

5.3. Experimental Results. First, the MFDFA and SPS algorithms were used to analyze the bearing signals in seven states, with ten samples for each state, and the corresponding multifractal spectrum and SPS spectrum were obtained, as shown in Figures 8–10.

Figure 8 shows multifractal spectra of seven signals based on MFDFA-1, where MFDFA-1 means the order of the polynomial of MFDFA is 1. It can be seen that the multifractal spectral lines in different states are hump-shaped and the spectrum widths and the location of the spectrum vertices of the corresponding multifractal spectra are different from each other. The location of the top of the multifractal spectrum of the normal bearing signal is in the rightmost part of Figure 8, and the length of the spectrum width is the shortest. The locations of the top of the spectra of the rest of faulty bearing signals are shifted to the left side of the spectrum of the normal signal. The spectrum widths of the rest of the faulty bearing signals are wider than the one of the normal signals, which indicates that the internal probability density of the faulty bearing signal has changed to different degrees.

The multifractal spectrum width of the normal bearing vibration acceleration signal is narrower than that of the faulty bearing signal, which indicates that the distribution of the probability of the amplitude of the normal operation signal is more even. An increase in the multifractal spectrum width usually implies an increase in the irregular properties within the signal. The purple solid line corresponds to a larger value of $f(\alpha_{\max})$ for slight outer-race fault signal, indicating that the signal has a larger number of subsets with small probability, and the $\alpha_{\min}$ value is the smallest, which corresponds to the subsets with maximum probability.

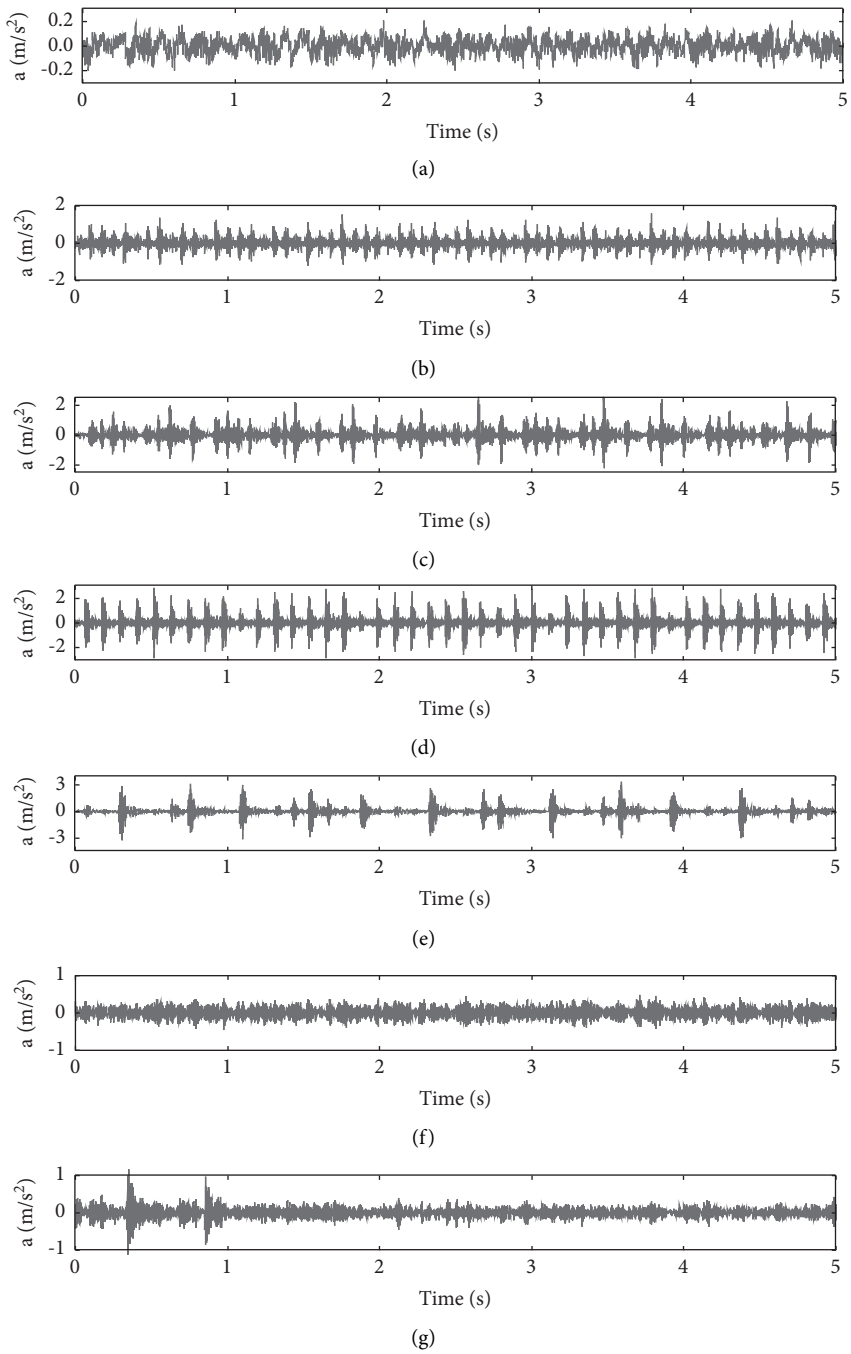


Figure 6: The waveforms of seven vibration acceleration signals in each state. (a) Normal. (b) 7 mils, slight inner-race fault. (c) 28 mils, serious inner-race fault. (d) 7 mils, slight outer-race fault. (e) 28 mils, serious outer-race fault. (f) 7 mils, slight ball fault. (g) 28 mils fault, serious ball fault.

Compared with the serious outer-race fault (olive green dotted line), the two multifractal spectra are in the shape of right hook, which indicates that the signal of the outer-race fault signal is predominated by the subset with small probability. The α_0 positions of the two signals are different, indicating that the most probable subsets are different. The spectrum widths of the two signals are similar to each other, which means that the probability density distributions of the corresponding signals are similar to each other, and the difference lies in the range of probability distributions of the two kinds of signals.

The values of $f(\alpha_{\min})$ corresponding to slight ball fault (blue-green dotted line) and serious ball fault (red solid line) are at a large level. The spectrum widths $\Delta\alpha$ of the spectra of the two kinds of signals are smaller than those of the other faults, indicating that the probability distribution of the ball

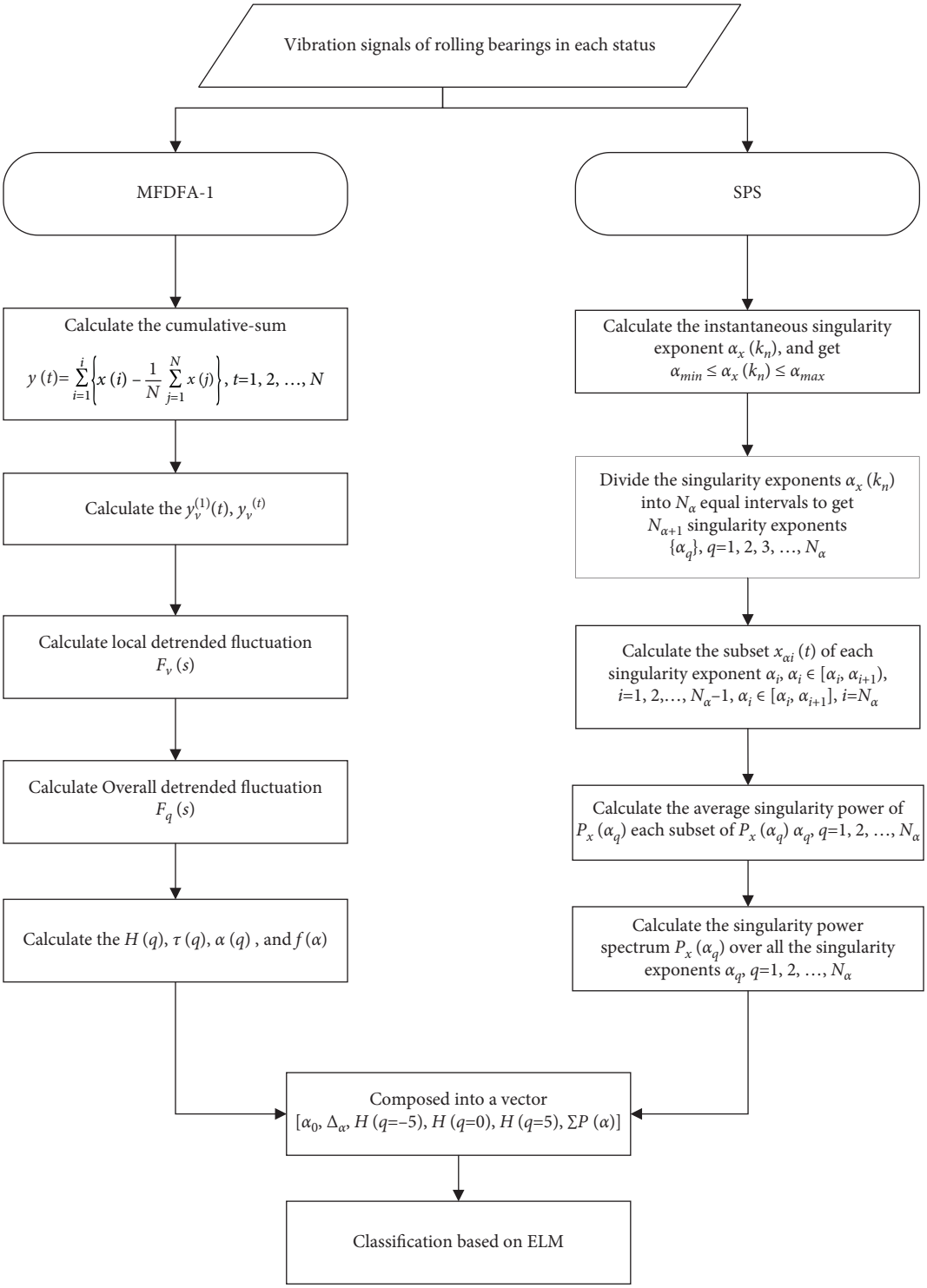


Figure 7: Flowchart of algorithm procedures.

fault signal is more uniform than the probability distributions of the other faults. The two spectrum lines are in the shape of left hook, which means that the large probability subsets dominate the signal.

The spectrum lines of slight inner-race fault (red solid line) and serious inner-race fault (blue solid line) are both left hook-shaped, and the α_0 positions of two curves are also close to each other. The values of $f(\alpha_{\max})$ and $f(\alpha_{\min})$ of serious inner-race fault signal are smaller than those of slight inner-race fault signal. This means that the number of both the small and large probability subsets of the serious inner-race fault signal is smaller. The value of $\Delta\alpha$ for serious inner-race fault signal is larger than that of slight inner-race fault signal, so the probability density distribution of serious

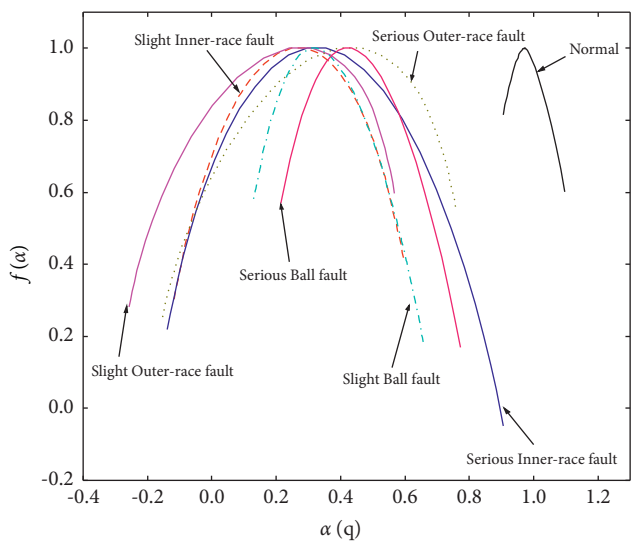


Figure 8: The multifractal spectra of seven signals based on MFDFA-1.

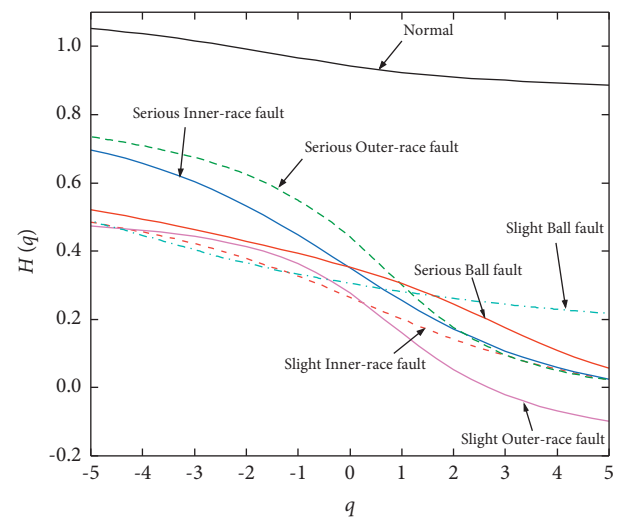


Figure 9: The Hurst exponents of seven signals based on MFDFA-1.

inner-race fault signal is more uneven than that of slight inner-race fault signal.

By comparing the multifractal spectra of signals with the same fault and different degrees of damage, it can be found that the multifractal spectra of signals with different degrees of damage for the same fault have similar shapes (left hook or right hook), but the difference lies in the difference in the uniformity of the probability density distribution (the size of $\Delta\alpha$ value).

Figure 9 shows the q-order Hurst exponent $H(q)$ calculated by MFDFA-1. It can be seen in Figure 9 that when q is changing from -5 to 5 by increasement of $\Delta q = 1/3$, the curve of $H(q)$ for the bearing vibration acceleration signal in the normal state is different from the line of $H(q)$ for the overall faulty bearing vibration acceleration signal distribution, indicating that once the bearing fails, the internal tortuous state of the signal will change. The q-order Hurst exponent $H(q)$ of serious bearing fault signal is generally larger than that of slight bearing faults. $H(q)$ for the normal state is calculated in the range of $q \in [-5, 5]$; it is uniformly distributed. The distinction is evident when comparing with the q-order Hurst exponent $H(q)$ values of faulty signals. When $q < 0$, the $H(q)$ values for the signals of serious inner-race fault, serious outer-race fault, and serious rolling element fault are greater than the $H(q)$ values of the corresponding slight fault signals. When $q > 0$, as q increases, there appear noticeable differences of the $H(q)$ values for signals of slight inner-race fault, slight outer-race fault, and slight ball fault. Therefore, the two points at the endpoints $H(q = -5)$ and $H(q = 5)$ with the points in the middle $H(q = 0)$ can be chosen as the feature values to distinguish the degree of fault for a given value of q. In the range $q \in [-5, -4]$, the $H(q)$ values for the signals of the serious outer-race fault and the serious inner-race fault are more distinguishable from the rest of the fault signals, and the curves of the $H(q)$ values for the signals of the two ball fault signals are difficult to be distinguished from the curves of $H(q)$ values for the slight outer-race fault and the slight inner-race fault in this interval. However, with the increase of q value, the $H(q)$ values for the signal of slight ball fault do not drop as rapidly as the $H(q)$ values for the rest of the fault signals. The $H(q)$ curve of slight ball fault signal is distinguishable from that of the serious ball fault in the interval $q \in [3, 5]$. Similarly, as the value of q increases, the difference between $H(q)$ curves of the slight outer-race fault signal and serious outer-race fault signal is almost the same except in the range $q \in [-5, -2]$. Conversely, the $H(q)$ curves of signals of the two inner-race faults and the serious outer-race faults in the range $q \in [2, 5]$ are difficult to be distinguished from each other. The difference between the $H(q)$ curves of the signals of the two kinds of inner-race faults is greater in the range $q \in [-5, 0]$ than that in the interval $q \in [1, 5]$. When $q > 0$, $H(q)$ reflects the degree of fluctuation of the major trend within the signal, and when $q < 0$, $H(q)$ reflects the degree of fluctuation of the minor trend within the signal. As a result, it can be obtained that the degree of fluctuation of minor trends within the signal of serious degree of bearing fault is usually greater than that of the signal of slight degree of bearing fault when $q \in [-5, 0]$.

Figure 10 shows the singularity power spectra of seven signals based on the SPS method. In Figure 10, the characteristics of the SPS spectra of the faulty bearing signals in different states are not the same and the power spectra at specific different singularity values changed to a larger extent due to the shift in the distribution range of the singularity values. The normal bearing signal has a large singularity exponent distribution and a uniform power distribution. The distribution range of the singularity exponents of the faulty bearing signal changes, and the power distribution becomes uneven compared to the normal state. According to the occurrence of such a phenomenon, extracting the features with a large degree of variation can effectively distinguish the bearing fault and the degree of failure. The specific distribution of the 5 feature parameters $(\alpha_0, \Delta\alpha, H(q = -5), H(q = 0), \text{and } H(q = 5))$ of the 70 samples (all training sets consisted of 10 sets of each state, so we get 70 sets for training) calculated by MFDFA-1 and

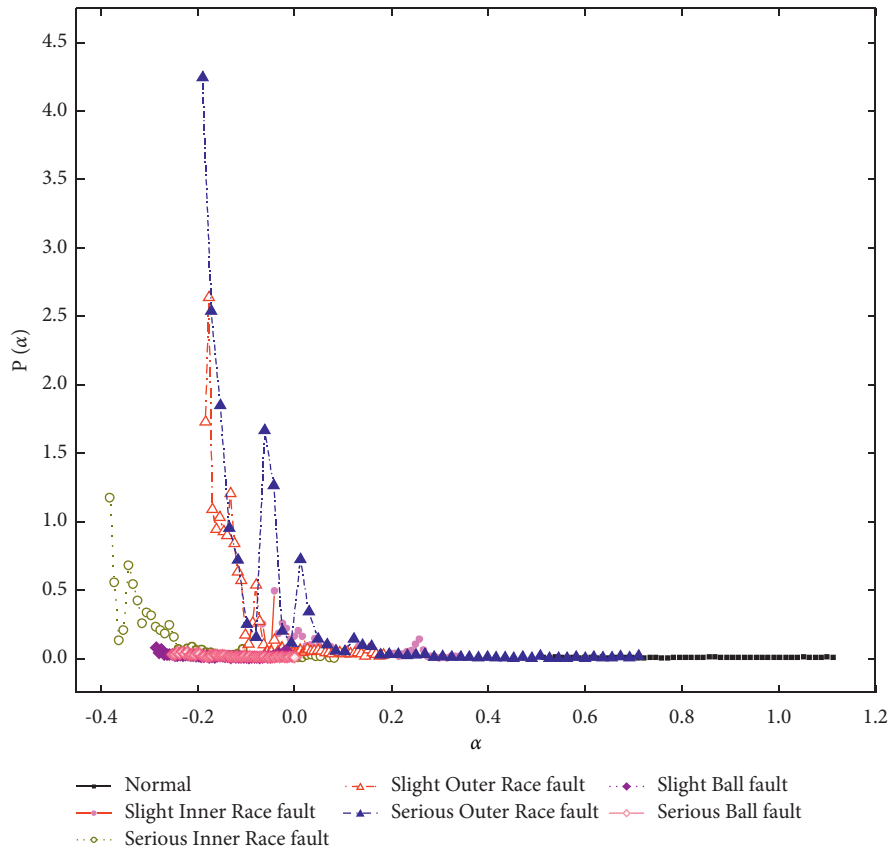


Figure 10: The singularity power spectra of seven signals based on the SPS method.

Table 2: The detailed descriptions of seven vibration signals.

Operational status	α_0	$\Delta\alpha$	$H(q=-5)$	$H(q=0)$	$H(q=5)$	$\sum P(\alpha)$
Normal	0.9428	0.1920	0.9709	1.0180	0.9719	0.164
	0.9015	0.0505	0.8890	0.9016	0.9220	0.1889
	0.8604	0.0810	0.8634	0.8810	0.8793	0.2001
	0.9444	0.1205	0.9494	0.9777	0.9601	0.2048
	0.879	0.2475	0.9031	1.0036	0.9064	0.187
	0.8412	0.0875	0.8521	0.8633	0.8646	0.1941
	0.8953	0.0358	0.8728	0.8839	0.9002	0.2125
	0.8845	0.1472	0.9252	0.9704	0.9296	0.1795
	0.8943	0.2320	0.9318	1.0088	0.9348	0.2152
	0.8801	0.2773	0.9592	1.0447	0.9659	0.1808
Slight inner-race fault	0.0212	0.7166	0.2469	0.4784	0.2680	3.8163
	0.017	0.7609	0.2410	0.4953	0.2626	3.4612
	0.0282	0.7446	0.2399	0.4968	0.2612	3.6263
	0.0194	0.7637	0.2446	0.4934	0.2655	3.9188
	0.0199	0.7514	0.2423	0.4926	0.2635	3.6843
	0.0338	0.7990	0.2349	0.5216	0.2558	3.9312
	0.0232	0.7452	0.2408	0.5013	0.2629	3.8348
	0.026	0.8390	0.2442	0.5325	0.2647	3.8426
	0.0206	0.7632	0.2394	0.4959	0.2606	3.7296
	0.0165	0.6942	0.2382	0.4722	0.2601	4.4186

Table 2: Continued.

Operational status	α_0	$\Delta\alpha$	$H(q=-5)$	$H(q=0)$	$H(q=5)$	$\sum P(\alpha)$
Serious inner-race fault	0.0179	1.0454	0.2957	0.6971	0.3247	6.5567
	0.0138	1.0007	0.3016	0.6499	0.3295	9.7837
	−0.0038	0.9971	0.2993	0.6581	0.3295	8.4325
	−0.0253	1.0316	0.2864	0.6547	0.3164	9.2739
	0.0035	1.0547	0.2926	0.6672	0.3211	6.15
	0.0027	0.9944	0.3019	0.6717	0.3329	7.4287
	−0.0039	1.1739	0.3056	0.7462	0.3361	5.5631
	0.0178	1.0191	0.2899	0.6604	0.3158	8.1085
	−0.0172	1.1845	0.3005	0.7466	0.3317	7.1877
	0.0032	1.1686	0.2995	0.7543	0.3297	7.1261
Slight outer-race fault	−0.113	0.8248	0.2408	0.4875	0.2780	15.47
	−0.1201	0.8391	0.2584	0.5113	0.3007	14.3396
	−0.1166	0.8249	0.2392	0.4847	0.2766	14.1815
	−0.1082	0.8311	0.2301	0.4799	0.2639	13.2106
	−0.1168	0.8655	0.2504	0.5129	0.2900	12.5522
	−0.1169	0.8559	0.2426	0.5009	0.2797	15.8954
	−0.1179	0.9478	0.2421	0.5299	0.2788	14.9106
	−0.1099	0.8490	0.2501	0.5178	0.2895	17.2239
	−0.1111	0.8175	0.2446	0.4951	0.2834	21.1568
	−0.1081	1.0547	0.2337	0.5815	0.2699	18.6271
Serious outer-race fault	−0.0035	0.9119	0.3916	0.6694	0.4321	15.9467
	−0.0135	1.1838	0.3972	0.8284	0.4442	13.4769
	−0.0468	1.0179	0.3876	0.6946	0.4304	19.988
	−0.0231	1.0009	0.3887	0.7131	0.4342	17.8346
	−0.03	1.0657	0.3965	0.7533	0.4433	17.228
	0.0041	0.9967	0.3921	0.7576	0.4387	11.4255
	−0.0169	0.9972	0.3727	0.7098	0.4144	17.3455
	−0.0311	1.0928	0.3804	0.7604	0.4261	22.7398
	−0.0226	1.0415	0.3886	0.7333	0.4342	23.2125
	−0.0111	0.9941	0.3875	0.7321	0.4340	22.9646
Slight ball fault	0.2136	0.5286	0.3091	0.4942	0.3175	0.9391
	0.2006	0.4239	0.2935	0.4254	0.3008	0.9505
	0.2255	0.3932	0.3093	0.4395	0.3166	0.8477
	0.2318	0.2892	0.2961	0.3894	0.3016	0.9875
	0.2353	0.4336	0.3075	0.4580	0.3144	0.9626
	0.245	0.2882	0.3101	0.4053	0.3158	0.9358
	0.2234	0.4132	0.3150	0.4592	0.3237	0.8414
	0.2308	0.4095	0.2942	0.4448	0.3009	0.934
	0.2447	0.4384	0.3159	0.4689	0.3227	0.8998
	0.2311	0.5342	0.3153	0.5191	0.3243	0.903
Serious ball fault	0.3011	0.5576	0.4128	0.6061	0.4229	0.6763
	0.2925	0.4627	0.3897	0.5352	0.3975	0.3966
	0.2533	0.6487	0.3998	0.5921	0.4105	0.5959
	0.2751	0.4947	0.3835	0.5475	0.3930	0.5489
	0.2931	0.4189	0.3927	0.5301	0.4013	0.6171
	0.2697	0.4043	0.3623	0.4861	0.3698	0.5054
	0.2823	0.4570	0.3861	0.5381	0.3952	0.5189
	0.2843	0.4149	0.3782	0.5047	0.3856	0.6314
	0.2821	0.3931	0.3873	0.4916	0.3944	0.8207
	0.2831	0.4511	0.3873	0.5228	0.3953	0.657

$\sum P(\alpha)$ calculated by using the SPS method are shown in Figures 10–15. The horizontal coordination denotes the number index of the sample signal. The vertical coordination denotes the feature parameter chosen for classification. The values of the six specific characteristic parameters are shown in Table 2.

In Figure 11, the normal state signal has the largest value of α_0 and the α_0 value of the signal for the occurrence of faults is smaller than the α_0 value of the normal state. Similarly, the α_0 values of the same faults at different degrees present a larger α_0 value for the signal with serious degree faults than for the signal with slight

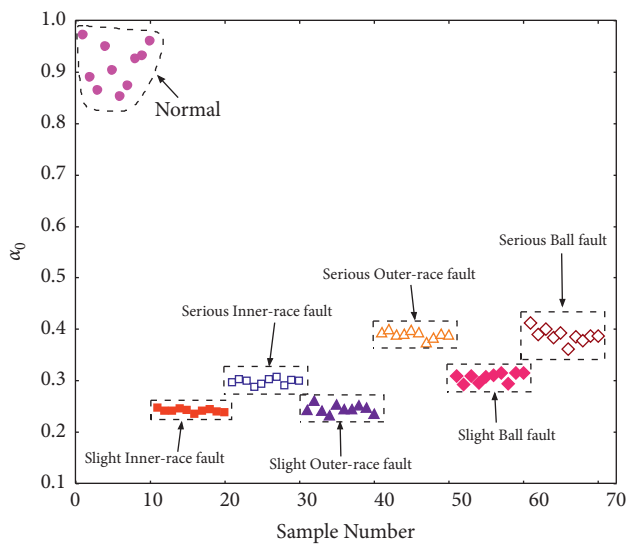


Figure 11: The distribution range of feature α_0 calculated by MFDFA-1.

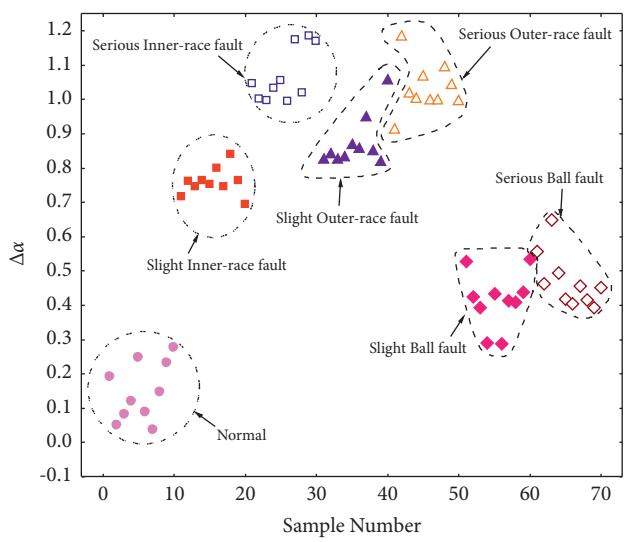


Figure 12: The distribution range of feature $\Delta\alpha$ calculated by MFDFA-1.

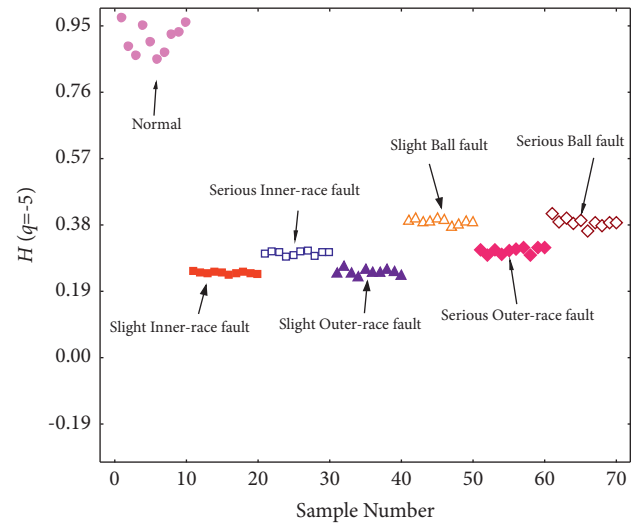


Figure 13: The distribution range of feature $H(q = -5)$ calculated by MFDFA-1.

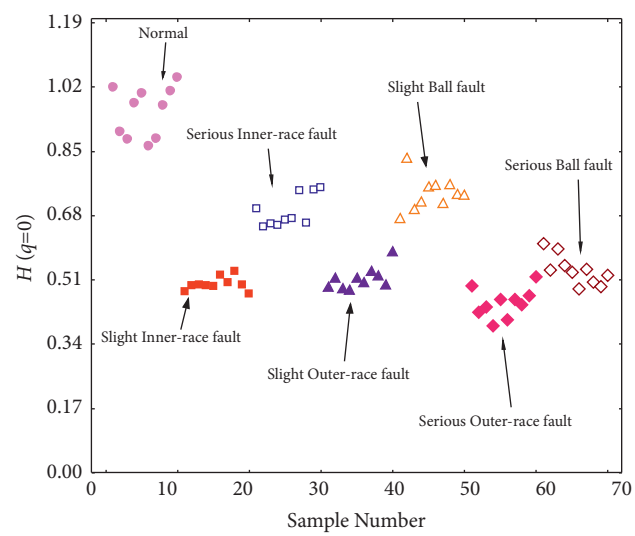


Figure 14: The distribution range of feature $H(q = 0)$ calculated by MFDFA-1.

degree faults. This indicates that the two parameters $\Delta\alpha$ and α_0 can be used as characteristic parameters to distinguish the normal state from fault states and also provides some possibilities to distinguish the degree of the fault as well.

In Figure 12, the distribution of $\Delta\alpha$ values for normal signals is at a low level, accompanied by an increase in $\Delta\alpha$ values in the case of faults, while the $\Delta\alpha$ values for faults of the same type but of different degrees are larger for serious faults than for slight faults.

In Figure 13, when $q = -5$, the values of the q-order Hurst exponent $H(q = -5)$ for different degrees of the same fault start to be distinguishable, especially for serious and slight faults in the inner-race as well as for serious and slight faults in the outer-race. Therefore, the q-order Hurst exponent value at $q = -5$ is sensitive to the degree of inner- and outer-race faults and is suitable for extraction as a feature parameter.

In Figure 14, when $q = 0$, $H(q = 0)$ corresponding to serious rolling element faults can be distinguished as a whole from those corresponding to slight rolling element faults. It also shows good differentiation between serious outer-race faults and slight outer-race faults.

In Figure 15, when $q = 5$, $H(q = 5)$ for the signal of rolling element fault is relatively larger than those for the inner-race fault and outer-race fault values, but it is less sensitive to the degree of damage of the fault.

The singularity power spectra of the seven signals also exhibit distinguishability between normal and fault states. In Figure 10, the distribution of singularity exponents α for the normal state lies in the rightmost of the figure. The

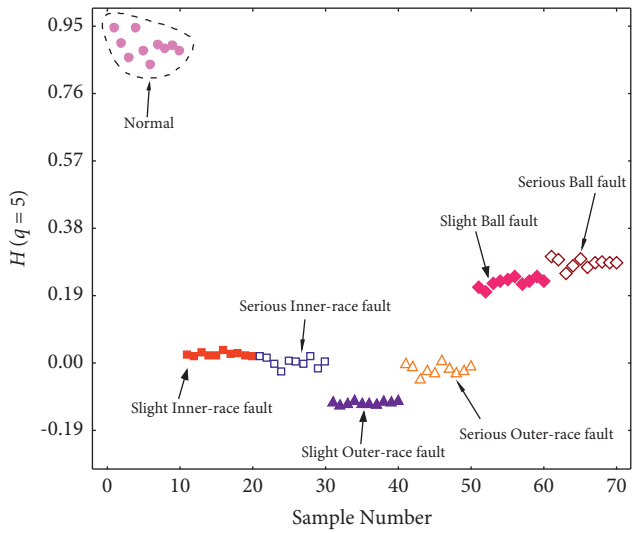


Figure 15: The distribution range of feature $H(q = 5)$ calculated by MFDFA-1.

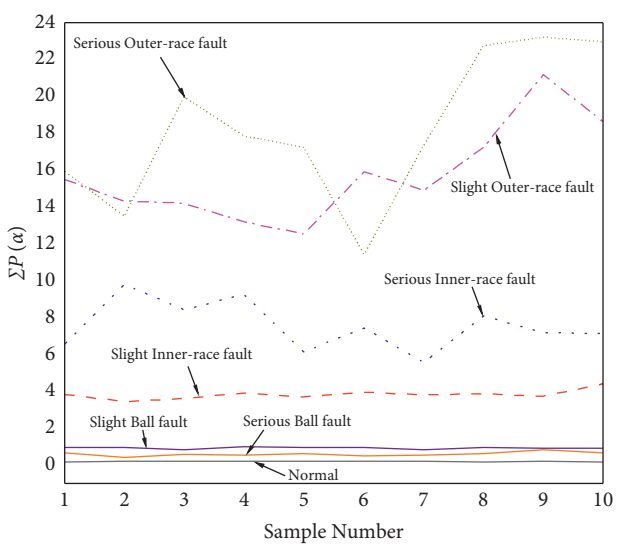


Figure 16: The distribution range of feature $\sum P(\alpha)$ calculated by the SPS method.

distribution of the singularity power of the signal in the normal state is flat compared to the singularity power of the signal in the faulty state. Among them, in the distribution of α, the serious inner-race fault signal lies in the leftmost of the figure, and the serious outer-race fault signal has the largest power extrema in its distribution range of α. The values of the singularity power of slight inner-race fault signal are overall smaller than those of severe degree inner-race fault signal. However, the singularity power spectra cannot be separated from each other evidently.

In Figure 16, the distribution of $\sum P(\alpha)$ for the seven states of the bearing signal is depicted. The distribution of $\sum P(\alpha)$ for serious outer-race fault signal is the widest; comparing to slight outer-race fault signal, there are crossovers existed in some samples, but the majority of

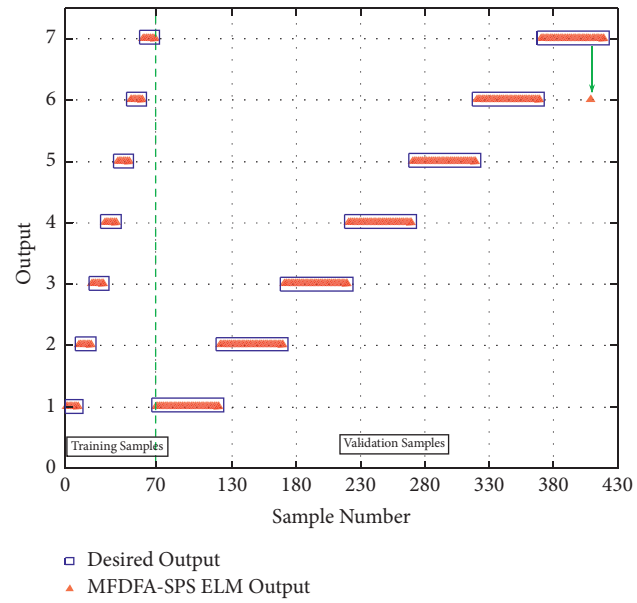


Figure 17: Classification results of the ELM model based on MFDFA-SPS.

samples have greater $\sum P(\alpha)$ values than slight degrees of outer-race faults. $\sum P(\alpha)$ distinguishes the degree of two inner-race fault signals, and the distributions of both feature lines are distinguished evidently in Figure 16. Since the singularity power spectrum distribution of ball fault signal is stable, there is only a minor difference between the $\sum P(\alpha)$ feature lines of serious ball fault signal and slight ball fault signal, but $\sum P(\alpha)$ is more stable for the samples.

Some valuable patterns can be drawn from these plots; the magnitude of the spectrum vertex position α_0 usually increases as the degree of damage increases for the same fault type. The multifractal spectrum width $\Delta\alpha$ is more likely to increase for severely damaged faults than for slightly damaged faults. Except for rolling element faults, the SPS summation $\sum P(\alpha)$ feature values for severely damaged inner and outer-race faults are usually larger than those for slightly damaged inner- and outer-race faults. The above analysis proves that the six characteristics selected in this paper can effectively quantify the types of bearing faults and the degrees of damage.

According to the representation law of these six features, the six parameters are composed into a vector to represent the different bearing state characteristics.

For classification of types using ELM, the number of hidden neurons is 50 and the transfer function is sigmoidal function. The training set consisting of 70 samples is fed into the ELM classifier, and the remaining samples after segmenting the data are used as the test set. The ELM classification results are shown in Figure 17. Similarly, the training set is trained using SVM and LSSVM tools, respectively, and the same is verified for the test set. The classification results obtained are shown in Figures 18 and 19. In the multi-classification problem of bearing faults using SVM, Support Vector Machine for Regression (SVR) is chosen. The kernel function is chosen as radial basis function with a gamma

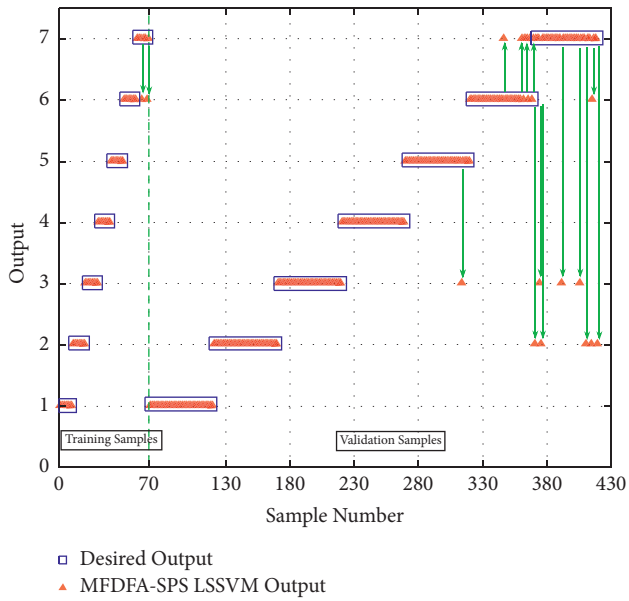


Figure 18: Classification results of the LSSVM model based on MFDFA-SPS.

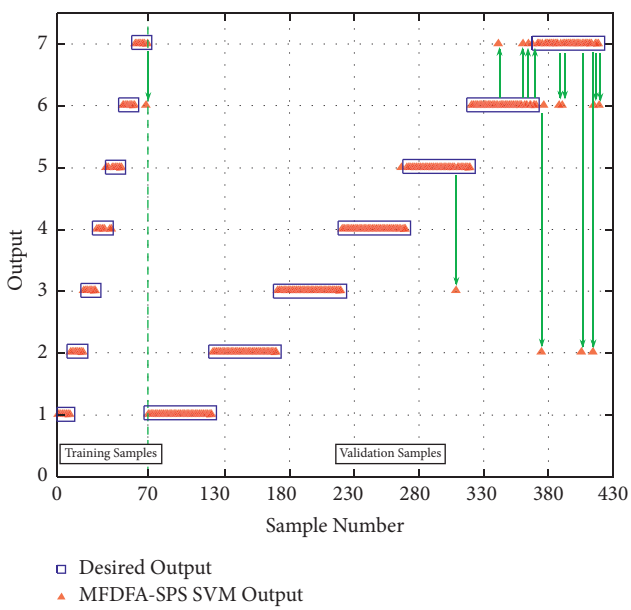


Figure 19: Classification results of the SVM model based on MFDFA-SPS.

parameter coefficient of 0.05. For classification of fault signals using LSSVM, the optimization function is chosen as simplex, and the training and test sets are classified using a cross-validation method.

As can be seen in Figures 20 and 21, we have 7 kinds of signals in different states; the validation acceleration data of each state are divided into 50 sets, so we can get 350 sets. All training sets consisted of 10 sets of each state, so we get 70 sets for training, and the number of total test sets is 350. The average runtime and correct recognition rates of five methods are shown in Table 3. The diagnostic results of the

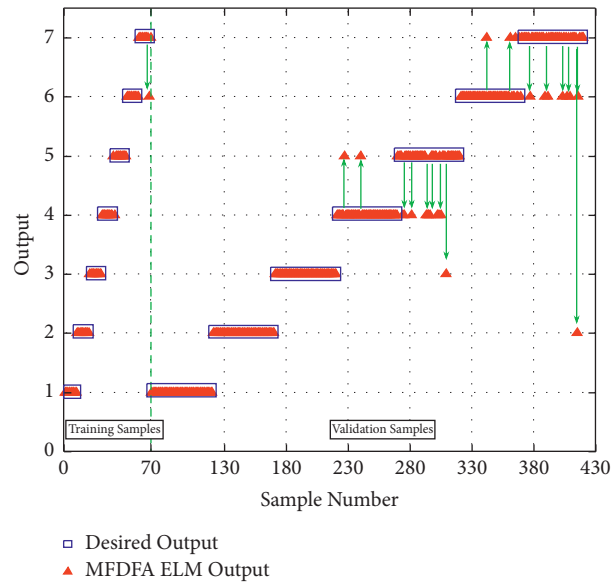


Figure 20: Classification results of the ELM model based on MFDFA.

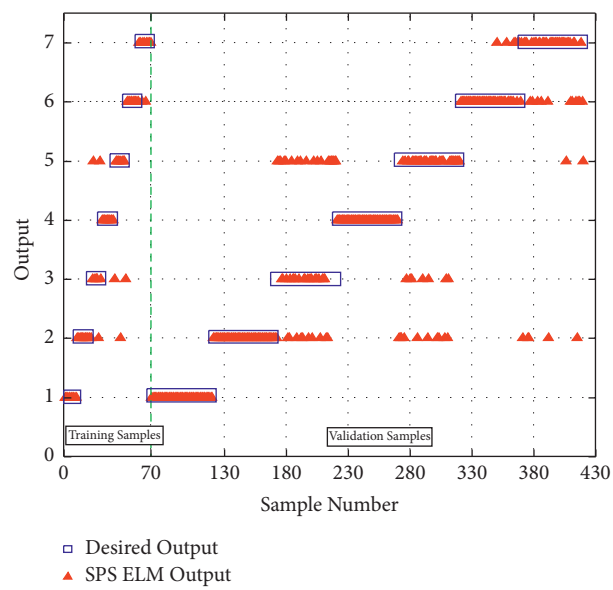


Figure 21: Classification results of the ELM model based on SPS.

Table 3: Average runtime and recognition accuracy of five methods.

Algorithm	Average runtime (s)	Average recognition accuracy (%)
MFDFA + ELM	54.41	92.46
SPS + ELM	2.05	80.83
MFDFA-SPS + ELM	54.56	99.25
MFDFA-SPS + LSSVM	55.52	95
MFDFA-SPS + SVM	54.22	95.71

five classification tools illustrate that this feature extraction method can distinguish different kinds of bearing fault signals with different faults. Among the ELM classifier

results, there are no errors in the test set except for only one data belonging to the serious rolling element fault in the test set is classified into the slight rolling element fault with a correct rate of 99.25%. In the results obtained by the MFDFA-SPS-based detection method with SVM classification, the highest probability of errors was found for ball faults, with an overall correct rate of 95%. In the LSSVM classification results, the correct rate also reached 95.71% due to the excessive number of parameters that needed to be preset and did not have optimization.

Therefore, based on the overall classification correctness, it can be obtained that the MFDFA-SPS-based feature extraction method can be applied to the fault diagnosis of bearings, and the classification method can be further optimized to achieve better classification results.

6. Conclusions

Based on the above study, the following conclusions can be drawn:

(1) By applying the bearing faults from the Case Western Reserve database, the bearing fault diagnosis method combining MFDFA with SPS and ELM can be applied to diagnose different faults as well as faults with different damage levels

(2) The MFDFA-SPS fault feature extraction method based on ELM does not require preprocessing of fault data and is suitable for real-time monitoring of the operation status of bearing faults

If only SPS is used to extract the specific feature parameters of the faulty bearing because the distribution range of the singularity values of different degrees of the faulty signal is different (except for some serious faulty signals), it will cause the stacking of different signal spectrum lines in different intervals of distribution, and the intuitive spectrum line discrimination is not strong only from the graph, so we need to combine other feature extraction methods and choose a more effective classifier for classification to achieve a better classification effect on the rolling bearing signal.

The proposed method has 3 advantages. First, the proposed method has no preprocessing procedure for the sampled data, which substantially reduces the complexity of the overall diagnostic process. Second, the feature extraction time using the MFDFA-SPS algorithm is short, which improves the efficiency of online fault diagnosis. The various bearing signal data used in this paper have only 2048 points, and the need for fault diagnosis can be achieved in a short time. Finally, the effect of classification can reach more than 99%, which ensures the reliability of the diagnosis method. Therefore, the proposed method is suitable for application to practical engineering problems.

With the approach proposed in the above section, we can still achieve good diagnostic results based on the Case Western Reserve bearing database using a smaller number of training set. The future direction of research is to test the sensitivity of this method to the feature extraction and fault diagnosis effect of different bearing vibration signals, as well as to realize the composite fault diagnosis of different faults superimposed.

Data Availability

The acceleration data of the bearings used to support the findings of this study have been deposited on the Case Western Reserve University Bearing Data Center website (https://engineering.case.edu/bearingdatacenter/).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was sponsored by the National Natural Science Foundation of China, under Grant 61901195, and the School Research Start-Up Foundation for on-the-job doctoral training, under Grant 1032931804.

References

[1] J. Gegner, *Tribological Aspects of Rolling Bearing Failures*, Tribology-Lubricants and Lubrication, Taiwan, 2011.

[2] R. B. Randall and J. Antoni, "Rolling element bearing diagnostics-A tutorial," *Mechanical Systems and Signal Processing*, vol. 25, no. 2, pp. 485–520, 2011.

[3] B. B. Mandelbrot, "How long is the coast of britain? Statistical self-similarity and fractional dimension," *Science*, vol. 156, no. 3775, pp. 636–638, 1967.

[4] B. B. Mandelbrot, *Multifractals and 1/f Noise: Wild Self-Affinity in Physics (1963–1976)*, Springer, New York, NY, USA, 2013.

[5] B. B. Mandelbrot and J. W. V. Van Ness, "Fractional brownian motions, fractional noises and applications," *SIAM Review*, vol. 10, no. 4, pp. 422–437, 1968.

[6] T. Harris and M. Kotzalas, *Advanced Concepts of Bearing Technology*, 2006.

[7] J. Antonia and R. B. Randall, "The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines," *Mechanical Systems and Signal Processing*, vol. 20, pp. 308–311, 2004.

[8] Q. Ni, J. C. Ji, and K. Feng, "Data-driven prognostic scheme for bearings based on a novel health indicator and gated recurrent unit network," *IEEE Transactions on Industrial Informatics*, p. 1, 2022.

[9] K. Feng and W. A. Smith, "Use of cyclostationary properties of vibration signals to identify gear wear mechanisms and track wear evolution," *Mechanical Systems and Signal Processing*, vol. 150, pp. 1–24, Article ID 107258, 2020.

[10] J. Shao, X. Pang, Z. Yang, and J. Li, "Research on Intelligent Recognition of axis Orbit Based on Hu Moment Invariants and Fractal Box Dimension," in *Proceedings of the 2017 14th International Conference on Ubiquitous Robots and Ambient IntelligenceURAI*, June 2017.

[11] H. Liu, X. Wang, and C. Lu, "Rolling bearing fault diagnosis based on LCD-TEO and multifractal detrended fluctuation analysis," *Mechanical Systems and Signal Processing*, vol. 60-61, pp. 273–288, 2015.

[12] J. Lin and Q. Chen, "Fault diagnosis of rolling bearings based on multifractal detrended fluctuation analysis and

Mahalanobis distance criterion," *Mechanical Systems and Signal Processing*, vol. 38, no. 2, pp. 515–533, 2013.

[13] Q. Chen, L. Hongru, and X. Baohua, "Fault prediction for bearings based on morphological fractal dimension and improved ELM," *Bearing*, vol. 4, pp. 45–49, 2014.

[14] Q. Xiong, W. Zhang, and T. Lu, "A fault diagnosis method for rolling bearings based on feature fusion of multifractal detrended fluctuation analysis and alpha stable distribution," *Shock and Vibration*, vol. 2016, Article ID 1232893, 12 pages, 2016.

[15] Y. Zhao and K. Yu, "Joint domain alignment and class Alignment method for cross-domain fault diagnosis of rotating machinery," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, Article ID 3526212, 2021.

[16] E. Ihlen, "Introduction to multifractal detrended fluctuation analysis in Matlab," *Frontiers in Physiology*, vol. 3, p. 141, 2012.

[17] G. Xiong, W. Yu, and S. Zhang, "Singularity power spectrum distribution," *Physica A: Statistical Mechanics and Its Applications*, vol. 431, pp. 63–73, 2015.

[18] G. Huang, Q. Zhu, and C. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1-3, pp. 489–501, 2006.

[19] I. J. Good and B. B. Mandelbrot, "Fractals: form, chance, and dimension," *Journal of the American Statistical Association*, vol. 32, pp. 73–438, 1978.

[20] H. Ocak and K. A. Loparo, "Estimation of the running speed and bearing defect frequencies of an induction motor from vibration data," *Mechanical Systems and Signal Processing*, vol. 18, no. 3, pp. 515–533, 2004.

[21] Cwru Dataset, "Case western Reserve university bearing data center," 2022, https://csegroups.case.edu/bearingdatacenter/home.

Hindawi
Mathematical Problems in Engineering
Volume 2022, Article ID 5266054, 12 pages
https://doi.org/10.1155/2022/5266054

Research Article

Application of PCA and SVM in Fault Detection and Diagnosis of Bearings with Varying Speed

Mushabi Pule,[1] **Oduetse Matsebe**,[1] **and Ravi Samikannu**[2]

[1]*Department of Mechanical, Energy and Industrial Engineering, Botswana International University of Science and Technology, Palapye, Botswana*
[2]*Department of Electrical, Computer & Telecommunications Engineering, Botswana International University of Science and Technology, Palapye, Botswana*

Correspondence should be addressed to Mushabi Pule; mushabi.pule@studentmail.biust.ac.bw

Received 22 March 2022; Revised 4 April 2022; Accepted 7 April 2022; Published 26 April 2022

Academic Editor: Ke Feng

Vibration analysis is widely used as an efficient condition monitoring (CM) tool for rotating machines in various industries. Fault detection and diagnosis (FDD) models play an important role in the development of any CM system. The use of artificial intelligence (AI) has since gained recognition in the development of fault detection and diagnosis systems. In this paper, a combination of principal component analysis (PCA) which is used for reducing the data dimensionality, and support vector machine (SVM) which is adopted for classification to carry out fault detection and diagnosis of faults in bearings using vibrations. The diagnostic feature design and machine learning toolbox in MATLAB are used to develop features and train the models, respectively. Real data from the Mendeley data depository is used to test and evaluate the models. Model training is carried out using data with varying speeds representing different conditions of bearing making it different from similar approaches involving SVM. The choice of data used proves that SVM can be able to classify faults with consideration of the varying operating speeds. Results have shown that the combination of PCA and SVM is effective in fault diagnosis of bearing faults under varying speeds such that a 97.4% classification accuracy was achieved. The result implies that PCA and SVM can be implemented in various industrial setups where variable speeds can occur both intentionally or nonintentionally. Furthermore, the method was able to differentiate between compounding faults and faults that occur at different times. The confusion matrix further proves the quality and accuracy of the trained model. Future work will focus on the development of models that can carry out the prognosis of faults in bearings as well as to model for other faults other than bearing faults.

1. Introduction

Condition monitoring abbreviated as CM is a form of predictive maintenance. It is a tool for estimating the current health of rotating equipment using techniques such as vibration analysis, motor current analysis, oil or wear debris analysis, and temperature analysis. Out of all these techniques, vibration analysis is well studied and has been accepted in most CM processes [1]. CM involves fault detection and diagnosis of machines in various industries. Before CM, time-based maintenance (TBM) was adopted through various industries. TBM leads to a waste of manpower, time, and money as maintenance is done based on predefined time [2]. The adverse conditions of industries and the need for continuous production put a strain on rotating machines making them susceptible to faults [3]. Bearings are common parts found in most rotating machines such as motors, turbines, engines, and pumps; they are very much susceptible to faults thus making their maintenance significant especially that their failure leads to machine failure causing disturbances in production [4]. Condition of bearings are categorized into the following: healthy bearing, inner race or ring damage fault, outer race or ring damage fault, and ball bearing damage fault. The faulty conditions are due to various factors such as overloading machines, misalignment, and improper mounting. Bearings are generally made of two concentric rings: an outer ring and an inner ring. In between the two rings, there are ball bearings

or roller bearings which reduce rotational friction in rotating equipment. The various parts of a bearing that usually fail are ball bearings, inner races, outer races, and cage faults, which are fully illustrated in Figure 1.

In this work, real vibration data is analyzed to detect and diagnose bearing faults using PCA and SVM techniques. It should be noted that the data used in the paper is of various speeds. Traditionally, fast Fourier transform (FFT) techniques are applied to infer faults from vibration data. FFT struggles to reveal the faults in the signal, hence the introduction of order tracking technique was introduced. These aforementioned techniques require a trained person to further analyze the signals to detect and diagnose faults using the fault characteristic frequencies. This leads to time-wasting and an increased likelihood of human errors. In recent years, intelligent approaches ranging from machine learning approaches to more advanced deep learning approaches have since been researched. However, most of the research surrounding the field does not cater for the changes in speed and load conditions of rotating equipment which have an effect in the equipment diagnosis. Furthermore, the effects of compounding faults or more than one fault occurring at the same time are somehow neglected. The application of PCA and SVM adopted in this work automates the fault detection and diagnosis process hence eliminates the need for trained personnel, avoid time-wasting, and reduce human errors. Even though the approach has been researched before, the method has not been tested for cases in which the data is of various speed conditions such as an increase in speed, a decrease in speed, increasing the speed, decreasing the speed, and lastly decreasing the speed of increasing speed. Note, the changes occur during the measuring or sampling period hence the purpose of this paper is to test the performance of PCA-SVM with K-fold cross validation on fault detection of bearing fault with various operating speed conditions and for a compounding faults.

This remainder of the paper structure is as follows: Section II presents the review of the related literature sources, section III presents the methods and techniques used, sections IV present the results, and lastly, the conclusion is reported in section V.

2. Related Work

FDD based on vibration analysis has been extensively studied and developed, from simple models to more complex mathematical algorithms. FFT is commonly adopted as a preprocessing technique for vibration data. Order tracking has been the most used technique for analyzing non-stationery signals to reveal the order characteristics necessary for fault diagnosis [6, 7]. In recent years, the autoregressive integrated moving average (ARIMA) has gained recognition [8]. An extended version of ARIMA has also been used [9] to solve the complexity in the modelling of multisensor condition monitoring. The use of artificial intelligence (AI) has also gained popularity in developing models and algorithms for fault detection and diagnosis. Various models have been developed in recent years. One notable tool that is globally accepted is the use of machine learning models or algorithms in fault detection and classification. Most utilized ML models include K-nearest neighbor [10], Naïve Bayes, artificial neural network, and support vector machines (SVMs). Kumar et al. used the novel convolutional neural network for bearing defects. The approach is unique as it can be able to cope with insufficient data unlike other convolutional neural networks, while creating the deep learning effect [11]. SVM has been used [8] to solve change detection problems in a gearbox. Multiple measurement vector compressive sampling, a combination of geodesic minimal spanning tree, stochastic proximity embedding, and neighborhood component analysis, and multiclass SVMs are used in [4] for condition monitoring of roller bearings in rotating machines using vibration signals. The approach has been proven to have reached high-bearing health classification accuracy while outclassing existing methods. Chowdhury adopted multiclass SVM's, Naïve Bayes, binary decision tree, discriminant analysis, nearest neighbour, and ensemble classifiers to test time-windowed extracted features in nonintrusive load monitoring [12]. As it has been stated, AI methods are the recent advances in the FDD field. One of the most used and studied AI methods in the field is SVM. Ahmed et al. [13] proposed a fault detection and diagnosis approach based on SVM for fault classification. Furthermore, the approach was incorporated with compressive sampling and Laplacian score to generate compressively sampled data from raw vibration data signals and to rank the sampled signals, respectively. The method was tested on bearing data sets and proved to be able to classify faults surpassing the other AI-based approaches. Fault diagnosis depends greatly on features and greatly depends on feature extraction. The separation of the two processes create inferior fault diagnosis accuracy. Zhang et al. [14] proposed a two-phase approach aiming to synchronously extract features and optimize SVM parameters for improved fault diagnosis. A hybrid filter and wrapper method were adopted for this purpose. The proposed method was trained and tested for bearing fault diagnosis and rotor fault diagnosis, while at it, the method proved to be

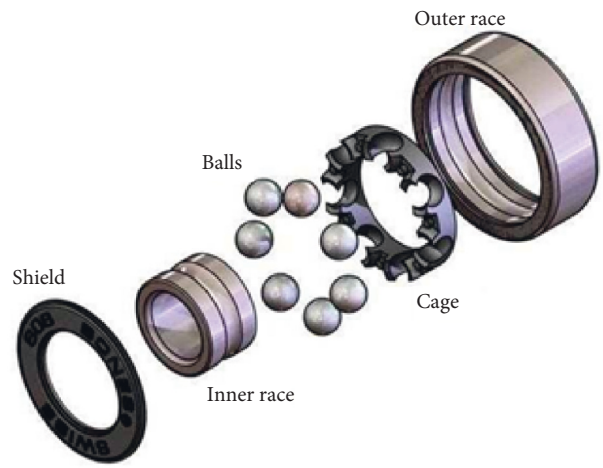


Figure 1: **Exploded diagram of a ball bearing** [5].

able to perform fault classification tasks perfectly. Amir et al. [15] did a study comparing the accuracy of the classical SVM and one class SVM (OC-SVM) for fault detection and diagnosis in bearings. Classical SVM tends to just create a hyperplane across two classes, whereas OC-SVM maximizes the distance between the origin and positive values in the absence of negative values. Furthermore, OC-SVM has the ability to learn given little data. The authors found that OC-SVM classification accuracy is higher than that of classical SVM and this is also affirmed by [16]. More research on the application of SVM in fault detection and diagnosis problems can be found in [17].

SVMs are also adopted for fault classification in [15, 16] and other classification problems in [18]. In [1], principal component analysis (PCA) is used on FFT data to reduce the dimensionality of the data. PCA can distinguish between various motor faults and provide an inexpensive and simple alternative [19]. The application of PCA in fault detection and diagnosis problems is also explored in [20–22]. Another well-researched approach is the extension of artificial neural networks including deep normalized convolutional neural networks [23] and deep neural networks with batch normalization (BN) [24] for the classification of bearing faults. Other neural network approaches are reported in [25–29]. In [26], they proposed a fault detection and diagnosis based on deep convolutional neural network (DCNN) with a mandate to combine the feature extraction and fault classification into a single stage. The approach is different from most methods such that it depends on feature learning rather than manual feature engineering hence eliminating the reliance on an individual's knowledge for feature extraction and selection. With raw data, the approach achieved 98% accuracy for the training sample and 91% for the test sample. In comparison with other approaches based on manual feature engineering, the approach is well on par. The approach seems to reserve most features as raw signals are used and there is less signal processing. Another notable insight about the approach is the ability to reflect on the unknown information relating to the bearings. However, the approach does not seem to cater for changes in speed rotating and situations where 2 more faults exist are not covered in the context. Proposed in [30] is the meta learning method which aims to learn prior from relevant tasks without learning from the start. The method is named meta learning fault diagnosis (MLFD) and the authors achieved an average classification accuracy of 97.28% for complex working conditions on case western data. Deep learning techniques which have a dependency of larger datasets are discussed in the following publications: [30–35].

In this paper, the capabilities of ML are tested in a combination of PCA and SVM with K cross-validation techniques are employed to detect and diagnose bearing faults using vibration analysis methods.

3. Methodology

This section presents the methods, tools, and techniques used in this research work. The whole process from data processing to model training is carried out in MATLAB. Since the fault labels are known already from the data publisher, order tracking was not done. A combination of SVM and PCA is adopted for the fault diagnosis of speed-varying bearing data. PCA is adopted to reduce the data dimensions as well as to reduce computation time, and SVM is used to classify the bearing faults using PCA data. As it is shown in Figure 2, raw time-domain vibration data is prepared using MATLAB. This includes labelling the data according to the type of fault they represent and creating an ensemble of the data to treat the data as a whole. The data is then taken through the preprocessing stage where the order spectrum and power spectrum are created. This enables the extraction of the features which follow next in both the time and frequency domains. The diagnostic feature designer toolbox in MATLAB is employed for the extraction and selection of the features which are then used in the classification learner to train and verify the PCA and SVM models. The PCA and SVM parameters are optimized for the desired results of high accuracy. The evaluation criteria are based on the performance of the algorithm on the classification of the various faults or classes.

3.1. Data Specification. Raw experimental data in the form of vibration signals from a bearing is used for training the model. The data is available through the Mendeley data repository and more information on the data can be found in [36]. The setup for the collection of data is shown in Figure 3. The sampling frequency is 200 kHz for 10 seconds. The bearing type is ER16 K with 9 balls, a ball diameter of 7.94 mm, and a pitch of 38.52 mm.

The data used in the method covers a wide range of different rotating speed conditions hence the effects of changes in speed during the sampling period on the bearing vibrations are taken into consideration. The operating speed conditions include increasing speed, decreasing speed, increasing speed then decreasing speed, and lastly, decreasing speed then increasing speed. The various speeds can be seen in [36]. The data represent various health conditions, including healthy bearing, ball damage fault, combined faults, an inner ring or race fault, and outer race or ring faults. In this paper, the health conditions are represented as follows: 0–healthy, 1–ball fault, 11–combined faults, 111–inner ring fault, and 1111–outer race fault. A total of 40 samples in the dataset were used, each class or fault represented by 8 samples. A combined fault is defined as a case in which ball fault, inner race fault, and outer race fault occur at the same time.

3.2. Preprocessing and Feature Extraction. Firstly, the data is preprocessed to ensure that it is ready for the extraction of condition indicators. The signals were prepared into an ensemble to process them all at once. Preprocessing of data is carried out using both signal-based functions and model-based functions. The time synchronous average (TSA) signal of the data is deduced from the original data to filter out noise and disturbances. Spectral analysis of vibration signals has been proven to be the most widely used for rotating machines. Spectral analysis allows for easy detection of the resonance frequency or the fault frequency [37]. The order

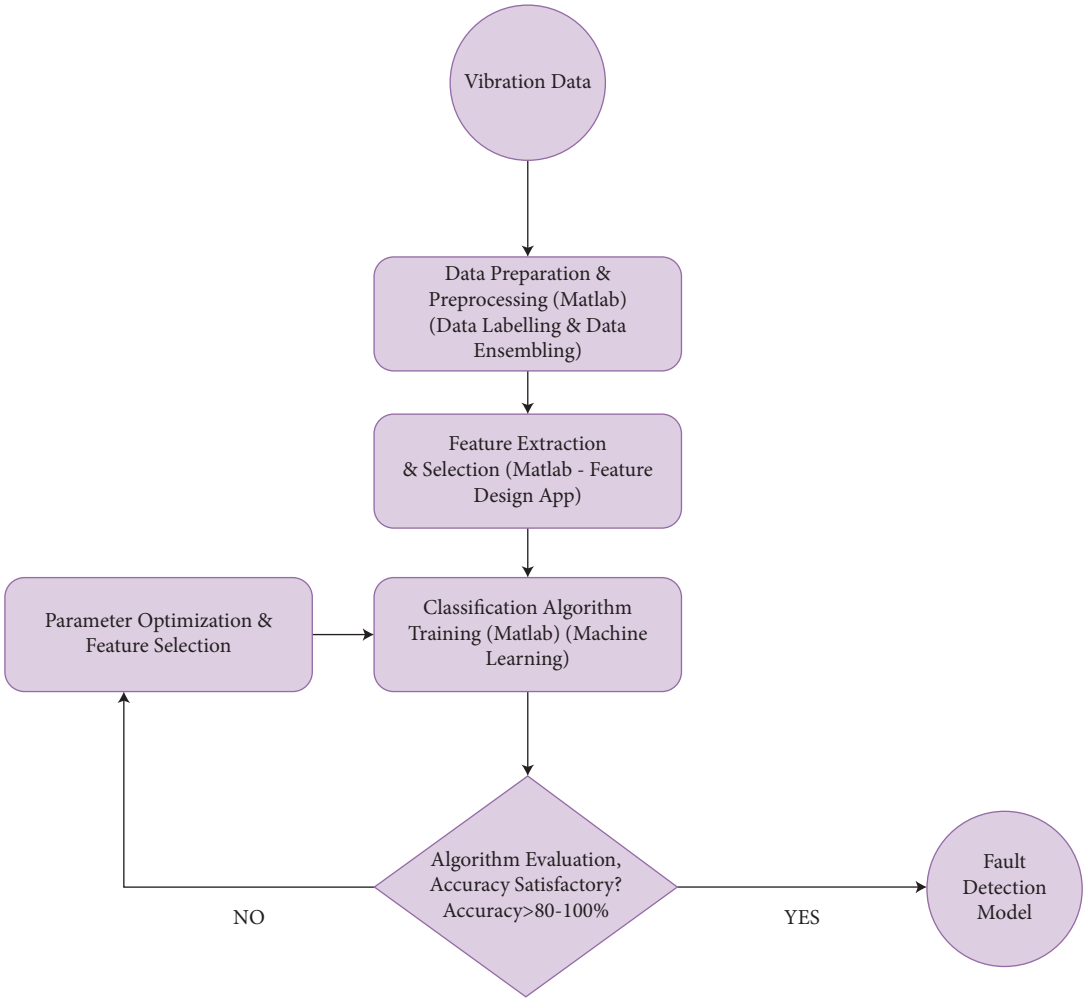


Figure 2: Process flow chart for the model training.

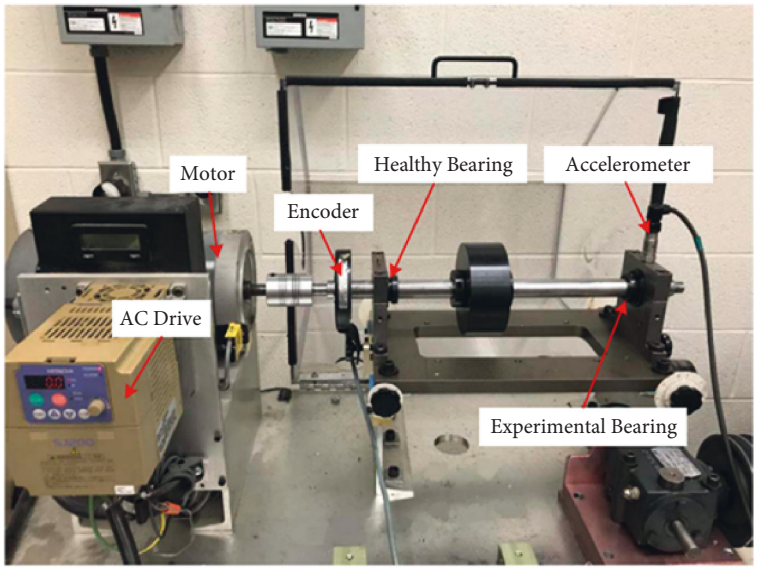


Figure 3: The setup used to collect the vibration data [36].

and power spectrums of the signals are extracted from the resulting TSA signal, as it is free from disturbances. The power spectrum allows for the characterization of frequency content and resonances within a system. Faults cause changes in the spectral signature making it easy for the extraction of features using the power spectrum. The order spectrum, similar to the power spectrum provides an extended understanding of harmonically interrelated systems in rotating machinery. Features are extracted from both the frequency and time series domains. The features are ranked according to their significance; less important features are not selected for classification. The data is then exported to the classification learner for model training. With PCA active, the optimized Gaussian SVM model is trained. Upon completion of training, a confusion matrix is generated which is used to check how the trained model performs for each class.

3.3. Condition Indicator Classification. PCA is used to reduce the dimensions of the training data features. SVM is adopted as a classification model. Features that are less significant in the preprocessed data are sidelined to increase the accuracy of our model.

3.3.1. Principal Component Analysis (PCA). PCA is one of the statistical learning algorithms used to reduce the features extracted from a signal. It has always been adopted for prediction, classification, and feature extraction problems [38, 39]. It changes a lot of related variables into new sets of uncorrelated variables and, in the interim, holds most of the information on the first signal. The principal components (PCs) are acquired from the uncorrelated variables to detect and confine process anomalies in a vigorous way [40]. For simplicity, any given normal data matrix $X_{(N \times m)}$, X is transformed into a new matrix $T_{(N \times r)}$ where r is greater than m. This is achieved by using a transformation matrix $P_{(m \times r)}$ [21].

$$T_{N\times r} = X_{N\times m} P_{m\times r}, \tag{1}$$

where P and T represent the orthogonal loading matrix and score matrix, respectively.

3.3.2. Support Vector Machine. Support vector machines popularly known as SVM is a supervised machine learning model used for both classification and regression problems, defined by a separating hyperplane/line for a given training set [41]. SVM has always been adopted for various classification problems such as Internet traffic classification [42]. Various hyperplanes separating the two or more classes (Figure 4) exist, but the SVM classifier depends primarily on the hyperplane or line that has the maximum separating margin among the fault classes (Figure 5) [44, 45]. The larger the separating or functional margin, the lower the classification error [46]. Originally, SVMs were designed for binary classification; however, they can also be used for nonlinear classification problems with the help of kernels [47], hence making them more versatile for classification problems [46].

Linear SVM is used since it is simple to implement. The linear classifier function equation (2) is expressed as follows:

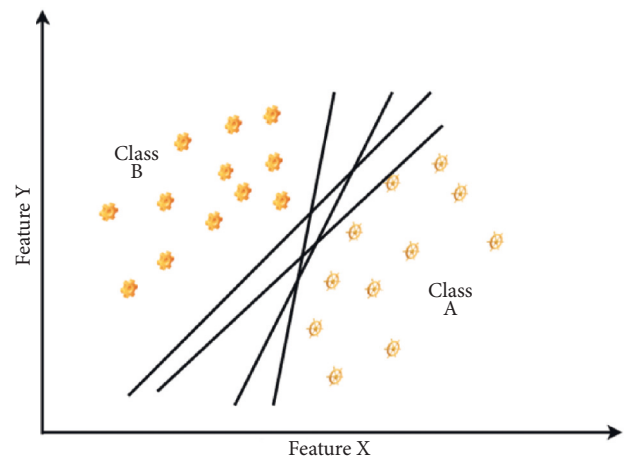


Figure 4: **Illustration of principle hyperplanes in SVM** [43].

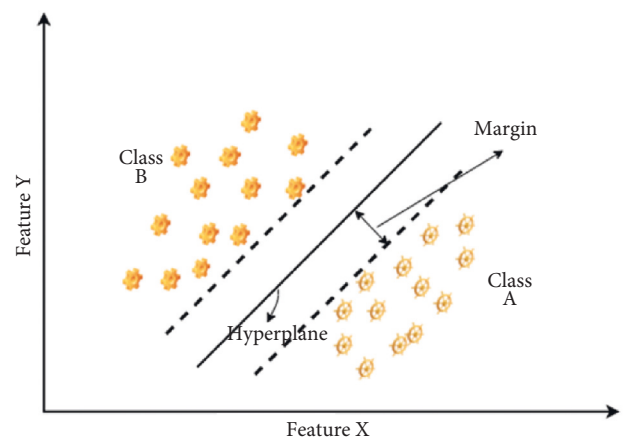


Figure 5: SVM optimal hyperplane between two classes.

$$f(x) = w^T x + b, \tag{2}$$

w and b are unknown coefficients that are determined from the minimization cost function shown in the following:

$$\min J(w, \xi) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N} \xi_i, \tag{3}$$

Subject to

$$\begin{gathered} y_i\left(w^T \Phi(x_i) + b\right) \geq 1 - \xi_i, \\ \xi_i \geq 0; i = 1, 2 \ldots, l. \end{gathered} \tag{4}$$

C is a user-specified, positive, regularization parameter adopted for control of trade-off among the model complexity and empirical risk [48].

To transform (4) for nonlinear classification, a kernel notion shown in (5) is used.

$$K\left(x_i, x_j\right) = \Phi^T(x_i)\Phi\left(x_j\right), \tag{5}$$

where Φ is the nonlinear operator. The following types of kernels are usually adopted: linear kernel, polynomial kernel, RBF kernel, and MLP kernel [49].

Table 1: SVM parameter used when training the algorithm.

Parameter	Detail
Kernel	Gaussian or RBF
Kernel scale	Automatic
Acquisition function	Bayesian
Optimization	Bayesian optimizer
Multiclass method	One vs one

Therefore, SVM is represented as follows:

$$f(x_i) = \sum_{k=1}^{N_i} y_i a_k K(x_i s_k) + b, \tag{6}$$

where $s_k, k = 1, \ldots, N_i$ represents support vectors that correspond to the training data samples, set during the training step, and y_i is the class label. The SVM parameters are stated in Table 1.

3.3.3. K-fold Cross-Validation. Cross-validation involves partitioning datasets into various equal k parts to reduce the overfitting and under fitting of classification models. It allows model training while reserving the kth fold or part for the validation of the training accuracy. After model or algorithm training, the fold is used to test the trained model [50].

4. Results and Analysis

This section presents the results from the preprocessing of the data and the training of the SVM model. A total of 60 vibration datasets were available. 40 datasets were used for the training purpose and 20 datasets were left for testing purposes.

Figure 6 shows the original vibration signals before preprocessing. The health conditions represented by the data are healthy, ball fault, combined faults, an inner ring or race fault, and outer ring or race fault. These are labelled as 0, 1, 11, 111, and 1111. The y-axis represents the vibration (peak to peak amplitude) in m/s^2 and the x-axis indicates the timestamps (samples) of the vibrations. The frequency-domain plots are transformed from the time domain.

Figures 7 and 8, respectively, depict the power and order spectrums of the vibration signals. Spectrum analysis is adopted due to its ability to show the underlying feature which cannot be seen in a simple time-domain signal and allows for easy detection of the resonance frequency or the fault frequency. The order and power spectrums of the signals are extracted from the raw vibration signal. The power spectrum allows for the characterization of frequency content and resonances within a system. In the power spectrum, the y-axis indicates the power of the signals in decibels, and the x-axis represents the frequency of the signals on a logarithmic scale. The logarithmic scale allows coverage of a wider range of frequencies. In the order spectrum, the y-axis indicates the power of the signals in decibels and the x-axis represents the orders, i.e., the frequency expressed in multiples of the running speed. The order spectrum is necessary to reveal the order characteristics of the signal [6].

Features are extracted from both the time-domain series and the frequency-domain series. The features are ranked according to their significance using analysis of variance (ANOVA). The feature ranking can be seen in Figure 9. Less important features were not selected for use in classification model training through a manual feature engineering process as seen in Table 2. A scatter plot was used to investigate the features by plotting the features against each other to see the combinations which can give the best classification results by showing variations in the classes. Figure 10 shows that a combination of the peak frequency of the order spectrum and skewness in time-domain features can be great for the classification of bearing faults. The scatter plot suggests that overfitting may be present between class 0 and class 1111 as they appear to be closer to each other. This may be a factor leading to the misclassification error of the algorithm. Figure 11 shows the legend for the scatter plot.

With the K-cross validation set to 10 to reserve a set of 10 data sets for the validation test, the PCA was set to keep 3 numeric components after training and a Gaussian kernel was selected for the optimizable SVM (one vs all). PCA, as it has been stated, it is used to reduce the dimensions of the feature dataset by giving a direction to keep features with maximum variance. The model achieved a classification accuracy of 97.4% with a training time of 88.315 seconds. The test accuracy was found to be 90% for the remaining 20 data samples. Figure 12 displays the confusion matrix which is used to check how the trained model performs for each class. The model can achieve at least 89% of positive predicted values (annotated in green) and achieve at most 11% of false discovery rates (annotated in pink) for some of the classes. A false discovery rate occurred for class inner race fault class where some of its data was predicted to be a combined fault class. This may be because class 11 represents inner race faults and some of the value of the features may be greatly present in the case of combined faults. The analysis of the results implies that PCA-SVM can be used to classify bearing faults when there are changes in speeds while sampling, to avoid representing changes in speed as faults.

The comparisons based on accuracy are documented in Table 3 [35, 51–54]. From Table 3, it can be taken that the method performed fairly compared to other approaches. However, SVM host an advantage of low computational time and the ability to cope with limited data.

The SVM model was tested on test vibration data resulting in the confusion matrix in Figure 13. The test accuracy was 97.8% and the test results have shown that

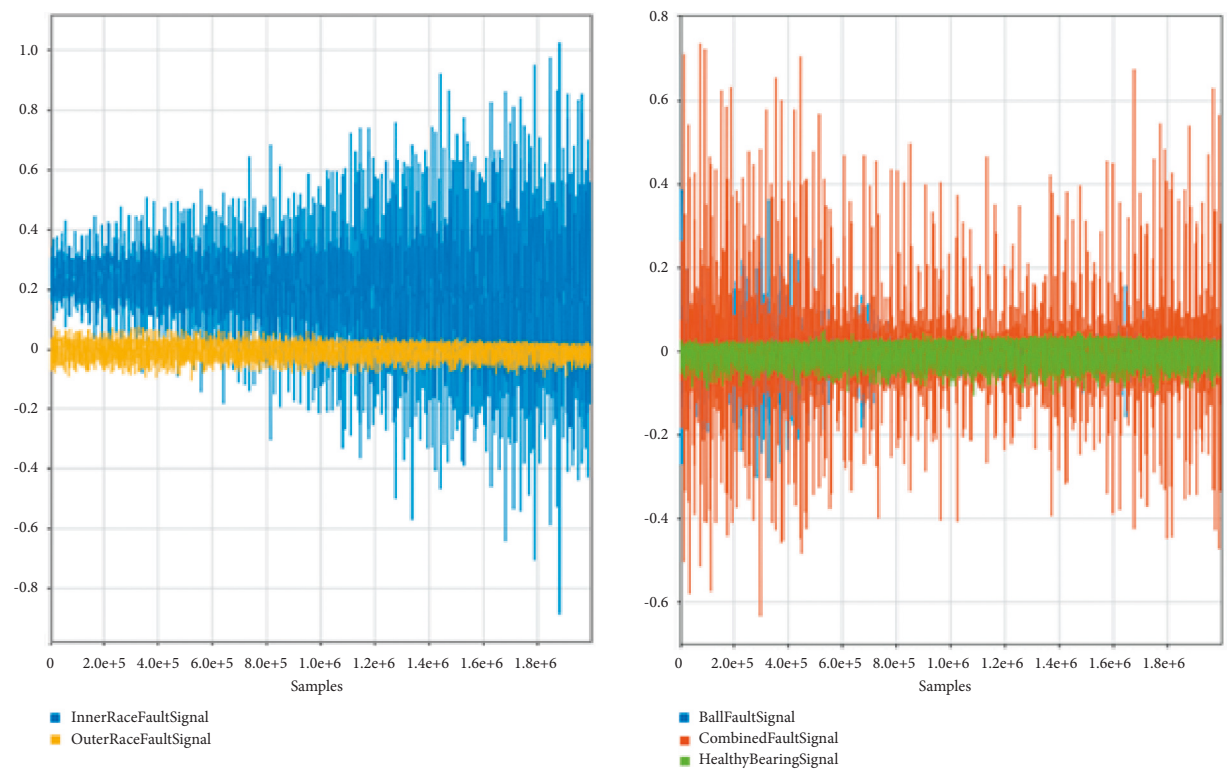


Figure 6: Ensemble view of raw vibration signals.

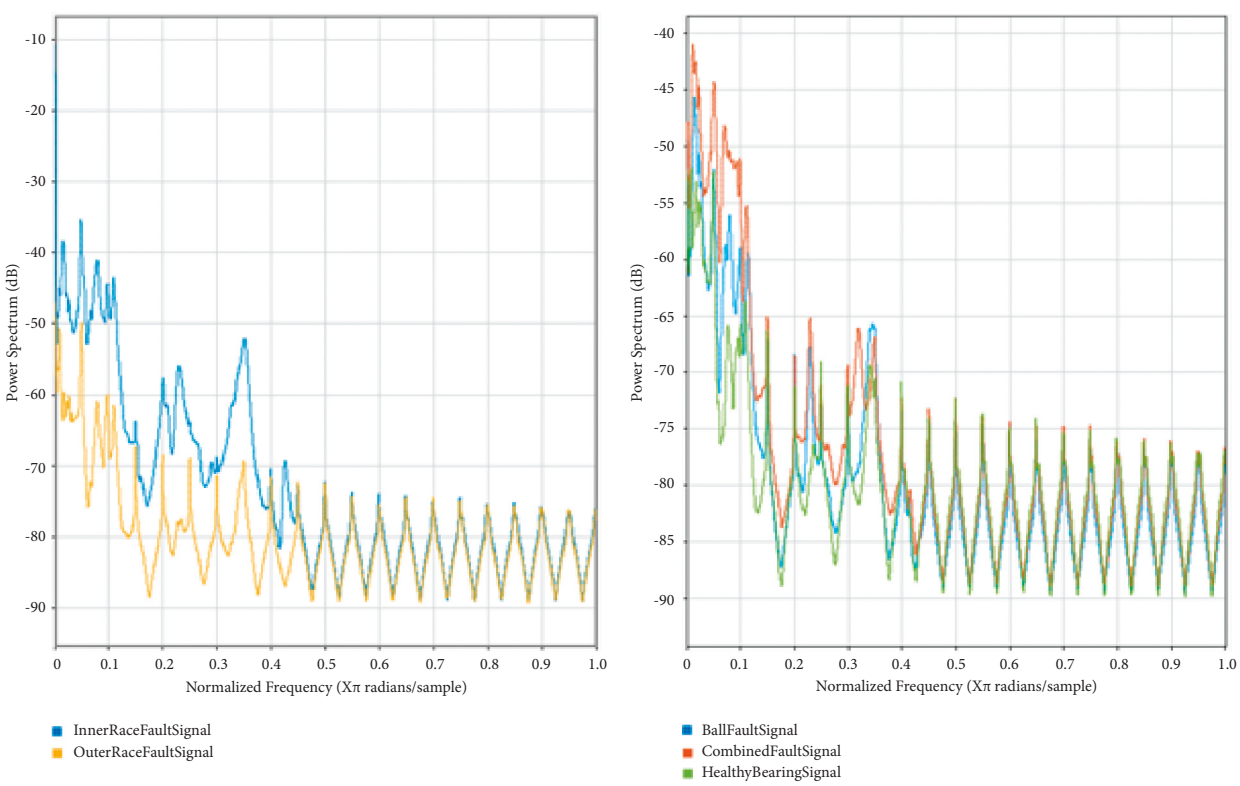


Figure 7: Power spectrum of vibration signals dataset samples.

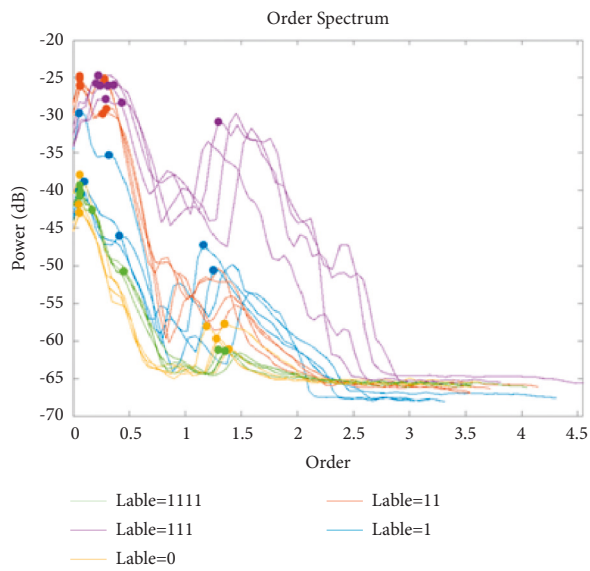


Figure 8: Ensemble view of order spectrum of various vibration signals.

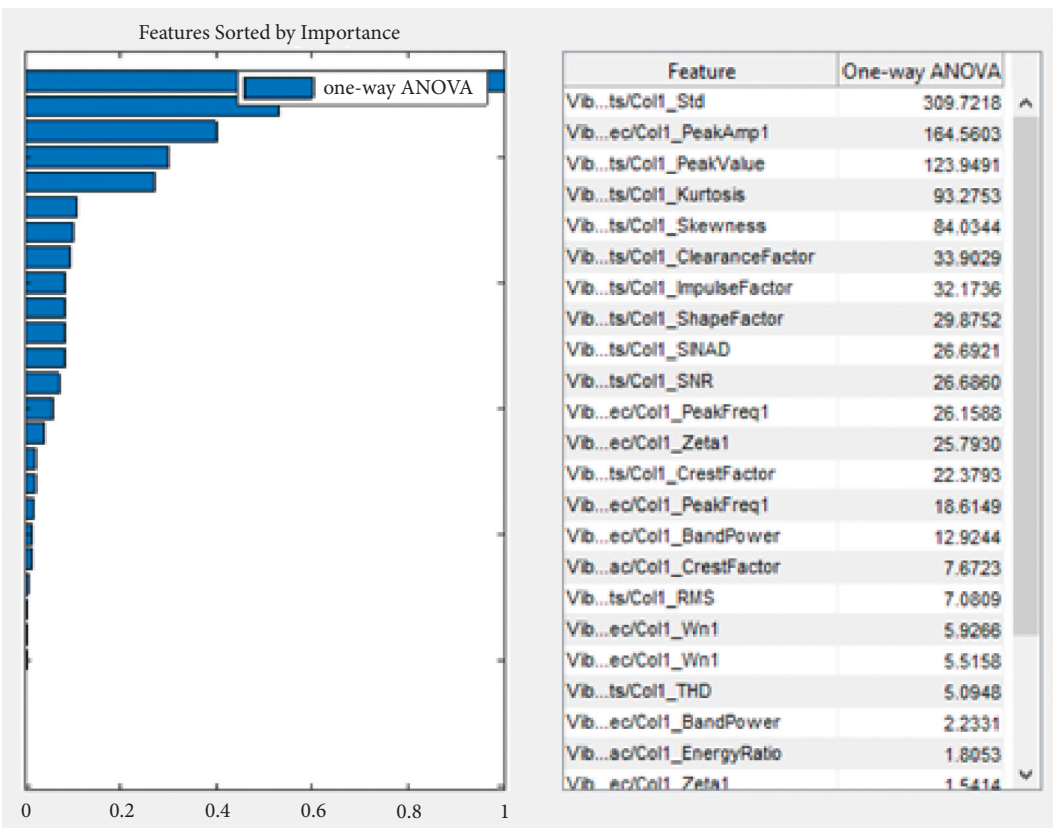


Figure 9: Extracted feature ranking using ANOVA.

Table 2: Selected features for model training.

Feature domain	Features selected
Time domain	Standard deviation, skewness
Frequency domain	Order spectrum's peak frequency and peak amplitude

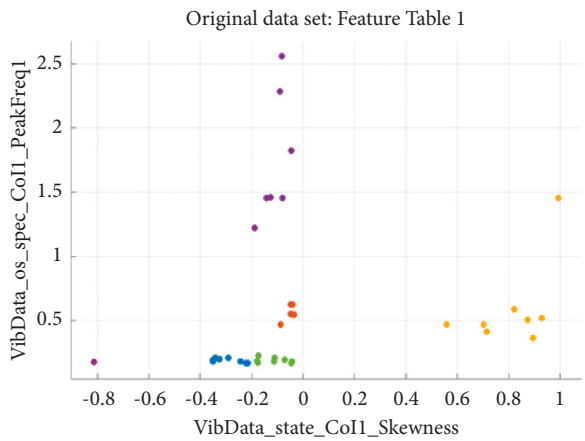


Figure 10: Scatter plot of peak frequency in order spectrum feature and time-domain skewness.

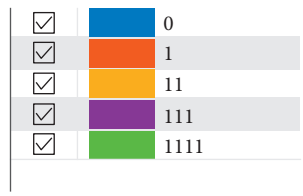


Figure 11: Legend for the scatter plot.

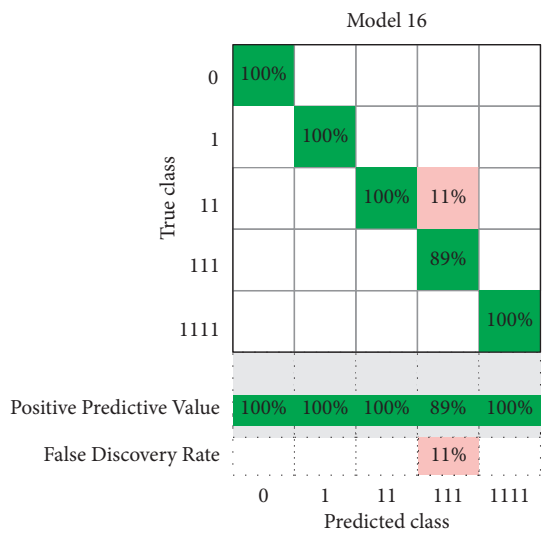


Figure 12: The confusion matrix of trained classification model.

Table 3: Comparison of PCA-SVM method with K-fold cross-validation with recent related AI fault classification models.

Author	Model	Classification accuracy (%)	Speed
Shuhui Wang, 2018	CNN based Markov model	98.125	Constant
Tang et. al, 2019	Adaptive learning rate deep belief network (ADDBN)	99.75	Constant
Sufi Tabassum Gul, 2018	SVM with 5 K-cross fold validation	91.95	Constant
Liu et al. 2018	1D-CNN	99.999	Constant
Our method	SVM + PCA and 10 K-fold CV	97.4	Various speed condition

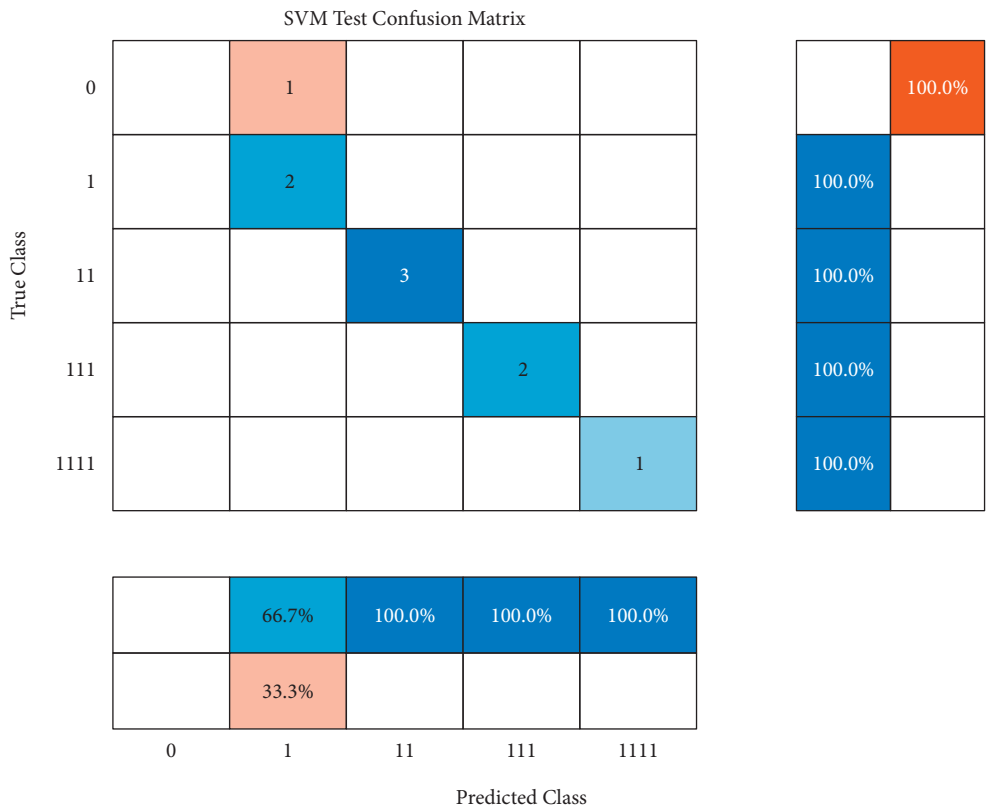


Figure 13: Confusion matrix for test data.

SVM can indeed be greatly considered for fault detection and diagnosis of rotating equipment using vibration data.

5. Conclusion

Vibration analysis has proven to be the most consistent technique when it comes to condition monitoring of rotating equipment. In this paper, a combination of PCA and SVM with cross-validation is used for fault classification of bearings using real data from the Mendeley data depository. The trained model achieves an accuracy of 97.4% with a training time of 88.315 seconds. The model shows an accuracy of at least 89% for all the health condition classes. These results imply that PCA and SVM can be combined to detect and classify faults; however, more work needs to be put into further optimization of the approach to improve the classification accuracy and reduce the dependence of individual knowledge on the feature selection process. The implication of the result based on the training and test accuracies suggests that PCA and SVM can be successfully employed for real engineering practice. The main advantage of the method is tied to the fact that SVM can achieve high accuracy in cases where there is limited data as depicted in most industrial setups. Industrial setups have limited data since the industrial sites are not collecting and archiving historical data. Hence, there is not enough data to train AI-based algorithms. Additionally, SVM and PCA do not provide computational complexity, hence they are suitable for application in real engineering scenarios. Future work includes testing the approach on other problems based on acoustic waves and also to expand the research on components with both bearings and gears.

Data Availability

The vibration data, consisting of tacho signal and accelerometer signals used in the research is found at [36]

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research work is funded by the Botswana International University of Science and Technology under the grant award number S00193.

References

[1] H. Ahmed and A. Nandi, "Three-stage method for rotating machine health condition monitoring using vibration signals," in *Proceedings of the 2018 Prognostics and System Health Management Conference (PHM-Chongqing)*, pp. 285–291, Chongqing, China, January 2019.

[2] P. K. Shukla, N. Ankur, and D. Abhishek, "Condition monitoring of 3-? A.C induction motor using PLC," *International Journal of Scientific Research*, vol. 4, no. 5, pp. 955–958, 2015, https://www.ijsr.net/archive/v4i5/SUB154207.pdf.

[3] H. Barksdale, Q. Smith, and M. Khan, "Condition monitoring of electrical machines with internet of things," in *Proceedings of the Conference Proceedings - IEEE SOUTHEASTCON*, St. Petersburg, FL, USA, October 2018.

[4] X. Li, Y. Yang, I. Bennett, and D. Mba, "Condition monitoring of rotating machines under time-varying conditions based on adaptive canonical variate analysis," *Mechanical Systems and Signal Processing*, vol. 131, pp. 348–363, 2019.

[5] A. Powell George, "Bearing Maintenance - Support - Bones Bearings," 2020, https://bonesbearings.com/support/maintenance/.

[6] K. Feng, K. Wang, M. Zhang, Q. Ni, and M. J. Zuo, "A diagnostic signal selection scheme for planetary gearbox vibration monitoring under non-stationary operational conditions," *Measurement Science and Technology*, vol. 28, no. 3, Article ID 035003, 2017.

[7] Y. Qin, C. Li, F. Cao, and H. Chen, "A fault dynamic model of high-speed angular contact ball bearings," *Mechanism and Machine Theory*, vol. 143, Article ID 103627, 2020.

[8] G. Chen, G. Lu, and P. Yan, "An integrated method for condition monitoring of rotating machines," in *Proceedings of the 2018 Prognostics and System Health Management Conference PHM-Chongqing*, pp. 144–149, Chongqing, China, January 2019.

[9] T. Wang, G. Lu, and P. Yan, "Multi-sensors based condition monitoring of rotary machines: an approach of multidimensional time-series analysis," *Measurement*, vol. 134, pp. 326–335, 2019.

[10] M. M. Hasan, D. Chowdhury, and M. Z. R. Khan, "Non-intrusive load monitoring using current shapelets," *Applied Sciences*, vol. 9, no. 24, 5363 pages, 2019.

[11] A. Kumar, G. Vashishtha, C. P. Gandhi, Y. Zhou, A. Glowacz, and J. Xiang, "Novel convolutional neural network (NCNN) for the diagnosis of bearing defects in rotary machinery," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–10, 2021.

[12] D. Chowdhury, M. Hasan, and M. Z. Rahman, "Statistical features extraction from current envelopes for non- intrusive appliance load monitoring," in *Proceedings of the Conference proceedings ieee southeastcon, 2020 southeastcon*, Raleigh, NC, USA, November 2020.

[13] H. O. A. Ahmed, M. L. Dennis Wong, and A. K. Nandi, "Classification of Bearing Faults Combining Compressive Sampling, Laplacian Score, and Support Vector Machine," in *Proceedings of the IECON 2017 - 43rd Annual Conference of the IEEE Industrial Electronics Society*, Beijing, China, December 2017.

[14] X. Zhang, Q. Zhang, M. Chen, Y. Sun, X. Qin, and H. Li, "A two-stage feature selection and intelligent fault diagnosis method for rotating machinery using hybrid filter and wrapper method," *Neurocomputing*, vol. 275, pp. 2426–2439, 2018.

[15] R. B. Amir, S. T. Gul, and A. Q. Khan, "A Comparative Analysis of Classical and One Class SVM Classifiers for Machine Fault Detection Using Vibration Signals," in *Proceedings of the 2016 International Conference on Emerging Technologies (ICET)*, Islamabad, Pakistan, January 2016.

[16] S. Zgarni and A. Braham, "Classification of bearing fault detection using multiclass SVM: a comparative study," in *Proceedings of the 2018 15th International Multi-Conference on Systems, Signals & Devices (SSD)*, pp. 888–892, Yasmine Hammamet, Tunisia, December 2018.

[17] S. Gunasekaran, S. E. Pandarakone, K. Asano, Y. Mizuno, and H. Nakamura, "Condition Monitoring and Diagnosis of Outer Raceway Bearing Fault Using Support Vector Machine," in *Proceedings of the 2018 Condition Monitoring and Diagnosis (CMD)*, Perth, WA, Australia, November 2018.

[18] M. Sarkar, D. Chowdhury, C. Shahnaz, and S. A. Fattah, "Application of electrical network frequency of digital recordings for location-stamp verification," *Applied Sciences*, vol. 9, no. 15, p. 3135, 2019.

[19] T. Plante, L. Stanley, A. Nejadpak, and C. X. Yang, "Rotating machine fault detection using principal component analysis of vibration signal," in *Proceedings of the 2016 Ieee Autotestcon*, October 2016.

[20] I. Jaffel, O. Taouali, E. Elaissi, and H. Messaoud, "A new online fault detection method based on PCA technique," *IMA Journal of Mathematical Control and Information*, vol. 31, no. 4, pp. 487–499, 2013.

[21] N. Atanasov, Z. Zhekov, I. Grigorov, and M. Alexandrova, "Application of principal component analysis for fault detection of DC motor parameters," *Advances in Intelligent Systems and Computing*, vol. 680, pp. 312–322, 2018.

[22] S. Ding, P. Zhang, A. Naik, P. Deng, and W. Gui, "On the Application of PCA Technique to Fault Diagnosis," *Tsinghua science and technology*, vol. 15, no. 2, 2010.

[23] F. Jia, Y. Lei, N. Lu, and S. Xing, "Deep normalized convolutional neural network for imbalanced fault classification of machinery and its understanding via visualization," *Mechanical Systems and Signal Processing*, vol. 110, pp. 349–367, 2018.

[24] J. Wang, S. Li, Z. An, X. Jiang, W. Qian, and S. Ji, "Batch-normalized deep neural networks for achieving fast intelligent fault diagnosis of machines," *Neurocomputing*, vol. 329, pp. 53–65, 2019.

[25] Z. Li, Y. Wang, and K. Wang, "A deep learning driven method for fault classification and degradation assessment in mechanical equipment," *Computers in Industry*, vol. 104, pp. 1–10, 2019.

[26] A. González-Muñiz, I. Díaz, and A. A. Cuadrado, "DCNN for condition monitoring and fault detection in rotating machines and its contribution to the understanding of machine nature," *Heliyon*, vol. 6, no. 2, Article ID e03395, 2020.

[27] C. Wu, P. Jiang, C. Ding, F. Feng, and T. Chen, "Intelligent fault diagnosis of rotating machinery based on one-dimensional convolutional neural network," *Computers in Industry*, vol. 108, pp. 53–61, 2019.

[28] X. Zhu, D. Hou, P. Zhou et al., "Rotor fault diagnosis using a convolutional neural network with symmetrized dot pattern images," *Measurement*, vol. 138, pp. 526–535, 2019.

[29] X. Li, X.-D. Jia, W. Zhang, H. Ma, Z. Luo, and X. Li, "Intelligent cross-machine fault diagnosis approach with deep auto-encoder and domain adaptation," *Neurocomputing*, vol. 383, pp. 235–247, 2020.

[30] C. Li, S. Li, A. Zhang, Q. He, Z. Liao, and J. Hu, "Meta-learning for few-shot bearing fault diagnosis under complex working conditions," *Neurocomputing*, vol. 439, pp. 197–211, 2021.

[31] Y. Yang, H. Zheng, Y. Li, M. Xu, and Y. Chen, "A fault diagnosis scheme for rotating machinery using hierarchical symbolic analysis and convolutional neural network," *ISA Transactions*, vol. 91, pp. 235–252, 2019.

[32] P. Chen, Y. Li, K. Wang, and M. J. Zuo, "A novel knowledge transfer network with fluctuating operational condition adaptation for bearing fault pattern recognition," *Measurement*, vol. 158, Article ID 107739, 2020.

[33] T. Han, C. Liu, W. Yang, and D. Jiang, "Learning transferable features in deep convolutional neural networks for diagnosing

unseen machine conditions," *ISA Transactions*, vol. 93, pp. 341–353, 2019.

[34] X. Yan, Y. Liu, and M. Jia, "Multiscale cascading deep belief network for fault identification of rotating machinery under various working conditions," *Knowledge-Based Systems*, vol. 193, Article ID 105484, 2020.

[35] S. Tang, C. Shen, D. Wang, S. Li, W. Huang, and Z. Zhu, "Adaptive deep feature learning network with Nesterov momentum and its application to rotating machinery fault diagnosis," *Neurocomputing*, vol. 305, pp. 1–14, 2018.

[36] H. Huang and N. Baddour, "Bearing vibration data collected under time-varying rotational speed conditions," *Data in Brief*, vol. 21, pp. 1745–1749, 2018.

[37] M. Vishwakarma, R. Purohit, V. Harshlata, and P. Rajput, "Vibration analysis & condition monitoring for rotating machines: a review," *Materials Today Proceedings*, vol. 4, no. 2, pp. 2659–2664, 2017.

[38] Y. Pei, "Linear principal component discriminant analysis," in *Proceedings of the 2015 IEEE International Conference on Systems, Man, and Cybernetics*, pp. 2108–2113, SMC, Hong Kong, China, January 2016.

[39] A. Kumar, C. P. Gandhi, G. Vashishtha et al., "VMD based trigonometric entropy measure: a simple and effective tool for dynamic degradation monitoring of rolling element bearing," *Measurement Science and Technology*, vol. 33, no. 1, Article ID 014005, 2021.

[40] F. Li, B. R. Upadhyaya, and S. R. P. Perillo, "Fault diagnosis of helical coil steam generator systems of an integral pressurized water reactor using optimal sensor selection," *IEEE Transactions on Nuclear Science*, vol. 59, no. 2, pp. 403–410, 2012.

[41] H. Guo and M. K. Liu, "Induction motor faults diagnosis using support vector machine to the motor current signature," in *Proceedings of the 2018 IEEE Industrial Cyber-Physical Systems, ICPS*, pp. 417–421, St. Petersburg, Russia, June 2018.

[42] R. Yuan, Z. Li, X. Guan, and L. Xu, "An SVM-based machine learning method for accurate Internet traffic classification," *Information Systems Frontiers*, vol. 12, no. 2, pp. 149–156, 2010.

[43] G. Rohith, "Support Vector Machine — Introduction to Machine Learning Algorithms," 2018, https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47.

[44] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, pp. 121–167, 1998.

[45] H. Can, L. Juelong, and X. Jianchun, "A new model for structural damage assessment using adaptive mutation particle swarm optimization and support vector machine," in *Proceedings of the 2018 Chinese Control and Decision Conference*, pp. 6711–6714, CCDC, Shenyang, China, July 2018.

[46] M. M. Hasan, D. Chowdhury, A. Shahir, and M. K. Hasan, "Statistical features extraction and performance analysis of supervised classifiers for non-intrusive load monitoring," *Engineering letters*, vol. 27, 2019.

[47] M. W. Ashour, F. Khalid, A. A. Halin, and L. N. Abdullah, "Machining process classification using PCA reduced histogram features and the Support Vector Machine," in *Proceedings of the 2015 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, pp. 414–418, Kuala Lumpur, Malaysia, February 2015.

[48] I. El-Naqa, Y. Yongyi Yang, M. N. Wernick, N. P. Galatsanos, and R. M. Nishikawa, "A support vector machine approach for detection of microcalcifications," *IEEE Transactions on Medical Imaging*, vol. 21, no. 12, pp. 1552–1563, 2002.

[49] B. O. Liu, Z. F. Hao, and X. W. Yang, "Nesting Support Vector Machinte for Muti-Classification," in *Proceedings of the 2005 International Conference on Machine Learning and Cybernetics*, pp. 18–21, Guangzhou, China, November 2005.

[50] F. Chang, H. C. Chen, and H. C. Liu, "Double K-folds in SVM," in *Proceedings of the 2015 9th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, pp. 384–387, Santa Catarina, Brazil, October 2015.

[51] S. Wang, J. Xiang, Y. Zhong, and Y. Zhou, "Convolutional neural network-based hidden Markov models for rolling element bearing fault identification," *Knowledge-Based Systems*, vol. 144, pp. 65–76, 2018.

[52] S. T. Gul, M. Imran, and A. Q. Khan, "An Online Incremental Support Vector Machine for Fault Diagnosis Using Vibration Signature Analysis," in *Proceedings of the 2018 IEEE International Conference on Industrial Technology (ICIT)*, Lyon, France, April 2018.

[53] L. Wang, Y. Yan, Y. Hu, and X. Qian, "Intelligent condition monitoring of rotating machinery through electrostatic sensing and signal analysis," in *Proceedings of the 2013 IEEE International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA)*, pp. 26-27, Kuala Lumpur, Malaysia, November 2013.

[54] X. Liu, Q. Zhou, and H. Shen, "Real-time fault diagnosis of rotating machinery using 1-D convolutional neural network," in *Proceedings of the 2018 5th International Conference on Soft Computing & Machine Intelligence (ISCMI)*, pp. 104–108, Nairobi, Kenya, November 2018.

Hindawi
Mathematical Problems in Engineering
Volume 2022, Article ID 3145402, 18 pages
https://doi.org/10.1155/2022/3145402

Research Article

Condition Monitoring of Mechanical Components Based on MEMED-NLOPE under Multiscale Features

Xuan Wang,[1] **Bo She**,[1] **Zhangsong Shi**,[1] **Shiyan Sun**,[1] **and Fenqi Qin**[2]

[1]*Department of Weaponry Engineering, Naval University of Engineering, Wuhan 430033, China*
[2]*713 Research Institute of China Shipbuilding, Zhenzhou 450000, China*

Correspondence should be addressed to Bo She; she1611@126.com

Received 28 February 2022; Accepted 18 March 2022; Published 11 April 2022

Academic Editor: Ke Feng

An increasing popularity of researches focuses on the vibration signal with the characteristics of nonstationary, nonlinear, and strong noise interference. A nonlinear dimension and feature reduction method called multiple empirical mode entropy decomposition-nonlocal orthogonal preserving embedding (MEMED-NLOPE) is proposed to implement condition monitoring in this paper. Different from multiple empirical mode decomposition (MEMD), MEMED adopts maximum entropy method, which can directly output the subsignal with the maximum correlation and realize nonlinear dimensionality reduction. Besides, multiscale feature extraction method is used during preprocessing nonlinear data process, which realizes feature reduction. Finally, nonlocal orthogonal preserving embedding algorithm-exponentially weighted moving average (NLOPE-EWMA) realizes the automatic detection of the fault. Taking the laboratory rolling bearing test and naval gun pendulum mechanism test as cases, the effectiveness of MEMED-NLOPE is verified.

1. Introduction

Mechanical components as the vital parts of mechanical equipment are prone to wear and cracks on the surface with long-term overload operation. Wear increases the mechanical components transmission error, generally resulting in increased vibration, noise, and dynamic loads [1]. If the early minor damage of components cannot be detected in time, once the fatigue deteriorates and the parts break, the mechanical equipment will be shut down. With the deterioration of the fault degree, the mechanical equipment may be shut down for a long time, resulting in catastrophic failures and unexpected economic losses [2]. Therefore, condition monitoring of mechanical components is an effective measure to avoid the continuous deterioration of parts after damage. Vibration signals are widely used to characterize the state of mechanical equipment because of their ease of acquisition, but usually the collected vibration signals have many interference components and have nonstationary and nonlinear characteristics, which also bring difficulties to fault diagnosis.

In recent years, the multivariate statistical process monitoring (MSPM) technology is often used to detect faults in industrial production processes, such as partial least squares (PLS) [3], principal component analysis (PCA) [4], and independent component analysis (ICA) [5]. Those traditional monitoring methods process the intermediate data by dimensionality reduction and extract a small number of components to construct the monitoring statistics that can reflect the characteristics of the original data. At this time, the performance of dimensionality reduction will affect the monitoring effect.

Different from the dimensionality reduction method that maintains the global data structure, manifold learning is used to maintain the characteristics of local data structure, such as locally linear embedding (LLE) [6], Laplacian eigenmap (LE) [7], local preserving projections (LPP) [8], and neighborhood preserving embedding (NPE) [9]. Both LPP and NPE belong to linear projection methods, but these methods may lose the key information contained in the global data structure because they only consider the neighborhood relationship to maintain the local

characteristics. Therefore, in order to consider the global and local data structure characteristics, a method combining LPP and PCA method is proposed [10, 11]. The test results show that its monitoring performance is better than that of single method. Besides, orthogonal neighborhood preserving embedding (ONPE) is developed from NPE [12]; by setting additional orthogonal constraints on the projection vector, it not only maintains the characteristics of local structure but also avoids the distortion defects of NPE [13]. In order to fully consider the global and local structure characteristics of data, combined with the basic principles of PCA and ONPE algorithm, a nonlocal orthogonal preserving embedding (NLOPE) algorithm is proposed [14]. However, those methods still belong to linear method and have limitations in dealing with nonlinear data.

In the data preprocessing stage, empirical mode decomposition (EMD) is often used to describe the characteristics of nonlinear and nonstationary signals [15]. However, when processing multiple signals (multichannel signals), EMD may lead to different number and frequency scale of IMF for signal decomposition of each channel [16]. The proposal of multivariate empirical mode decomposition (MEMD) [17] ensures the matching of IMF components in quantity and scale. However, in the process of data preprocessing, the dimension of subsignals and features may increase, which will affect the effect of condition monitoring. To realize subsignals and nonlinear dimensionality reduction, entropy has been widely developed and used in this field, which can measure the correlation, uncertainty, and complexity of signals and features [18, 19].

In this paper, a linear dimension and feature reduction method called multiple empirical mode entropy decomposition-nonlocal orthogonal preserving embedding (MEMED-NLOPE) is proposed on the basis of MEMD and NLOPE. To reduce the redundancy of subsignal set and reduce the complexity of the system, MEMED takes both advantages of MEMD and maximum entropy method into account. To verify the effectiveness of MEMED-NLOPE, MEMED-NLOPE and MEMD-NLOPE are employed to detect the faults of naval gun pendulum mechanism, and MEMED-NLOPE is verified by the experimental data set of rolling bearing in laboratory.

The rest of the paper is organized as follows. MEMD, PCA, ONPE, and NLOPE are reviewed and analyzed in Section 2. The proposed MEMED-NLOPE is developed in Section 3. In Section 4, two cases are used to demonstrate the effectiveness of the proposed method. Finally, conclusions are drawn in Section 5.

2. Background Techniques

2.1. MEMD. EMD is suitable for one-dimensional real signals. For the processing of multichannel signals (multichannel signals), the EMD method often needs to solve the single channel signals, respectively, which may lead to the different number and frequency scale of IMF decomposed by each channel signal; that is, there is the problem of oscillation mode calibration of different channels, which is not conducive to the synchronous correlation analysis between multichannel signal channels. Although CMED [20], BEMD [21], and TEMD [22] are Multivariate Applications of EMD methods in multivariate data, they are limited to multivariate data: only binary and ternary signals. For real multivariate signals, it is still impossible to decompose the signal on the premise of correctly analyzing the physical meaning of the signal. The proposal of MEMD realizes the multichannel synchronous joint analysis of multichannel signal oscillation modes, obtains the common modes of different channels, ensures the matching of intrinsic mode function (IMF) components in quantity and scale, and solves the problem of mode calibration of multichannel signals.

The specific implementation of MEMD can be summarized as the following steps.

Let an n-dimensional vector group sequence $\{V(t)\}_{t=1}^{T} = \{v_1(t), v_2(t), \ldots, v_n(t)\}$ represent a n-tuple signal, the length of the signal sequence is T, and $X^{\theta_k} = \{x_1^k, x_2^k, \ldots, x_n^k\}$ represents the direction vector set of the corresponding angle $\theta^k = \{\theta_1^k, \theta_2^k, \ldots, \theta_{n-1}^k\}$ on the $n-1$-dimensional sphere. If you want to establish K direction vectors in spherical space, then $k = 1, 2, 3, \ldots, K$.

(1) The Hammersley sequence sampling method is used to obtain a suitable set of uniform sampling points on the $n-1$-dimensional sphere, that is, the direction vector of the n-dimensional space.

(2) The mapping $p^{\theta_k}(t)$ of the input signal $v(t)$ on each direction vector X^{θ_k} is calculated.

(3) Determine the instantaneous time $\{t^{\theta_k}\}_{k=1}^{K}$ corresponding to the extreme value of the mapping signal $\{p^{\theta_k}(t)\}_{k=1}^{K}$ of all direction vectors, and l represents the extreme point position, $l \in [1, T]$.

(4) The extreme point $[t_l^{\theta_k}, v(t_l^{\theta_k})]$ is interpolated by multivariate spline interpolation function to obtain K multivariate envelopes $\{e^{\theta_k}(t)\}_{k=1}^{K}$.

(5) For K direction vectors in spherical space, the mean $m(t)$ of n-tuple signal is as follows:

$$m(t) = \frac{1}{K} \sum_{K=1}^{K} e^{\theta_k}(t). \tag{1}$$

(6) Extract the intrinsic mode function $h(t)$ through $h(t) = v(t) - m(t)$. If $h(t)$ meets the judgment standard of multivariate IMF, take the $v(t) - h(t)$ result as the input signal in step (2), continue the iterative calculation in steps (2)~(6), and extract a new multivariate IMF component $h(t)$; otherwise, take $h(t)$ as the input signal of step (2) and continue the iteration of steps (2)~(6).

After a series of MEMD decomposition processes, similar to the EMD algorithm, the original n-tuple signal $\{V(t)\}_{t=1}^{T} = \{v_1(t), v_2(t), \ldots, v_n(t)\}$ is decomposed into a series of addition forms of IMF $(h_i(t)\}_{i=1}^{q})$ and Residual $r(t)$, as follows:

$$V(t) = \sum_{i=1}^{q} h_i(t) + r(t), \tag{2}$$

where q represents the decomposed multivariate IMF function, $h(t)$ is $\{h_i^1(t), h_i^2(t), \ldots, h_i^n(t)\}_{t=1}^T$, $r(t)$ is $\{r^1(t), r^2(t), \ldots, r^n(t)\}_{t=1}^T$, corresponding to n groups of IMF components and n margins of n-tuple signals, respectively. The number of IMF decomposed by each channel of multivariate signal is the same, and the frequencies of IMF in each layer are different. The first decomposed IMF frequency is high, and then the decomposed IMF frequency is low, and the decomposed residual frequency is the lowest. The IMF corresponding to each variable of n-tuple signal is aligned according to the frequency scale in n channels to form multiple IMF.

2.2. Principal Component Analysis. PCA, namely, principal component analysis, is one of the most widely used data dimensionality reduction algorithms. In this study, the PCA algorithm is implemented based on eigenvalue decomposition covariance matrix. The specific steps are as follows:

Step 1: input data set $X = \{x_1, x_2, x_3, \ldots, x_n\}$, which needs to be reduced to k dimension.

Step 2: deaveraging (i.e., decentralization), that is, each feature subtracts its own average.

Step 3: calculate the covariance matrix $(1/n)XX^T$. Note: dividing or not dividing the number of samples n or $n-1$ has no effect on the calculated eigenvector.

Step 4: find the eigenvalue and eigenvector of covariance matrix $(1/n)XX^T$ by eigenvalue decomposition method.

Step 5: sort the eigenvalues from large to small, and select the largest k of them. Then, the corresponding x eigenvectors are used as row vectors to form the eigenvector matrix P.

Step 6: convert the data into a new space constructed by k eigenvectors, i.e., $Y = PX$.

2.3. Orthogonal Neighborhood Preserving Embedding. Given data set $X = \{x_1, x_2, \ldots, x_N\} \in R^m$, as a kind of linear dimensionality reduction method, the goal of the orthogonal neighborhood preserving embedding (ONPE) algorithm is to reduce the dimension of high-dimensional data X to low-dimensional data $Y = \{y_1, y_2, \ldots, y_N\} \in R^d$, that is, $Y = A^T X$, using a transformation matrix $A = [a_1, a_2, \ldots, a_d] \in R^{m\times d}$ $(d < m)$, and the low-dimensional data can express the essential characteristics of the original high-dimensional data. The NPE algorithm is the basic form of ONPE. NPE maintains the local characteristics in the data structure by constructing the neighborhood graph between adjacent samples. Therefore, each sample can be expressed as a linear combination of adjacent samples and their corresponding weight coefficients. The weight coefficient matrix W minimizes the following objective functions:

$$\min \sum_i \left\| x_i - \sum_j W_{ij} x_j \right\|^2. \tag{3}$$

In order to fully maintain the local characteristics of the data structure, the high-dimensional spatial data x_i are mapped to the low-dimensional feature space to obtain y_i, and the weight coefficients between x_i and its nearest neighbors will be projected to the low-dimensional feature space to be saved to characterize the connection relationship between y_i and its nearest neighbors. The low-dimensional mapping Y of high-dimensional data X can calculate the following loss functions:

$$\begin{cases} \min \quad \sum_i \left\| y_i - \sum_j W_{ij} y_j \right\|^2, \\ \text{s.t.} \quad Y^T Y = A^T X X^T A = I, \end{cases} \tag{4}$$

where $\sum_{j=1}^k W_{ij} = 1$, $i = 1, 2, \ldots, N$, k is the number of nearest neighbors in the neighborhood of x_i. If x_j is not the nearest neighbor of x_i, there is $W_{ij} = 0$.

ONPE adds an orthogonal constraint on the basis of NPE, that is, mapping high-dimensional data to low-dimensional feature space through an orthogonal projection matrix A. According to (3) and (4), the projection matrix is calculated by the following formulas:

$$\begin{cases} a_1 = \arg\min_a \sum_i \left\| y_i - \sum_j W_{ij} y_j \right\|^2 = \arg\min_a A^T X M X^T A, \\ \text{s.t.} A^T X X^T A = I, \end{cases}$$
$$\begin{cases} a_k = \arg\min_a \sum_i \left\| y_i - \sum_j W_{ij} y_j \right\|^2 = \arg\min_a A^T X M X^T A, \\ \text{s.t.} a_k^T a_1 = a_k^T a_2 = \cdots = a_k^T a_{k-1} = 0, \\ A^T X X^T A = I, \end{cases} \tag{5}$$

where $k = 2, 3, \ldots, d$, $M = (I - W)^T (I - W)$. Through iterative calculation by Lagrange operator, the expression of orthogonal matrix A is as follows:

(a) a_1 is the eigenvector corresponding to the minimum eigenvalue of matrix $(XX^T)^{-1} X M X^T$;

(b) a_k is the eigenvector corresponding to the minimum eigenvalue of matrix $Q^{(k)}$, where $Q^{(k)}$ is

$$Q^{(k)} = \left\{ I - \left(XX^T\right)^{-1} A^{(k-1)} \left[\left(A^{(k-1)}\right)^T \left(XX^T\right)^{-1} A^{(k-1)} \right]^{-1} \left(A^{(k-1)}\right)^T \right\} \cdot \left(XX^T\right)^{-1} X M X^T, \tag{6}$$

where $A^{(k-1)} = [a_1, a_2, \ldots, a_{k-1}]$.

2.4. Objective Function of Nonlocal Orthogonal Preserving Embedding. In order to fully consider the global and local structure characteristics of data, combined with the basic principles of PCA and ONPE algorithm, a nonlocal orthogonal preserving embedding (NLOPE) algorithm is proposed. Assuming data set $x = \{x_1, x_2, \cdots, x_N\} \in R^{m\times N}$, the objective function of NLOPE is as follows:

$$\begin{aligned} J(a)_{\text{NLOPE}} &= \eta J(a)_{\text{Local}} - (1-\eta) J(a)_{\text{Global}} \\ &= \eta \min_a a^T x M x^T a - (1-\eta) \max_a a^T C a \\ &= \min_a a^T \left(\eta x M x^T - (1-\eta) C\right) a \\ &= \min_a a^T \left(\eta L' - (1-\eta) C\right) a, \end{aligned} \tag{7}$$

$$\begin{aligned} \text{s.t. } & a_k^T a_1 = a_k^T a_2 = \cdots = a_k^T a_{k-1} = 0, \\ & a^T \left[\eta x x^T + (1-\eta) I\right] a = 1, \end{aligned} \tag{8}$$

where $C = (1/N)\sum_{i=1}^{N}(x_i - \overline{x})(x_i - \overline{x})^T$, $\overline{x} = (1/N)\sum_{i=1}^{N} x_i$.

Using Lagrange operator, projection matrix A can be obtained by calculating the following feature decomposition problem:

(a) a_1 is the eigenvector corresponding to the minimum eigenvalue of matrix $S^{-1}D$;

(b) a_k is the eigenvector corresponding to the minimum eigenvalue of matrix $Q^{(k)}$, where $Q^{(k)}$ is

$$Q^{(k)} = \left\{ I - (S)^{-1} a^{(k-1)} \left[\left(a^{(k-1)}\right)^T (S)^{-1} a^{(k-1)} \right]^{-1} \left(a^{(k-1)}\right)^T \right\} S^{-1} D, \tag{9}$$

where $k = 2, 3, \ldots, d$, d is the dimensions of data in NLOPE feature space, + $a^{(k-1)} = [a_1, a_2, \ldots, a_{k-1}]$, $S = \eta x x^T + (1-\eta) I$, and $D = \eta x M x^T - (1-\eta) C$.

For the new sample x_{new}, the mapping in the low-dimensional NLOPE feature space is

$$y_{\text{new}} = A^T x_{\text{new}}, \tag{10}$$

where $A = [a_1, a_2, \ldots, a_k]$.

The detailed derivation and calculation of projection matrix A are shown in [12]; there is

$$\left\{ I - S^{-1} a^{(k-1)} \left[\left(a^{(k-1)}\right)^T S^{-1} a^{(k-1)} \right]^{-1} \left(a^{(k-1)}\right)^T \right\} S^{-1} D a_k = \lambda a_k. \tag{11}$$

Thus, a_k is the eigenvector corresponding to the minimum eigenvalue of matrix $Q^{(k)}$, and the expression of $Q^{(k)}$ is as follows:

$$Q^{(k)} = \left\{ I - (S)^{-1} a^{(k-1)} \left[\left(a^{(k-1)}\right)^T (S)^{-1} a^{(k-1)} \right]^{-1} \left(a^{(k-1)}\right)^T \right\} S^{-1} D. \tag{12}$$

2.5. Calculation of Parameters. In the construction of the NLOPE model, parameter η makes the global data structure characteristics and local data structure characteristics occupy different components in the above model. The selection of parameter η affects the extraction of potential features in the data and then affects the effect of mechanical equipment fault detection, fault detection, and degradation performance evaluation.

It can be seen from (12) that the objective function of NLOPE is composed of two subobjective functions. Therefore, the objective function optimization problem of the NLOPE model is essentially a double objective optimization problem. Usually, it is difficult to obtain the optimal solution of the two subobjective functions at the same time. However, by balancing the two subobjective functions, a relatively better solution can be obtained.

By balancing the global data structure characteristics and local data structure characteristics of the model, the calculation of parameter η is as follows:

$$\eta S_{\text{Local}} = (1-\eta) S_{\text{Global}}, \tag{13}$$

where $S_{\text{Global}} = \rho(C)$ and $S_{\text{Local}} = \rho(L')$ represent the energy changes of $J(a)_{\text{local}}$ and $J(a)_{\text{local}}$, respectively.

According to (7), parameter η is used to balance matrix L' and matrix C in the NLOPE model, which can be regarded as the energy change of balance L' and C. Based on the principle of the PCA method, the eigenvectors corresponding to the first few large eigenvalues can characterize the distribution of matrix energy. Therefore, the maximum eigenvalues of matrix L' and matrix C can be used to estimate the energy change.

In the NLOPE model, parameter η is calculated as follows:

$$\eta = \frac{\rho(C)}{\rho(L') + \rho(C)}, \tag{14}$$

where $\rho(\cdot)$ is the spectral radius of the matrix, and matrix L' and matrix C are defined in (7).

2.6. Detection Index. Hotelling's T^2 and *SPE* statistics are often used as indicators of industrial process fault detection to judge whether the production process is abnormal. Hotelling's T^2 is used to measure the change of sample variables in the potential variable space, and *SPE* is mainly used to measure the change of sample variables in the residual space. When the statistics T^2 or *SPE* exceed their respective control limits, it indicates that the process may be abnormal. T^2 and *SPE* are calculated as follows:

$$T^2 = y^T \Lambda^{-1} y, \tag{15}$$

where y is the low-dimensional feature sample of sample x projected in NLOPE feature space, and $\Lambda = (yy^T/(N-1))$ is the covariance matrix of the projection vector of training sample in NLOPE feature space.

$$\begin{aligned} SPE &= \langle \Phi(x), \Phi(x) \rangle - \langle y, y \rangle \\ &= k(x,x) - \frac{2}{N}\sum_{i=1}^{N} k(x_i, x) + \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} k(x_i, x_j) - y^T y \\ &= 1 - \frac{2}{N}\sum_{i=1}^{N} k(x_i, x) + \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} k(x_i, x_j) - y^T y. \end{aligned} \tag{16}$$

Among them, $y_{\text{new}} = A^T x_{\text{new}}$.

In order to detect the early faults of mechanical equipment more accurately and reliably, the exponential weighted moving average (EWMA) is proposed by integrating T^2 and SPE statistics. The statistic U is a linear combination of T^2 and SPE statistics, including

$$U = \frac{T^2}{LT_2} + \frac{SPE}{L_{SPE}}, \tag{17}$$

where L_{T^2} and L_{SPE} are the control limits of statistics T^2 and SPE, respectively, which can be calculated by kernel density estimation (KDE). (T^2/L_{T^2}) and (SPE/L_{SPE}) normalize T^2 and SPE to (0, 1), respectively.

The detection index $EWMA$ is calculated as follows:

$$W_t = (1-\gamma)W_{t-1} + \gamma U_t, \tag{18}$$

where W_t represents the detection index, which is composed of the current index quantity and the historical index quantity, and γ is the smoothing coefficient between (0, 1). When γ takes a larger value, the current detection quantity U_t has a larger proportion in the detection quantity W_{t-1} than the historical detection quantity W_t. The control limit of the detection amount $EWMA$ is also calculated by the kernel density estimation method. In this chapter, the smoothing coefficient γ takes the empirical value of 0.2.

3. Proposed Condition Monitoring Model

Based on the analysis of the above background techniques, the MEMED-NLOPE model is proposed. Firstly, the architecture of MEMED-NLOPE is proposed; secondly, the preprocessing stage of nonlinear data is described in detail; finally, the steps of the automatic fault detection model based on NLOPE-EWMA is described.

3.1. Proposed Architecture. The MEMED-NLOPE model is mainly divided into two parts. The first part is nonlinear data preprocessing, and the second part is automatic fault detection model.

For the first part, MEMED decomposes the signals collected by each sensor, quantitatively analyzes the correlation and orthogonality between the multiscale subsignal and the original signal, selects the subsignal with the maximum correlation for the preliminary extraction of multi domain features, and uses the feature measurement criterion based on mutual information to optimize and eliminate redundant features, as the input of the fault detection model.

For the second part, it proposes condition monitoring model, adopts PCA which extracts the correlation between multidimensional variables from the historical normal operation data, and diagnoses abnormalities through their unexpected changes, but PCA only considers the global structure relationship between samples and ignores the local structure relationship. Therefore, based on the PCA multivariate statistical process monitoring method, combined with the local orthogonal preserving embedding (ONPE) algorithm, this project proposes NLOPE, which uses exponential weighted moving average (EWMA) statistics as detection index to realize the construction of the condition monitoring model. The research scheme is shown in Figure 1.

3.2. Preprocessing Nonlinear Data

3.2.1. MEMED

(1) Maximum Entropy Method. Based on the information entropy theory, the mutual information between different subsignals and source signals is measured to characterize the correlation of subsignals, reduce the redundancy of subsignal set, and reduce the complexity of the system. The formula of information entropy is as follows:

$$\begin{aligned} H(X) &= E[I(X)] \\ &= -\sum_{i=1}^{n} p(x_i) I(x_i) \\ &= -\sum_{i=1}^{n} p(x_i) \log_b p(x_i), \end{aligned} \tag{19}$$

where $I(x_i)$ represents the amount of information of x_i:

$$\begin{aligned} I(x_i) &= \log_b\left(\frac{1}{p(x_i)}\right) \\ &= -\log_b p(x_i). \end{aligned} \tag{20}$$

$p(x_i)$ is the probability of occurrence of x_i. The number of information ontologies contained in a randomly generated event is only related to the probability of occurrence of the event. The lower the probability of an event, the larger the information ontology contained in the received information when the event really occurs. The meaning is that the event with probability 0 has a large amount of information; on the contrary, it has a small amount of information. The reason for taking logarithm is to make the product sum. Two independent events x, y: $p(x,y) = p(x) * p(y)$ and $I(x,y) = I(x) + I(y)$.

Information entropy is the mathematical expectation of information.

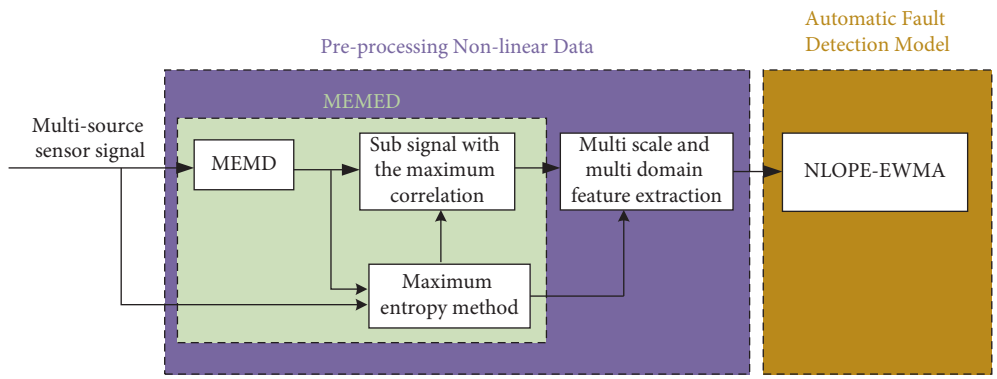


Figure 1: Research scheme of MEMED-NLOPE.

Mutual information is

$$I(x_i, y_i) = \log \frac{p(x_i, y_i)}{p(x_i)p(y_i)}. \tag{21}$$

Average mutual information is the mathematical expectation of mutual information:

$$\begin{aligned} I(X,Y) &= E[I(x_i, y_i)] \\ &= \sum_i \sum_j p(x_i, y_i) \log \frac{p(x_i, y_i)}{p(x_i)p(y_i)}. \end{aligned} \tag{22}$$

From the formula,

$$I(X,Y) = H(X) + H(Y) - H(X,Y). \tag{23}$$

(2) MEMED Flow Chart. Based on MEMD, MEMED adopts the maximum entropy method to output the subsignal with the maximum correlation, which realizes the function of dimension reduction and avoid data explosion. The specific flow of MEMD is shown in Figure 2.

3.2.2. Multiscale Feature Extraction Method

(1) Fault Feature Construction Method. The vibration signal is used to evaluate the running state of mechanical equipment. Generally, the corresponding features are extracted from the time domain, frequency domain, and time-frequency domain of the signal as the basis of diagnosis. Time-domain analysis is to describe the change of signal waveform and amplitude with time. Frequency domain analysis is to describe the change of signal power or energy with frequency. Time-frequency analysis is to study the change of signal spectrum with time and represent the distribution of signal strength or energy in both time and frequency dimensions.

Time-domain and frequency-domain features generally include root mean square, kurtosis, skewness, peak factor, spectral mean square deviation, and envelope spectral variance. Table 1 contains 11 times domain characteristic parameters ($p_1 - p_{11}$) and 13 frequency domain characteristic parameters ($p_{12} - p_{24}$). In each characteristic expression, $x(n)$ is the time-domain signal sequence, $n = 1, 2, \ldots, N$, N are the number of samples, $s(k)$ is the spectrum of signal $x(n)$, $k = 1, 2, \ldots, K$, K are the number of spectral lines, and f_k is the frequency value of the k spectral line. The time-domain characteristic parameters p_1 and $p_3 - p_5$ describe the amplitude and energy changes of the time-domain signal; p_2 and $p_6 - p_{11}$ describe the time series distribution of time-domain signals. The frequency domain characteristic parameter p_{12} describes the change of frequency domain energy; $p_{13} - p_{15}$, p_{17} and $p_{21} - p_{24}$ reflect the concentration and dispersion of the spectrum; p_{16}, $p_{18} - p_{20}$ reflects the change of the position of the main frequency band.

The time-frequency domain features include sample entropy, permutation entropy, wavelet energy entropy, and EEMD information entropy, which are generally calculated by time-frequency analysis methods such as wavelet analysis and empirical mode decomposition.

(2) Fault Feature Selection Method. Similarly, based on the information entropy theory, the mutual information between different features is measured to characterize the correlation between features and reduce the redundancy of feature sets.

3.3. Automatic Fault Detection Model Based on NLOPE-EWMA. The offline modeling steps based on NLOPE are as follows:

Step 1: the features of training samples after MEMD adaptive decomposition, multiscale subsignal selection, and multi-scale feature extraction are constructed as feature samples, and the feature samples are standardized

Step 2: calculate the projection coefficient matrix from (9)

Step 3: calculate the sum SPE statistics of all training samples, and calculate the sum of control limits, so as to calculate the detection index EWMA and its control limits

The steps of online detection based on NLOPE are as follows:

Step 1: after multiscale feature extraction, the feature samples of each test sample are constructed, and the feature test samples are standardized by using the mean and variance of the training feature samples

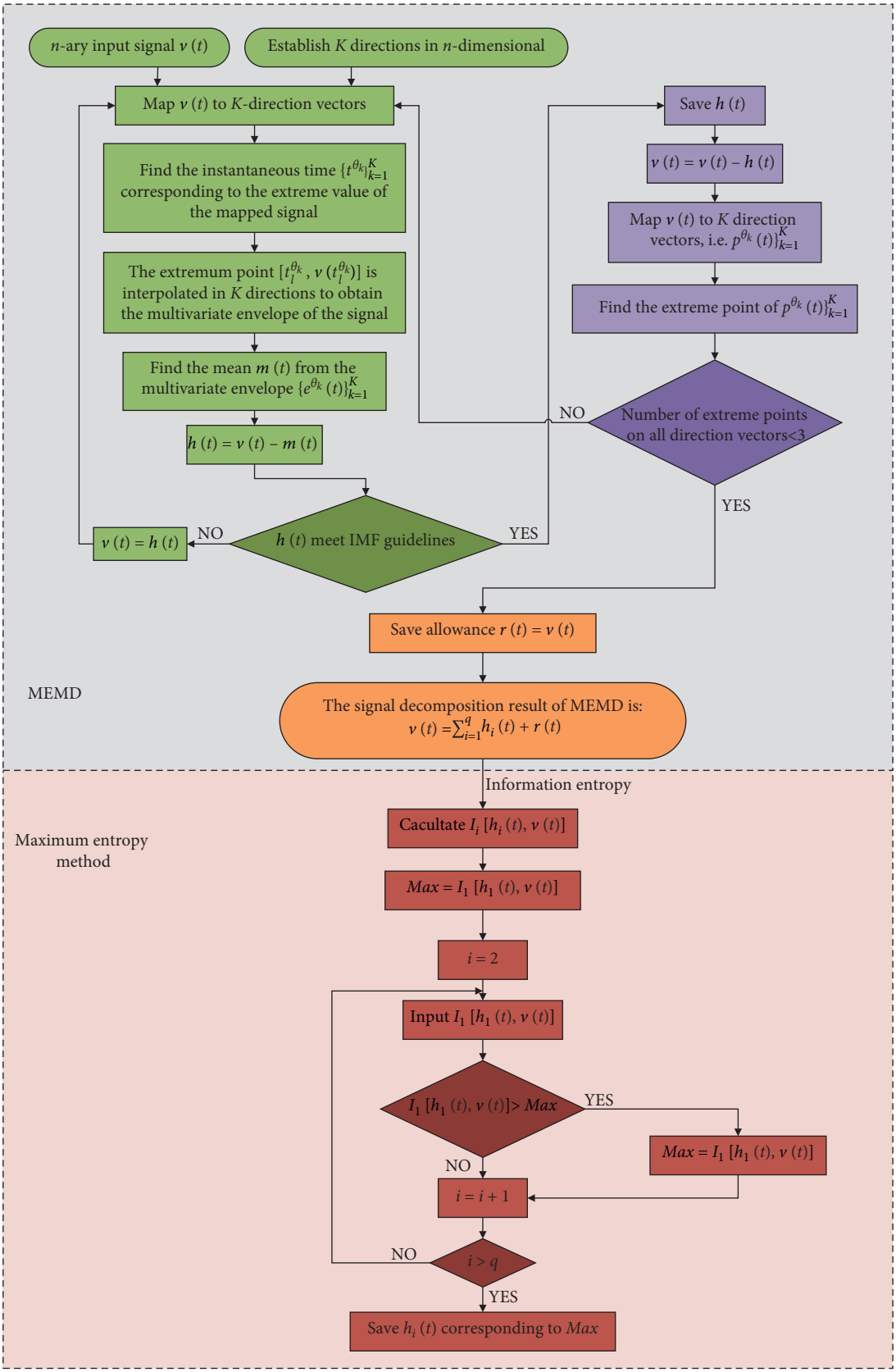


Figure 2: MEMED flow chart.

Table 1: Characteristic parameters.

No.	Characteristic expression	No.	Characteristic expression	No.	Characteristic expression
1	$p_1 = (\sum_{n=1}^{N} x(n)/N)$	9	$p_9 = (p_5/p_3)$	17	$p_{17} = \sqrt{\sum_{k=1}^{K} (f_k - p_{16})^2 s(k)/K}$
2	$p_2 = \sqrt{\sum_{n=1}^{N} (x(n) - p_1)^2/N - 1}$	10	$p_{10} = (p_4/1/N \sum_{n=1}^{N} \lvert x(n)\rvert)$	18	$p_{18} = \sqrt{\sum_{k=1}^{K} f_k^2 s(k)/\sum_{k=1}^{K} s(k)}$
3	$p_3 = (\sum_{n=1}^{N} \sqrt{\lvert x(n)\rvert}/N)^2$	11	$p_{11} = (p_5/1/N \sum_{n=1}^{N} \lvert x(n)\rvert)$	19	$p_{19} = \sqrt{\sum_{k=1}^{K} f_k^4 s(k)/\sum_{k=1}^{K} f_k^2 s(k)}$
4	$p_4 = \sqrt{\sum_{n=1}^{N} x(n)^2/N}$	12	$p_{12} = (\sum_{k=1}^{K} s(k)/K)$	20	$p_{20} = \sum_{k=1}^{K} f_k^2 s(k)/\sqrt{\sum_{k=1}^{K} s(k) \sum_{k=1}^{K} f_k^4 s(k)}$
5	$p_5 = \max\lvert x(n)\rvert$	13	$p_{13} = (\sum_{k=1}^{K} (s(k) - p_{12})^2/K - 1)$	21	$p_{21} = (p_{17}/p_{16})$
6	$p_6 = (\sum_{n=1}^{N} (x(n) - p_1)^3/(N-1)p_2^3)$	14	$p_{14} = (\sum_{k=1}^{K} (s(k) - p_{12})^3/K(\sqrt{p_{13}})^3)$	22	$p_{22} = (\sum_{k=1}^{K} (f_k - p_{16})^3 s(k)/K p_{17}^3)$
7	$p_7 = (\sum_{n=1}^{N} (x(n) - p_1)^4/(N-1)p_2^4)$	15	$p_{15} = (\sum_{k=1}^{K} (s(k) - p_{12})^4/K p_{13}^2)$	23	$p_{23} = (\sum_{k=1}^{K} (f_k - p_{16})^4 s(k)/K p_{17}^4)$
8	$p_8 = (p_5/p_4)$	16	$p_{16} = (\sum_{k=1}^{K} f_k s(k)/\sum_{k=1}^{K} s(k))$	24	$p_{24} = (\sum_{k=1}^{K} (f_k - p_{16})^{(1/2)} s(k)/K\sqrt{p_{17}})$

Step 2: calculate the projection of the test sample in the low-dimensional feature space from (10)

Step 3: calculate the detection amount EWMA corresponding to the test sample and judge whether it exceeds the monitoring limit

The fault detection process based on the NLOPE method is shown in Figure 3. Using the NLOPE method, the offline detection model is constructed by using normal vibration signals, the new data samples are input into the detection model, and the fault detection of mechanical equipment can be carried out by calculating the detection indexes.

4. Experimental Verification and Analysis

To verify the effectiveness of MEMED-NLOPE, the laboratory rolling bearing test and naval gun pendulum mechanism test are taken as cases. Besides, the programming software used in the experiment is MathWorks Matlab R2018a, and the computer configuration is Core i7-10875H CPU @ 2.30 GHz.

4.1. Experimental Data Set of Naval Gun Pendulum Mechanism

4.1.1. Experimental Design. The life test of typical mechanical parts of naval gun is carried out by using the test bench of energy storage mechanism of single 130 mm naval gun pendulum. In the test, the data of the health and fracture damage of the pressing plate and the health and crack damage of the roller are collected. The damage of mechanical parts is shown in Figures 4 and 5. In the test, six vibration acceleration sensors, numbered a1–a6, and acoustic sensors are arranged near the sliding plate and pressing plate mechanism of the pendulum. The location of the sensor measuring points is shown in Figure 6. Two vibration acceleration sensors (No. a7-a8) are arranged near the roller track, and the measuring point positions of the sensors are shown in Figures 7 and 8.

The test data collected are composed of the following:

(1) *Platen Data.* The composition of platen data collected in the test is shown in Table 2, and the sampling frequency is 10 kHz. The number of times of one test in the table indicates that the artillery test bench has completed a complete action cycle of latch closing, recoil, reentry, latch opening, lower swing, and upper swing.

(2) *Roller Data.* The roller test adopts rollers in two states, and the test data composition is shown in Table 3.

4.1.2. Condition Monitoring and Analysis of Pendulum Mechanism. MEMD-NLOPE and MEMED-NLOPE are used to monitor the condition of platen in different status, and the results are shown in Figure 9.

It can be seen from Figure 9 that under the condition monitoring model constructed by MEMD-NLOPE, when the monitoring object is the healthy platen and roller, some of the detection indicators of the test sample exceed the monitoring limit; the part exceeding the monitoring limit indicates that the detection model generates false alarm. While when the monitoring object is the platen and roller tending to be damaged, some of the detection indicators of the test sample are below the monitoring limit. Under the condition monitoring model constructed by MEMED-NLOPE, most of the detection indicators of the test samples are within the normal monitoring range, and the collected training samples are consistent with the state of the naval gun platform.

4.1.3. Summary. Through the action cycle test of mechanical mechanism on the naval gun test bench, the data information of key mechanical parts in the damaged state is obtained. Using the proposed detection model, the normal operation state and abnormal operation state of the naval gun test bench are detected and analyzed. The performance evaluation results based on MEMD-NLOPE and MEMED-NLOPE are shown in Table 4, respectively.

According to the results in Table 4, the performance of condition monitoring based on MEMED-NLOPE is better than that based on MEMD-NLOPE, and the average accuracy of normal and damage detection of platen and roller is greater than 90%, indicating that MEMED-NLOPE can determine the normal operation state of pendulum mechanism and detect the faults of mechanical components. However, when the platen is in a healthy state, the false

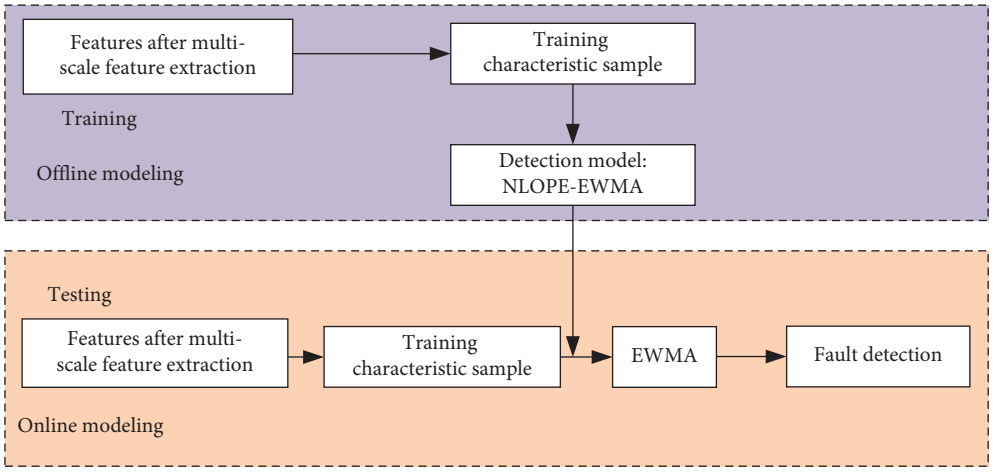


Figure 3: Fault detection process based on NLOPE-EWMA.

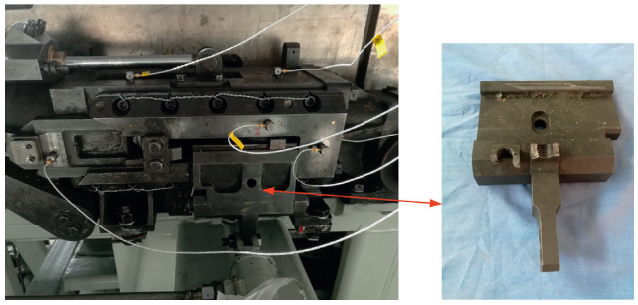


Figure 4: Damage diagram of pressing plate.

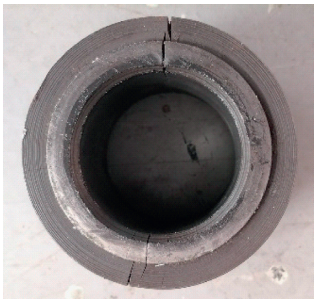


Figure 5: Crack damage of roller.

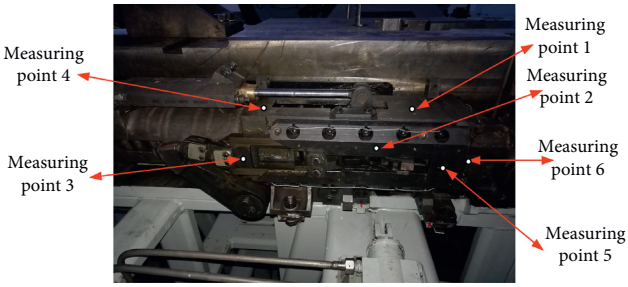


Figure 6: Layout of measuring points of acceleration sensor.

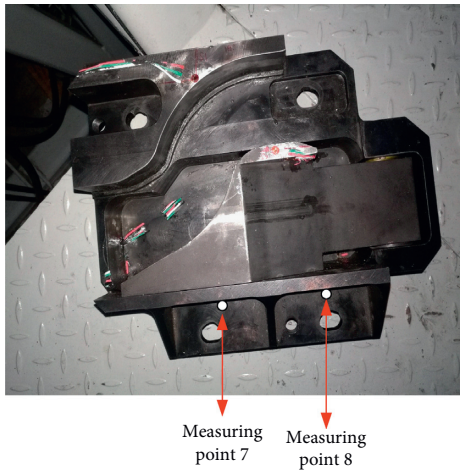


Figure 7: Layout of measuring points of acceleration sensor.

(a) (b)

Figure 8: Test bench sensor measuring points.

Table 2: Composition of platen test data.

No.	Data	Number of tests	Platen status	Data type
1	Yaban_Data1	20	Damaged status	Test data
2	Yaban_Data2	20	Damaged status	Test data
3	Yaban_Data3	20	Healthy status	Test data
4	Yaban_Data4	20	Healthy status	Test data
5	Yaban_Data5	20	Healthy status	Training data

Table 3: Composition of roller test data.

No.	Data	Number of tests	Roller status	Data type
1	Gunlun_Data1	20	Damaged status	Test data
2	Gunlun_Data2	20	Damaged status	Test data
3	Gunlun_Data3	20	Healthy status	Test data
4	Gunlun_Data4	20	Healthy status	Test data
5	Gunlun_Data5	20	Healthy status	Training data

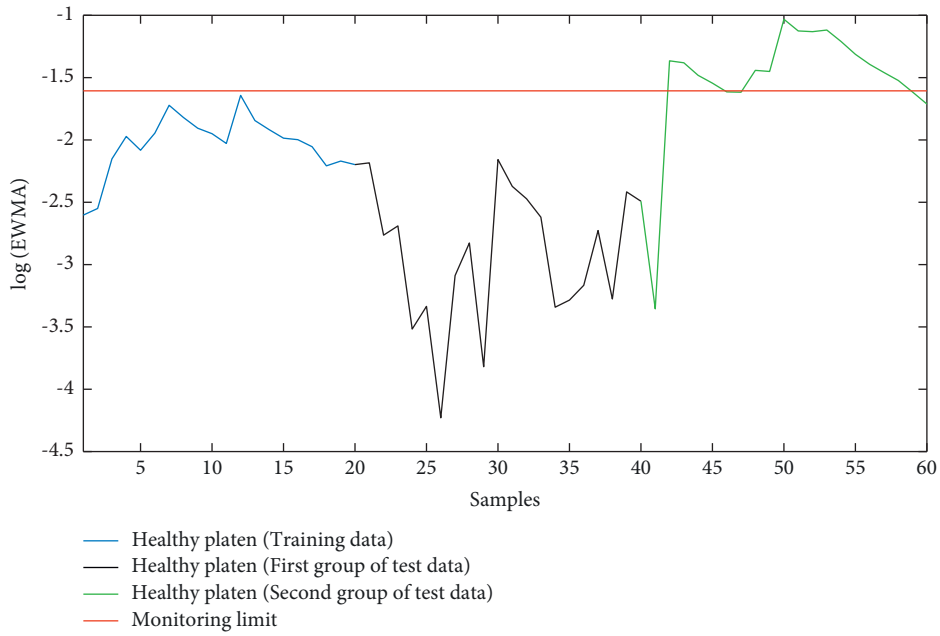
alarm rate is little high, which reflects that MEMD-NLOPE has certain instability in this test.

4.2. Experimental Data Set of Rolling Bearing in Laboratory

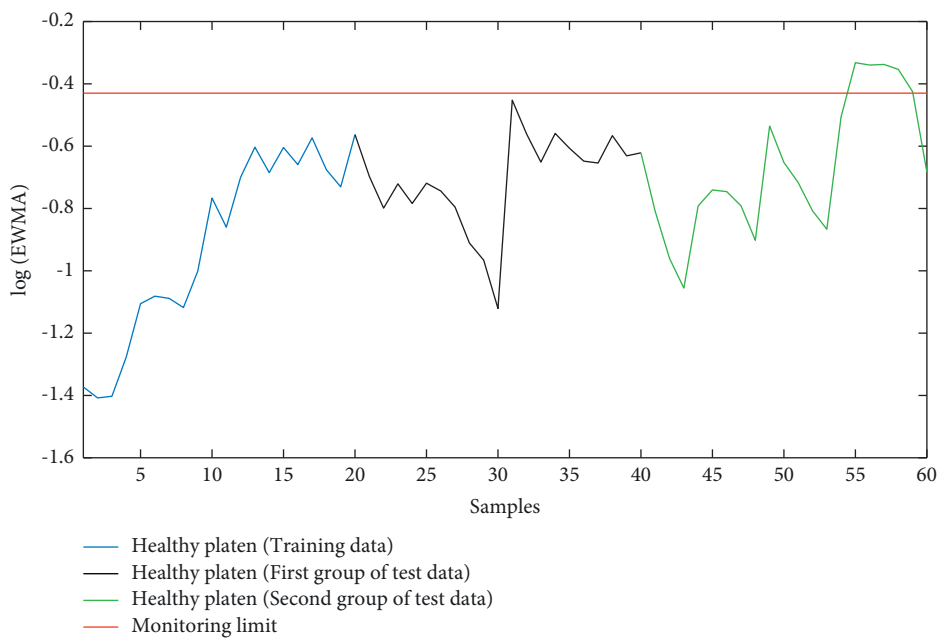
4.2.1. Experimental Design. The mechanical failure test bench used in the laboratory is purchased from Anhui Chaokun Testing Equipment Co., Ltd. The test of bearing is shown in Figure 10.

The data used in the study are shown in Table 5.

4.2.2. Condition Monitoring and Analysis. MEMED-NLOPE is used to monitor the condition of bearing in different status, and the results are shown in

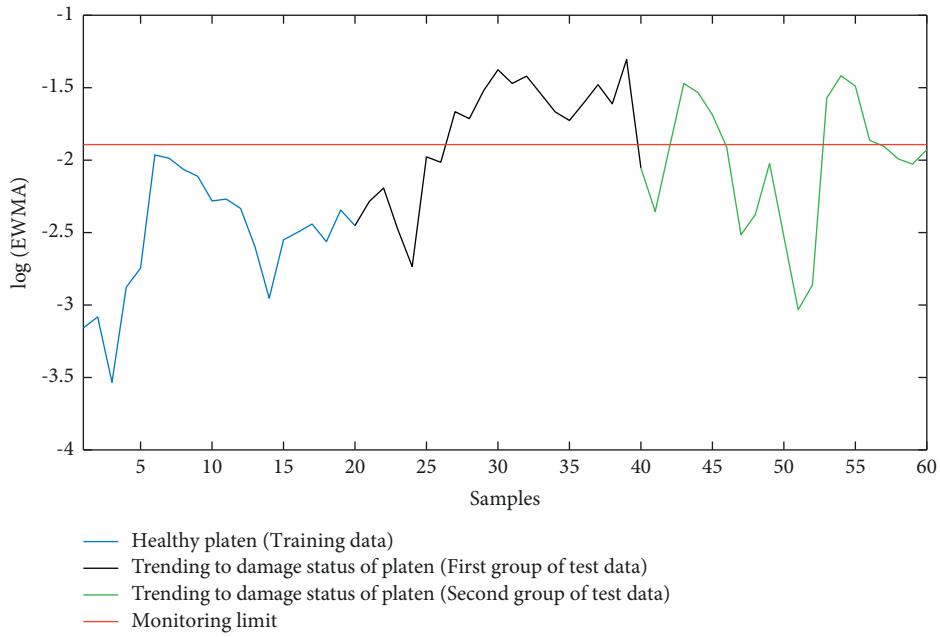


(a)

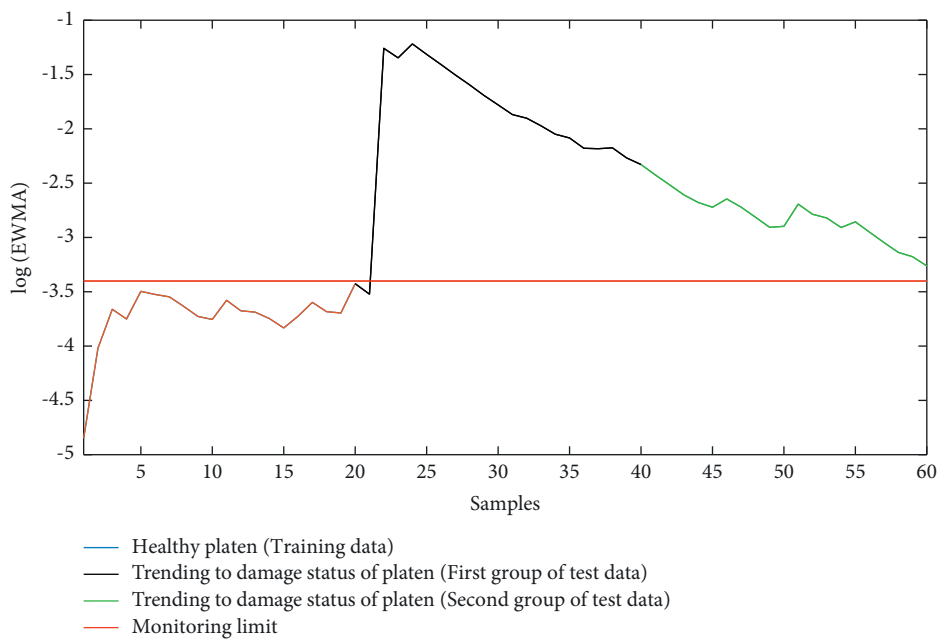


(b)

Figure 9: Continued.

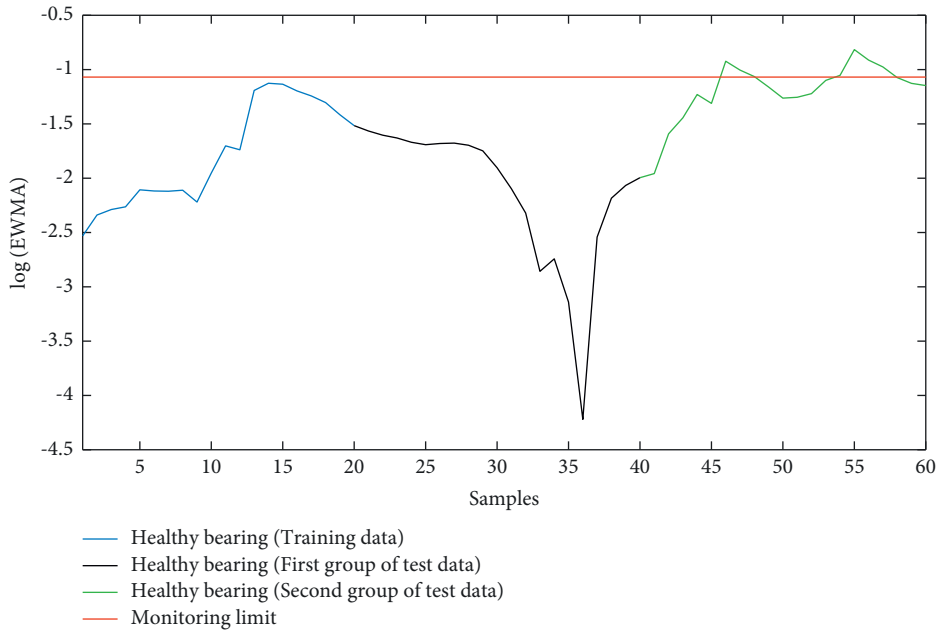


(c)

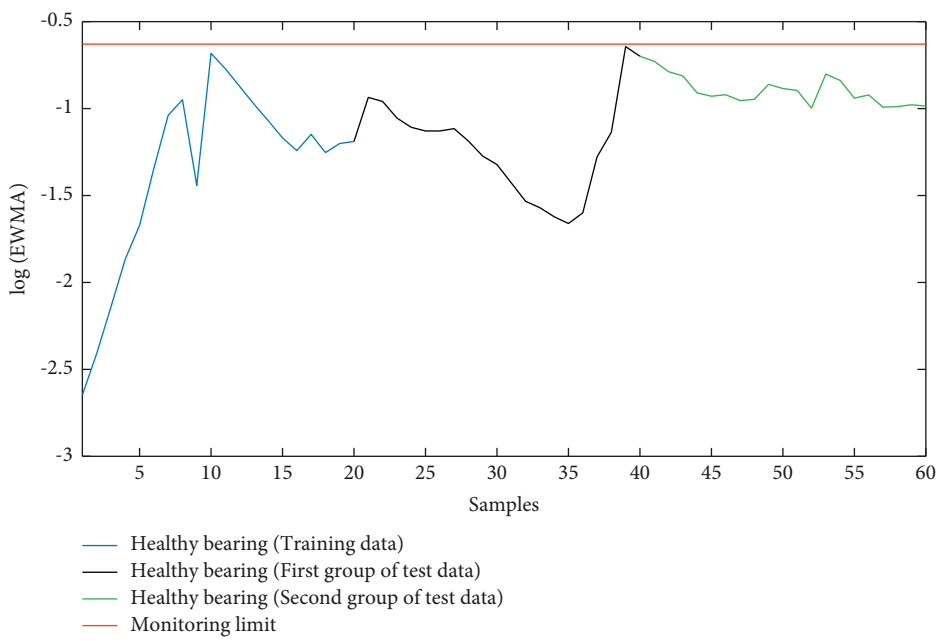


(d)

Figure 9: Continued.

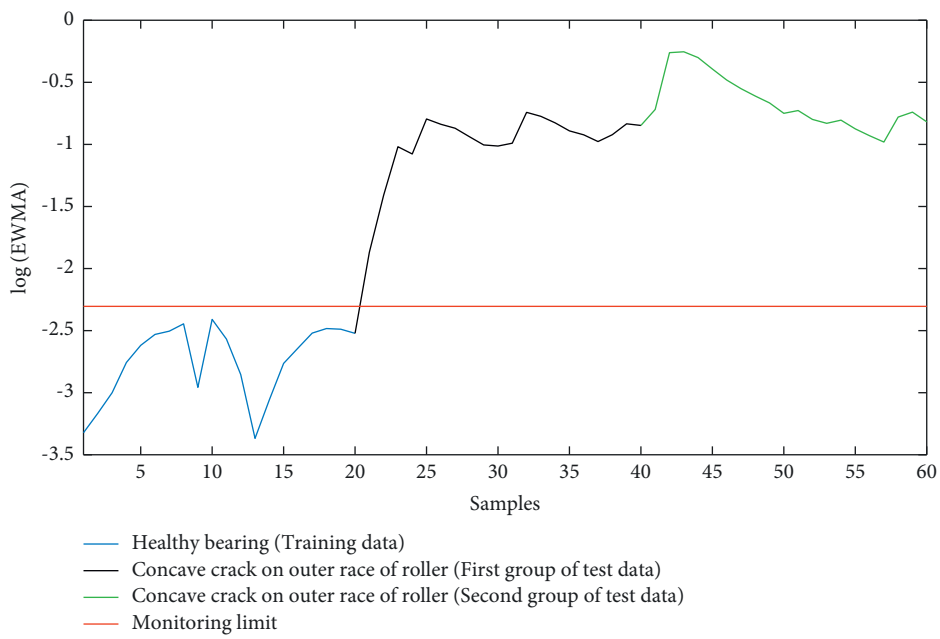


(e)

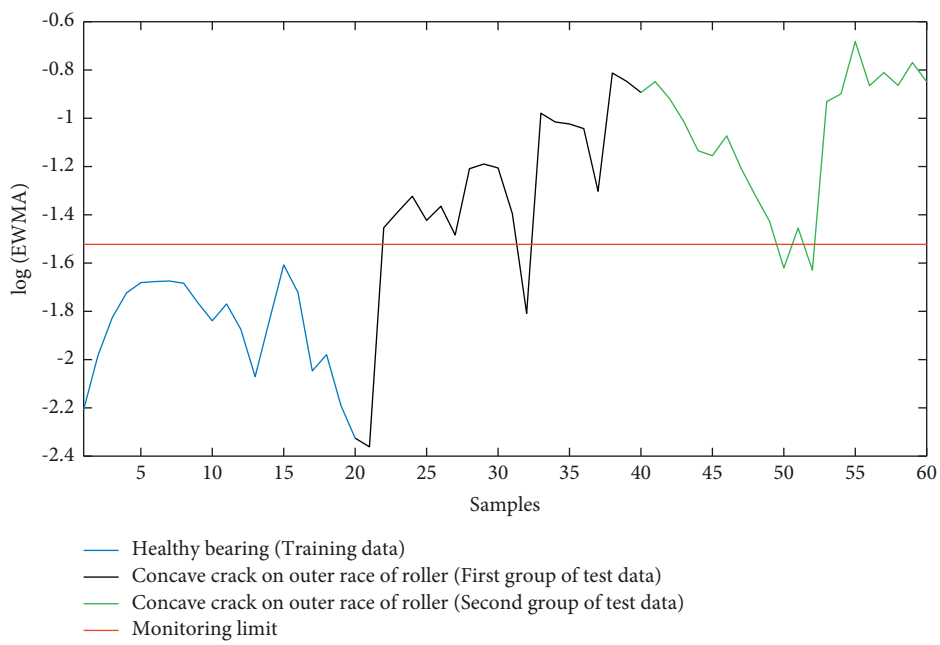


(f)

Figure 9: Continued.



(g)



(h)

Figure 9: Condition monitoring results using MEMD-NLOPE and MEMED-NLOPE. (a) Condition monitoring of healthy platen using MEMD-NLOPE. (b) Condition monitoring of healthy platen using MEMED-NLOPE. (c) Condition monitoring of trending to damage status of platen using MEMD-NLOPE. (d) Condition monitoring of trending to damage status of platen using MEMED-NLOPE. (e) Condition monitoring of healthy bearing using MEMD-NLOPE. (f) Condition monitoring of healthy bearing using MEMED-NLOPE. (g) Condition monitoring of concave crack on outer race of roller using MEMD-NLOPE. (h) Condition monitoring of concave crack on outer race of roller using MEMED-NLOPE.

Figure 11; the corresponding data of different test groups are as follows:

(a) Zc_Data1 (Taining data), Zc_Data2 (First group of test data), and Zc_Data3 (Second group of test data)

(b) Zc_Data1 (Training data), Zc_Data4 (First group of test data), and Zc_Data5 (Second group of test data)

(c) Zc_Data1 (Training data), Zc_Data6 (First group of test data), and Zc_Data7 (Second group of test data)

(d) Zc_Data1 (Training data), Zc_Data8 (First group of test data), and Zc_Data9 (Second group of test data)

(e) Zc_Data1 (Training data), Zc_Data10 (First group of test data), and Zc_Data11 (Second group of test data)

Table 4: Performance evaluation results based on MEMD-NLOPE and MEMED-NLOPE.

Mechanical parts	Status	Cumulative number of tests	False alarm times/ missed detection times based on MEMD-NLOPE	False alarm rate/ missed detection rate based on MEMD-NLOPE (%)	False alarm times/ missed detection times based on MEMED-NLOPE	False alarm rate/ missed detection rate based on MEMED-NLOPE (%)
Platen	Healthy status	40	15	37.50	4	10.00
	Tending to damaged status	40	19	47.50	1	2.50
Roller	Healthy status	40	6	15.00	0	0.00
	Damaged status	40	4	10.00	1	2.50

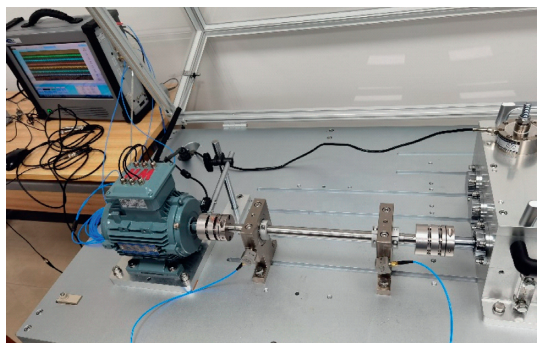
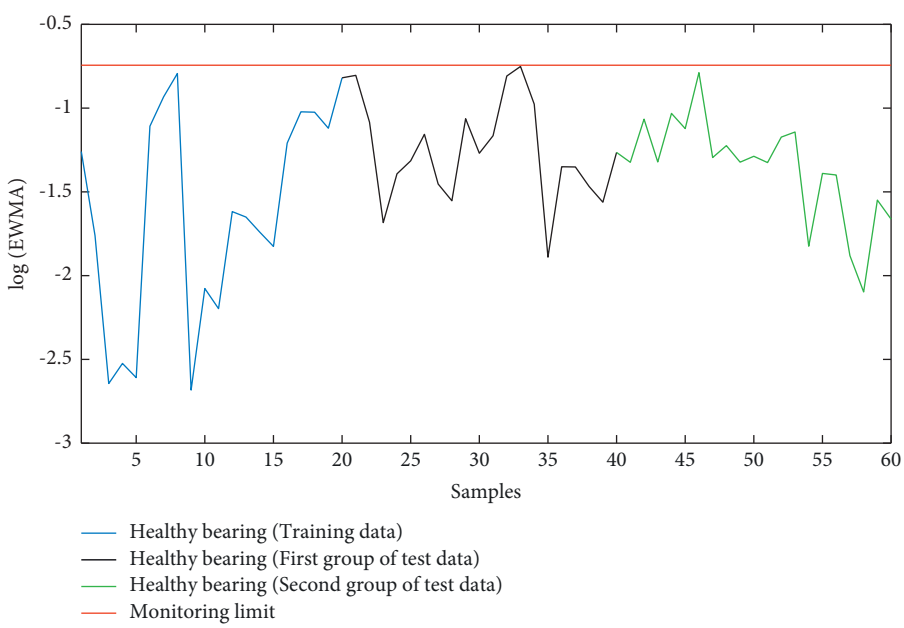


Figure 10: Mechanical failure test bench for bearing.

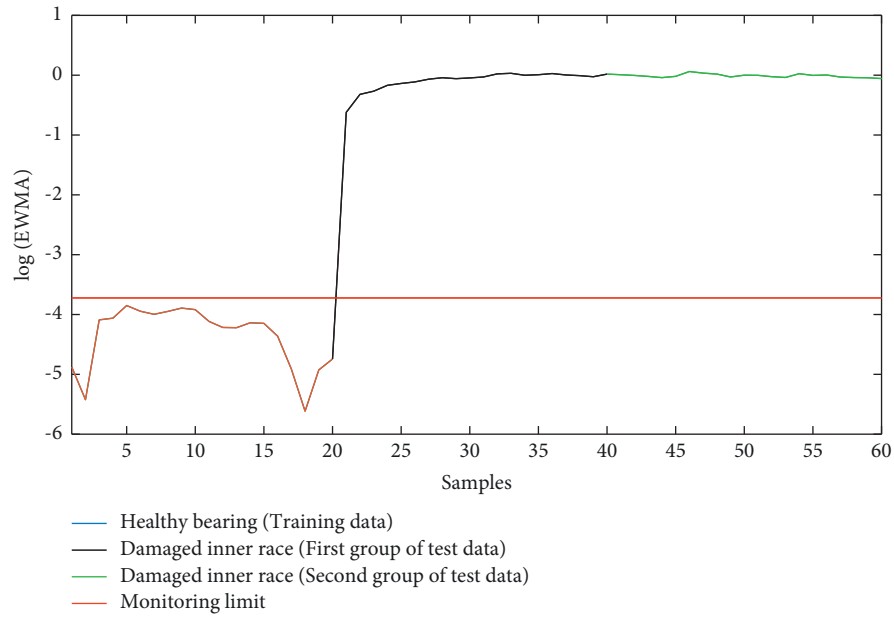
Table 5: Data related to mechanical failure test of bearing.

Experimental group	Rated speed	Fault design of bearing	Vibration signal data	Sampling frequency of vibration signal
Experimental group 1	1000 r/min	Healthy status	Zc_Data1—Zc_Data3	200 ks/S
Experimental group 2	1000 r/min	Inner race damaged status	Zc_Data4—Zc_Data5	200 ks/S
Experimental group 3	1000 r/min	Outer race damaged status	Zc_Data6—Zc_Data7	200 ks/S
Experimental group 4	1000 r/min	Ball damaged status	Zc_Data8—Zc_Data9	200 ks/S
Experimental group 5	1000 r/min	Mixed damaged status	Zc_Data10—Zc_Data11	200 ks/S

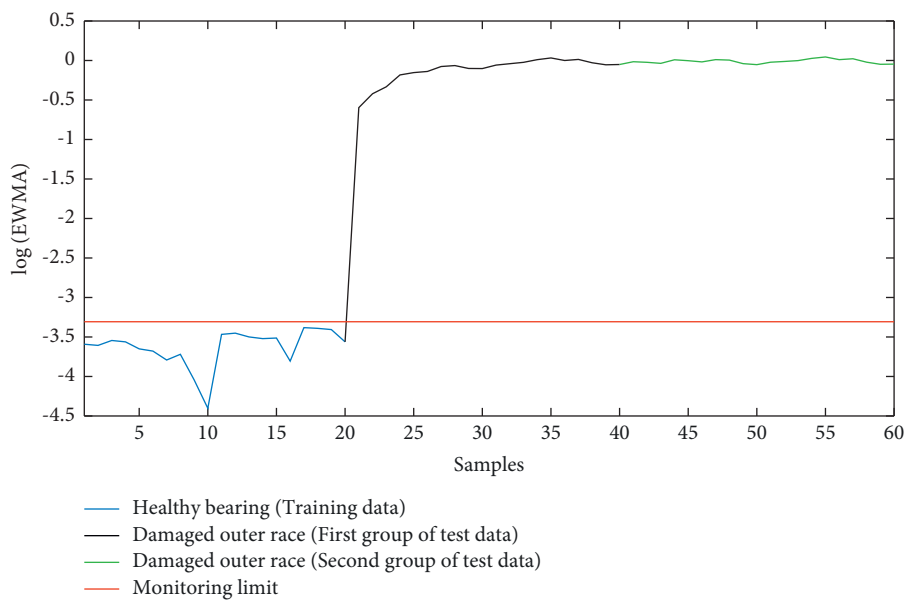


(a)

Figure 11: Continued.



(b)



(c)

Figure 11: Continued.

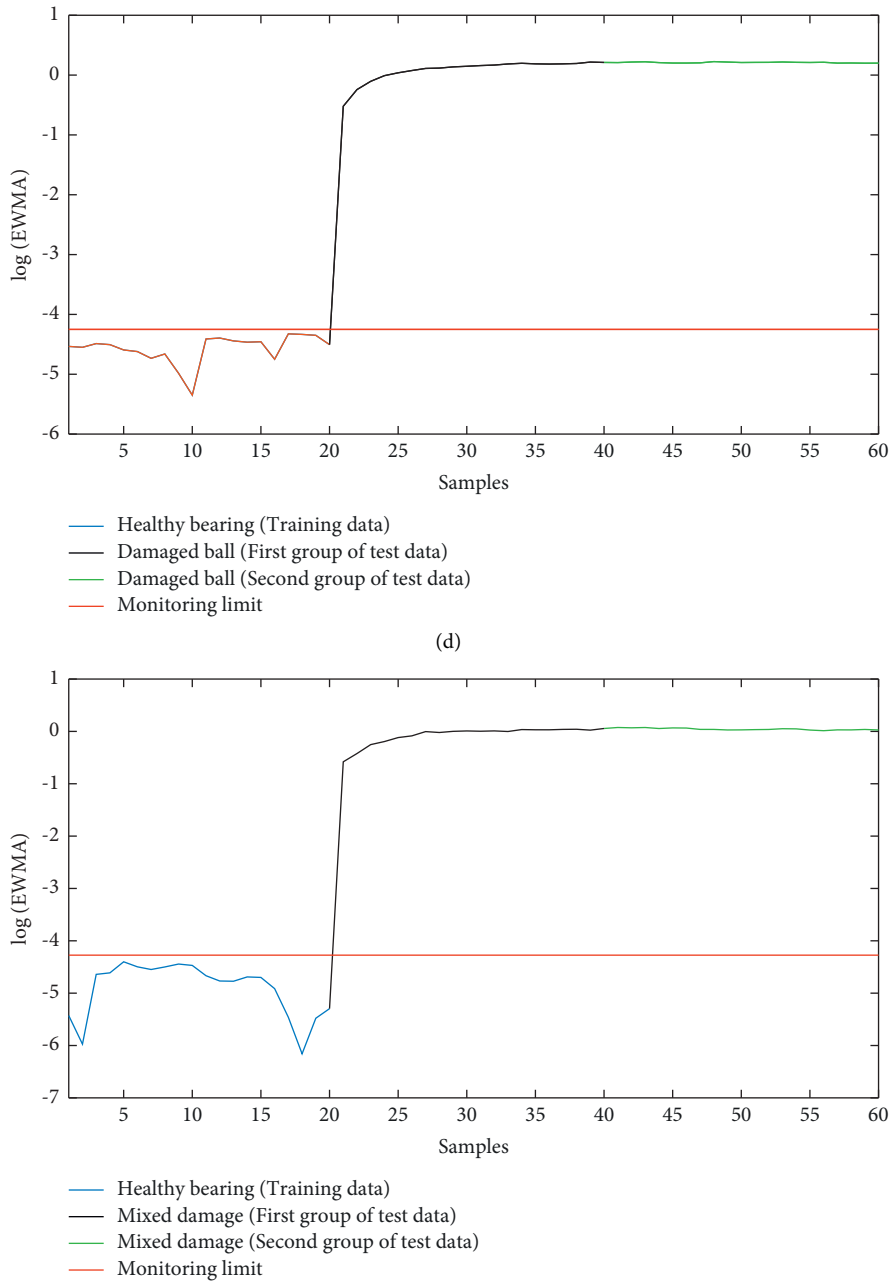


Figure 11: Condition monitoring of different states of bearings. (a) Condition monitoring of healthy bearing. (b) Condition monitoring of damaged inner race. (c) Condition monitoring of damaged outer race. (d) Condition monitoring of damaged ball. (e) Condition monitoring of mixed damage.

Table 6: Performance evaluation results of the fault detection model.

Status	Cumulative number of tests	False alarm times/missed detection times	False alarm rate/missed detection rate (%)
Healthy status	40	0	0.00
Inner damage	40	0	0.00
Outer damage	40	0	0.00
Ball damage	40	0	0.00
Mixed damage	40	0	0.00

4.2.3. Summary. Through the rolling bearing test on the mechanical fault test-bed in the laboratory, the data information of the bearing under different states is obtained, and the normal operation state and abnormal operation state of the bearing are detected and analyzed by using the proposed detection model. The performance evaluation results of MEMED-NLOPE are shown in Table 6.

It can be seen that MEMED-NLOPE can detect the bearing in different states, and its performance is verified.

5. Conclusions

In this paper, a linear dimension and feature reduction method called multiple empirical mode entropy decomposition-nonlocal orthogonal preserving embedding is proposed. In order to reduce the dimension of multivariate signals and consider the correlation between sub signals and source signals, MEMED adopts the maximum entropy method to directly output the subsignal with the maximum correlation. Then, the multiscale feature extraction method reduces the redundancy of feature set by describing the correlation between features. Finally, the automatic fault detection model based on NLOPE-EWMA is proposed to realize condition monitoring. Based on the results of two cases, the performance of condition monitoring based on MEMED-NLOPE is verified, in which the average accuracy of normal and damage detection is higher in comparison with MEMD-NLOPE. For the future work, the massive amounts of data from multiple sensors could be considered for naval gun in health condition monitoring and fault diagnostics.

Data Availability

The experimental data set of naval gun pendulum mechanism data and rolling bearing in laboratory used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by National Nature Science Foundation of China (nos. 61640308 and 61573364) and Nature Science Foundation of Naval University of Engineering (no. 20161579).

References

[1] K. Feng, P. Borghesani, W. A. Smith et al., "Vibration-based updating of wear prediction for spur gears," *Wear*, vol. 426-427, pp. 1410–1415, 2019.

[2] K. Feng, W. A. Smith, P. Borghesani, R. B. Randall, and Z. Peng, "Use of cyclostationary properties of vibration signals to identify gear wear mechanisms and track wear evolution," *Mechanical Systems and Signal Processing*, vol. 150, Article ID 107258, 2020.

[3] S. J. Qin, "Survey on data-driven industrial process monitoring and diagnosis," *Annual Reviews in Control*, vol. 36, no. 2, pp. 220–234, 2012.

[4] Q. Jiang, X. Yan, and B. Huang, "Performance-driven distributed PCA process monitoring based on fault-relevant variable selection and bayesian inference," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 1, pp. 377–386, 2016.

[5] Z. Ge, "Review on data-driven modeling and monitoring for plant-wide industrial processes," *Chemometrics and Intelligent Laboratory Systems*, vol. 171, pp. 16–25, 2017.

[6] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.

[7] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Computation*, vol. 15, no. 6, pp. 1373–1396, 2003.

[8] X. He and P. Niyogi, "Locality preserving projections," *Advances in Neural Information Processing Systems*, vol. 16, pp. 153–160, 2004.

[9] X. He, D. Cai, S. Yan, and H. Zhang, "Neighborhood preserving embedding," in *Proceedings of the 10th IEEE International Conference on Computer Vision (ICCV '05)*, pp. 1208–1213, Beijing, China, October 2005.

[10] J. Yu, "Local and global principal component analysis for process monitoring," *Journal of Process Control*, vol. 22, no. 7, pp. 1358–1373, 2012.

[11] J. Wang, J. Feng, and Z. Y. Han, "Locally preserving PCA method based on manifold learning and its application in fault detection," *Control and Decision*, vol. 22, no. 5, pp. 683–687, 2013.

[12] X. Liu, J. Yin, Z. Feng, J. Dong, and L. Wang, "Orthogonal neighborhood preserving embedding for face recognition," in *Proceedings of the 14th IEEE International Conference on Image Processing, ICIP 2007*, pp. I133–I136, San Antonio, TX, USA, September 2007.

[13] D. Cai, X. He, J. Han, and H.-J. Zhang, "Orthogonal laplacianfaces for face recognition," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3608–3614, 2006.

[14] B. She, F. Tian, W. Liang, and G. Zhang, "Nonlinear model for condition monitoring and fault detection based on nonlocal kernel orthogonal preserving embedding," *Shock and Vibration*, vol. 2018, no. 5, pp. 1–16, 2018.

[15] A. Dorostghol and M. Dorfeshan, "Intelligent fault diagnosis via EMD method," *Journal of Applied Sciences*, vol. 12, no. 18, pp. 1960–1965, 2012.

[16] D. Looney and D. P. Mandic, "Multiscale image fusion using complex extensions of EMD," *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1626–1630, 2009.

[17] N. Rehman and D. P. Mandic, "Multivariate empirical mode decomposition," *Proceedings of the Royal Society A: Mathematical, Physical & Engineering Sciences*, vol. 466, no. 2117, pp. 1291–1302, 2010.

[18] J. Xu, M. Yuan, and Y. Ma, "Feature Selection Using Self-Information and Entropy-Based Uncertainty Measure for Fuzzy Neighborhood Rough Set," *Complex & Intelligent Systems*, pp. 1–19, 2021.

[19] Y. Li, X. Gao, and L. Wang, "Reverse dispersion entropy: a new complexity measure for sensor signal," *Sensors*, vol. 23, no. 19, 2019.

[20] T. Tanaka and D. P. Mandic, "Complex empirical mode decomposition," *IEEE Signal Processing Letters*, vol. 14, no. 2, pp. 101–104, 2007.

[21] G. Rilling, P. Flandrin, P. Goncalves, and J. M. Lilly, "Bivariate empirical mode decomposition," *IEEE Signal Processing Letters*, vol. 14, no. 12, pp. 936–939, 2007.

[22] A. Hemakom, A. Ahrabian, D. Looney, N. u. Rehman, and D. P. Mandic, "Nonuniformly sampled trivariate empirical mode decomposition," in *Proceedings of the 2015 IEEE International Conference on Acoustics*, South Brisbane, QLD, Australia, April 2015.