

RESEARCH ARTICLE

SVMWPE-HLLE Based Fault Diagnosis Approach for Rolling Bearing

CHUNHUI LI¹, YOUFU TANG², AND GUOLIN XU¹¹College of Engineering, Heilongjiang Bayi Agricultural University, Daqing 163319, China²Mechanical Science and Engineering College, Northeast Petroleum University, Daqing 163318, China

Corresponding author: Chunhui Li (li_chh0816@163.com)

This work was supported in part by the National Natural Science Foundation of Heilongjiang, China, under Grant LH2019E073; and in part by the “Three Vertical” Scientific Research Support Program under Grant ZRCPY202207.

ABSTRACT Multi-scale permutation entropy (MPE) is an analytical method describing the complexity of time series, which has been applied to the fault diagnosis of rolling bearings. To solve the problems of MPE in coarse-grained process and permutation entropy calculation, multi-scale weighted permutation entropy based on sliding variance (SVMWPE) was proposed in this paper. By analyzing WGN signal and 1/f noise signal, the parameter selections for SVMWPE were studied, and the stability and superiority were investigated by comparing SVMWPE with MPE, MWPE, and SVMPE. The high-dimensional matrix obtained by MPE feature extraction to pattern recognition was solved by introducing Hessian local linear embedding (HLLE) dimension reduction method, and the feature extraction method based on SVMWPE-HLLE was proposed. The clustering effect was studied by comparing SVMWPE-HLLE with SVMWPE-LLE through the analysis of three simulation signals. Fault diagnosis method for rolling bearing was proposed by combining SVMWPE-HLLE with extreme learning machine (ELM), which was applied to two experimental cases of rolling bearings for analysis. The experimental results showed that the proposed method can realize intelligent diagnosis of different fault types and degrees of rolling bearing, and the fault recognition rate of the proposed method was higher than other methods.


INDEX TERMS SVMWPE, hessian local linear embedding, extreme learning machine, rolling bearings, fault diagnosis.

I. INTRODUCTION

As the basic industry and backbone industry of mechanical industry, the development level of bearing industry often represents or restricts the development level of a country's mechanical industry and other related industries. Rolling bearing is a widely used and strictly required accessory and basic part in the machinery industry. It is a supporting element of the rotating shaft or movable part of various machinery, and also an influencing factor of equipment performance and health status [1], [2].

The diagnosis method of rolling bearing health state is more widely used in the practical application of vibration method. There are two parts to judge the health state of bearing based on vibration method: one is to extract the fault feature information in bearing vibration signal by relying

on signal processing technology, and the other is pattern recognition. Among them, the time domain and frequency domain methods are more suitable for analyzing the vibration signal in ideal environment by analyzing the time domain diagram and spectrum diagram to monitor whether the bearing has a fault. In the actual operation of rolling bearings, non-stationary and nonlinear vibration signals are often generated due to the influence of load changes and speed fluctuations. How to extract feature information from such signals is the key to studying fault diagnosis [3], [4]. The time-frequency analysis method can effectively overcome the shortcomings of the two methods of time domain and frequency domain analysis, and analyze the local characteristics of non-stationary vibration signals more accurately. Commonly used time-frequency analysis methods are short-time Fourier transform (STFT) [5], wavelet analysis [6], empirical mode decomposition (EMD) [7] and so on. In literature [8], STFT and convolutional neural network are combined to diagnose

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

rolling bearings, and end-to-end fault pattern recognition is realized. In literature [9], wavelet analysis is used to denoise the vibration signal of rolling bearing. In literature [10], EMD is used as a preprocessor to extract the energy of each frequency band of the rolling bearing vibration signal as a characteristic parameter, which is used as the input parameter of the neural network to identify the fault type of the rolling bearing. The wavelet analysis or EMD of the signal is to decompose the non-stationary signal into the sum of several simple stationary signals, and then process each component to extract the time-frequency domain information, and then obtain the complete time-frequency information of the original signal. However, due to the factors such as friction, vibration and load in the process of mechanical operation, the vibration signal of mechanical system often shows nonlinear behavior. It is difficult to decompose the signal into stationary signal by time-frequency analysis.

With the development of non-linear analysis methods, especially with Shannon introduced the concept of entropy into the field of information electronic engineering, many non-linear analysis methods have been applied in the field of fault diagnosis, such as approximate entropy (AE) [11], sample entropy (SE) [12], fuzzy entropy (FE) [13], Dispersion entropy (DE) [14] and permutation entropy (PE) [15]. In literature [16], AE was used to measure valve fault complexity. In literature [17], feature information of rolling bearing fault signal was represented by PR component after FE calculation as feature vector. In literature [18], PE was used to predict the whole life of rolling bearings, which can show the early fault detection time of rolling bearings. In literature [19], DE was used for pattern recognition of three states of normal, planetary gear fault and sun gear fault of tank planetary gearbox, and the conclusion of high recognition rate and fast calculation speed is obtained. However, PE, SE, DE and FE can only describe the complexity of signals at a single scale, resulting in incomplete fault information. Therefore, the analysis methods of multi-scale sample entropy (MSE) [20], multi-scale fuzzy entropy (MFE) [21], multi-scale dispersion entropy [22], and multi-scale permutation entropy (MPE) [23] that can describe the complexity of signals at different scales have been proposed. In literature [24], MPE was used to extract the fault features of rolling bearings, and MPE had better performance than PE.

The coarse-grained process of MPE is based on the calculation of the mean value, which can only reflect the average amplitude of the sequence. MPE is lack of description of signal fluctuation and unable to describe the difference between failure degrees of mechanical equipment. In the process of coarse-grained, the information of adjacent points is ignored, which makes the sequence length shorter, the stability lower, and the fault information incomplete. To address these problems, sliding variance method is introduced to the coarse-grained calculation instead of the mean method. Order variation of mode amplitude is considered for the calculation of PE, but not the difference of amplitude of the same serial number mode. The weighted idea is introduced for

PE calculation in the proposed method. Multi-scale weighted permutation entropy based on sliding variance (SVMWPE) is proposed in this paper.

Dimension of the obtained feature matrix is high after MPE feature extraction. To improve the accuracy of pattern recognition, it is necessary to reduce the dimension of the high-dimensional matrix for extracting the main features of low-dimensionality [25]. For non-linear signals, local linear embedding (LLE) is one of the most commonly used dimension reduction methods, which has the advantages of low time complexity and few parameters [26]. In literature [27], the combination of LLE and LDA was applied to the rolling bearing fault diagnosis model, results showed that the dimensionality reduction sample had the best feature separability in the new feature space. When LLE processes data, there are problems such as hole phenomenon, being too sensitive to nearby points and noise, etc. Donoho proposed HLLE on the basis of the LLE algorithm, which uses the Hessian operator to improve the LLE, and replaces the linear representation of the local weights of the LLE with local isometrics to realize the dimensionality reduction process [28].

In this paper, the advantages of SVMWPE in multi-scale information processing are combined with the good adaptability of HLLE in non-linear signal data, and a fault diagnosis method for rolling bearings based on SVMWPE-HLLE and ELM was proposed. The validity and superiority of the method were verified and analyzed by two rolling bearing data.

II. SVMWPE

A. MPE

MFE is a nonlinear analysis method that combines multi-scale analysis with FE to measure the complexity and randomness of time series at different scales [29]. MPE algorithm is as follow.

1) For time series $x_i = \{x_1, x_2, \dots, x_N\}$, the coarse-grained sequence can be expressed by using (1).

$$y_j^\tau = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i, 1 \leq j \leq \frac{N}{\tau} \quad (1)$$

where τ is scale factor, then $\tau = 1, y_j^1$ is original sequence. For the scale factor $\tau \neq 0$, the original sequence is split into N/τ coarse-grained vectors each of length y_j^τ . Taking $\tau = 2$ as an example, the calculation method of coarse-grained sequence is shown in Fig. 1.

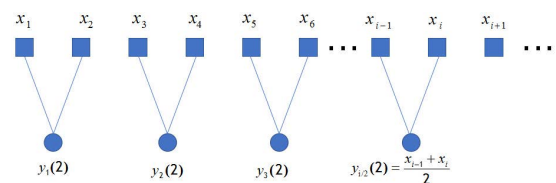


FIGURE 1. Coarse-grained sequence computing segmentation diagram of MPE at $\tau = 2$.

We observed from (1) that the length of the coarse-grained sequence becomes shorter with the increase of the scale factor, which leads to the decrease of stability and the increase of entropy deviation. In Fig.1, the coarse-grained sequence represents the information between x_1 and x_2, x_3 and x_4, \dots, x_{i-1} and x_i , whereas ignoring the information between x_2 and x_3, x_4 and x_5, \dots, x_{i-2} and x_{i-1} , resulting in the incomplete failure information. The coarse-grained computing process is essentially for calculating the mean value between the data, which can reflect the average amplitude of the sequence, ignoring the volatility of the signal.

2) PE of the coarse-grained sequence at each scale is separately calculated to obtain the MPE, which can be expressed by using (2).

$$MPE(x, \tau, m, t) = PE(y_j^\tau, m, t) \quad (2)$$

where m is the embedding dimension, and t is delay time.

B. SVMPE

SVMPE is firstly developed to solve the above problems of MPE, which uses the sliding variance method to improve the coarse-grained process, and its steps are as follows.

1) For time series $x_i = \{x_1, x_2, \dots, x_N\}$, and the coarse-grained sequence can be expressed by using (3).

$$y_j^\tau = \frac{1}{\tau} \sum_{i=j}^{j+\tau-1} (x_i - \bar{x})^2, 1 \leq j \leq N - \tau + 1 \quad (3)$$

When $\tau \neq 0$, the original sequence is split into $N - \tau + 1$ coarse-grained vectors each of length $y_j^\tau(\tau)$. Taking $\tau = 2$ as an example, the calculation method of coarse-grained sequence is shown in Fig. 2.

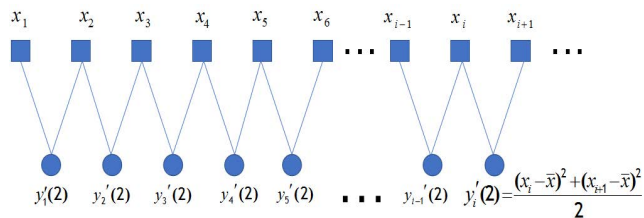


FIGURE 2. Coarse-grained sequence computing segmentation diagram of SVMPE at $\tau = 2$.

We observed from (3) that the length change of the coarse-grained sequence is small, which improves the stability of the entropy value. SVMPE considers the information between the adjacent data in the sequence avoiding information leakage, which can show the volatility of the original sequence.

2) PE at each scale for the new sequence $y_j^\tau(\tau)$ is calculated to obtain SVMPE, which is expressed by using (4).

$$SVMPE(x, \tau, m, t) = PE(y_j^\tau, m, t) \quad (4)$$

C. SVMWPE

The SVMWPE algorithm introduces the weighted idea on the basis of SVMPE, and fully considers the difference of

different amplitudes in the same sequence number ranking mode, and its specific calculation steps are as follows.

1) For a given time series $x_i = \{x_1, x_2, \dots, x_N\}$, the coarse-grained sequence y_j^τ is calculated by (3).

2) Phase space reconstruction of coarse-grained sequence y_j^τ is performed by adding a weighting coefficient to each reconstruction component, which can be expressed by (5).

$$w_r(s) = \frac{1}{m} \sum_{k=1}^m (y_{j+(k-1)t} - \bar{y}_j)^2 \quad (5)$$

Of which:

$$\bar{y}_j = \frac{1}{m} \sum_{k=1}^m y_{j+(k-1)t} \quad (6)$$

where s represents the number of identical symbol sequences.

3) The probability of different symbols by using (7).

$$p_w(\pi_r) = \frac{f_w(\pi_r)}{\sum_{i=1}^{m!} f_w(\pi_r)} \quad (7)$$

Of which:

$$f_w(\pi_r) = \sum_{i=1}^s f(\pi_r(i))w_r(i) \quad (8)$$

where π_r indicates the possible sorting patterns, $f_w(\pi_r)$ represents the frequency of the r -th order.

4) WPE is obtained according to the concept of Shannon entropy, which can be expressed by (9).

$$H_w = - \sum_{i=1}^{m!} p_w(\pi_l) \ln p_w(\pi_l) \quad (9)$$

5) The SVMWPE representation of the original time series, which is expressed by using (10).

$$SVMWPE(x, \tau, m, t) = \frac{1}{\tau} \sum_{k=1}^{\tau} H_w^k(y_j^\tau, m, t) \quad (10)$$

D. ANALYSIS OF PARAMETER SELECTION

Four parameters need to be set in SVMWPE, including scale factor τ , embedding dimension m , time delay t , and time series length N . For the scale factor τ , in order to ensure that the length of the coarse-grained sequence cannot be too short, and τ cannot be too large. We used $\tau_{\max} \geq 10$ and $\tau_{\max} = 20$ in this study. The permutation entropy is highly dependent on the embedding dimension m . The larger the value of m , the more information the sequence needs to be reconstructed, and the subtle changes in the sequence are ignored [30]. If the selection of m is too small, the dimension of the reconstructed signal will become lower, the contained states will be reduced, and the dynamic changes of the sequence will be difficult to detect. Bandt recommended that the embedding dimension m should be selected between 3 and 7 [15]. The research on permutation entropy

by many scholars showed that the selection of time delay t had almost no effect on the results [31]. Hence, t is set to 1 in this paper. For the time series length N , whether the length is too long or too short will have a greater impact on the entropy value.

MPE, MWPE, SVMPE and SVMWPE of WGN and $1/f$ noise at $m = 4 - 7$ were calculated by selecting $t = 1$ and $N = 3000$, the results are shown in Fig. 3 and Fig. 4. We observed that when $m = 4$ or 5 , most of the entropy values of the four methods hovered within 0.9-1, and the variation range was small. This shows that when m was small, the arrangement was less, resulting in a lower frequency of the same symbolic pattern. The probability of the corresponding state vector after weighting was higher, and the entropy value was larger. When $m = 6$ or 7 , the entropy obtained by the four methods varied greatly, which can more accurately reflect the variation range of entropy with scale factor. When $m = 7$, the four entropy curves had high overlap and were difficult to distinguish. Based on these analyses, therefore m was set to 6. In addition, compared with the other three methods, SVMWPE had smaller fluctuation in entropy value

and smoother curve, which reflected that SVMWPE had good stability and superiority.

The parameters $m = 6$ and $t = 1$ were selected to investigate the influence of the selection of N on SVMWPE. The SVMWPE of WGN and $1/f$ noise for $N = 1000, 1500, 2000, 2500, 3000, 3500, 4000$ are calculated. The corresponding entropy values were recorded as EV1-EV7. The results are shown in Fig. 5 and Fig. 6, and the differences at each scale are shown in Fig. 7 and Fig. 8. Fig. 5 and Fig. 6 show that the overall trend of each entropy curve was roughly the same, with the increase of N , the entropy corresponding to each scale increases. Combining Fig. 7 and Fig. 8, we observed that when $N \leq 3000$, the entropy value difference between the curves was large, when $N \geq 3000$, the fluctuation of each curve was smaller, the entropy value difference was very close, and the stability was higher. Therefore, $N \geq 3000$ was selected in the next stage of research.

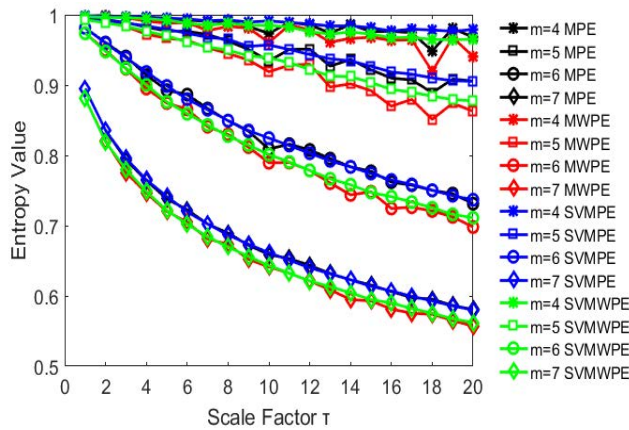


FIGURE 3. The entropy values of WGN under different embedding dimensions of MPE, MWPE, SVMPE and SVMWPE algorithms.

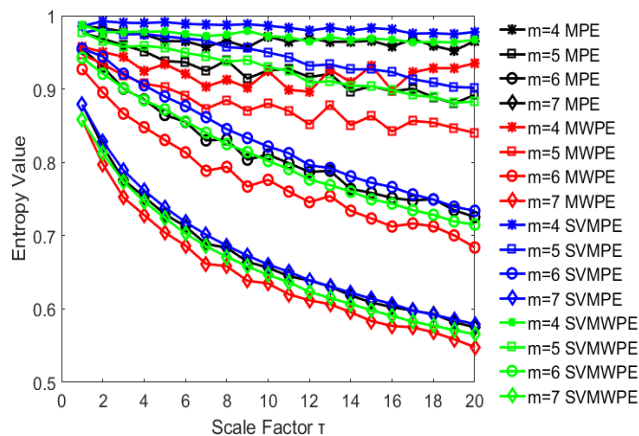


FIGURE 4. The entropy values of $1/f$ noise under different embedding dimensions of MPE, MWPE, SVMPE and SVMWPE algorithms.

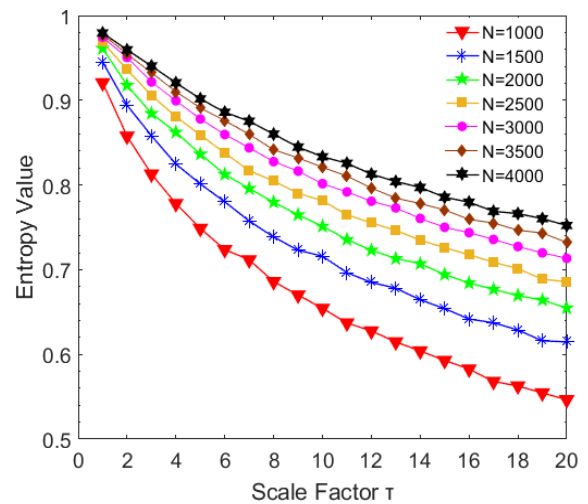


FIGURE 5. The entropy values of WGN under different time series lengths of SVMWPE algorithms.

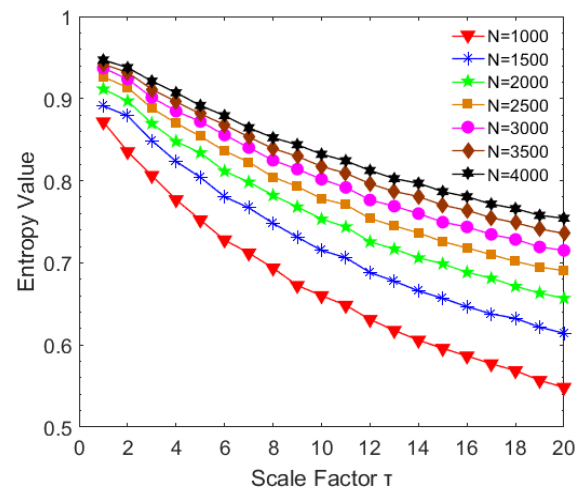


FIGURE 6. The entropy values of $1/f$ noise under different time series lengths of SVMWPE algorithms.

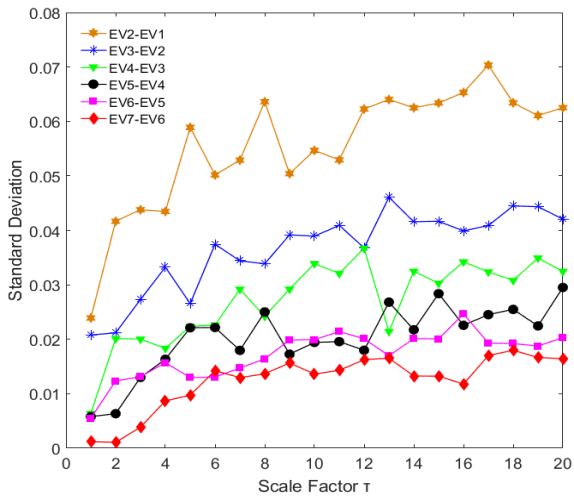


FIGURE 7. Difference of SVMWPE of WGN with different lengths under different scale Factors.

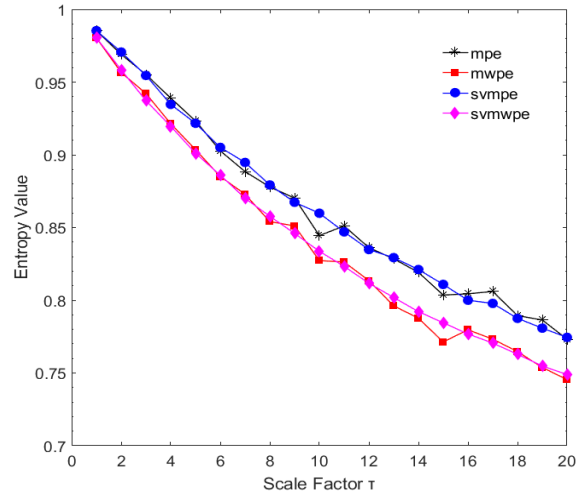


FIGURE 9. MPE, MWPE, SVMPE, and SVMWPE of WGN under the same parameters.

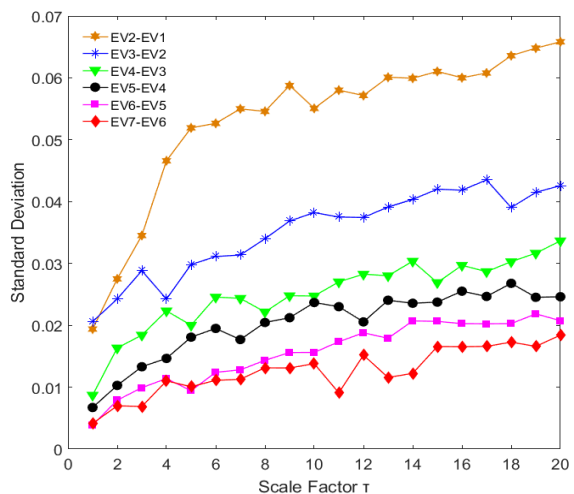


FIGURE 8. Difference of SVMWPE of 1/f noise with different lengths under different scale factors.

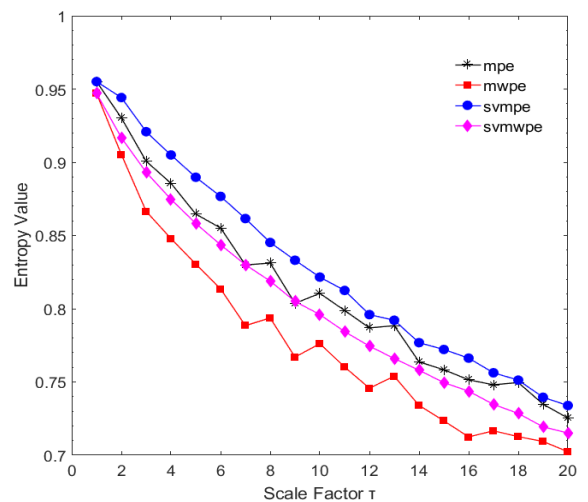


FIGURE 10. MPE, MWPE, SVMPE, and SVMWPE of 1/f noise under the same parameters.

E. ANALYSIS OF PARAMETERS IN SVMWPE

Parameters of SVMWPE based on the previous section were analyzed to illustrate the superiority of SVMWPE with $m = 6$, $t = 1$ and $N = 4000$. MPE, MWPE, SVMPE, and SVMWPE of WGN and 1/f noise was calculated, and results are shown in Fig. 9 and Fig. 10. The entropy curves of SVMPE and SVMWPE were significantly smoother than those of MPE and MWPE, indicating that the MPE had been improved by sliding variance method, which can retain more useful information in processing signals and improve the insufficient calculation of original coarse graining. As the scale factor increases, the entropy fluctuation of SVMWPE was much smaller than that of SVMPE, especially in the high-scale segment, indicating that SVMWPE was superior and more stable.

III. SVMWPE-HLLE BASED FEATURE EXTRACTION

The research on SVMWPE in the previous section found that the dimension of the feature matrix obtained after feature

extraction is too high, which is not conducive to pattern recognition, and it is necessary to reduce the dimension of the high-dimensional feature matrix. In this section, we combine SVMWPE with HLLE to propose a feature extraction method based on SVMWPE-HLLE.

A. LLE

LLE is a non-linear dimensionality reduction method proposed by Roweis and Saul, which has the advantages of low time complexity and few parameters (only one predetermined parameter). The core idea of LLE is to ensure that the manifold structure remains unchanged by keeping the weight values unchanged during dimension reduction [32].

For high-dimensional data set $X = \{x_i \in R^D, i = 1, 2, \dots, N\}$, where D is the sample dimension, and N is the sample number, the mapping of low-dimensional space

$Y = \{y_i \in R^d, i = 1, 2, \dots, N\}$, where d is the dimension of low-dimensional space ($d \ll D$).

The LLE algorithm is as follows.

1) Find k neighbor points for each sample by using Euclidean distance between sample points, and get neighborhood $V_i = \{x_{i1}, x_{i2}, \dots, x_{ik}\}$.

2) The local weight W is calculated based on the minimum principle of local linear structural error, and the reconstruction error function by using (11).

$$\varepsilon(w) = \sum_{i=1}^N \left\| x_i - \sum_{j=1}^N w_{ij} x_{ij} \right\|_2^2 \quad (11)$$

Equation (11) is requires two constraints: (1) $\sum_{j=1}^k w_{ij} = 1$;

(2) Sample x_j is not in the neighborhood of x_i , and $w_{ij} = 0$. Take the minimum value of $\varepsilon(w)$ under condition (1) to obtain the best weight vector $w_i = (w_{i1}, w_{i2}, \dots, w_{ik})^T$. Then the weight matrix is $W = (w_1, w_2, \dots, w_N)^T$.

3) To obtain low-dimensional coordinates, the reconstruction error function Y is constructed under the condition that the weight matrix W is constant by using (12).

$$\varepsilon(Y) = \sum_{i=1}^N \left\| y_i - \sum_{j=1}^N w_{ij} y_{ij} \right\|_2^2 \quad (12)$$

Equation (12) has two restrictions: (1) $\sum_{j=1}^N y_i = 0$,

(2) $\frac{1}{N} \sum_{j=1}^N y_i y_i^T = I$. Equation(12) is transformed into eigen-

value decomposition of the matrix $M = (1-W)^T(1-W)$, and the eigenvector corresponding to the smallest eigenvalue is obtained. The eigenvector $P = (p_2, p_3, \dots, p_{d+1})$ corresponding to the first 2 to $d + 1$ eigenvalues is taken, and the final low-dimensional matrix $Y = P^T$ is obtained.

B. HLLE

HLLE is proposed on the basis of LLE, which uses the local Hessian matrix to decompose the curvature of the sample data, and its steps are as follows.

1) Find k neighbor points for each sample by using Euclidean distance between sample points, and get neighborhood $V_i = \{x_{i1}, x_{i2}, \dots, x_{ik}\}$.

2) Obtain the centralization matrix $\tilde{V}_i = V_i - \frac{V_i}{k} e_k^T$ for each neighborhood V_i , and then obtain the singular value U_i for V_i . The first d column vectors of U_i are the matrix tangent space coordinates, and denoted as $Z_i = [z_1, z_2, \dots, z_d]^T$.

3) For each neighborhood set to $M_i = [e, Z_i, Z_i^{k \times s} \cdot Z_i^{k \times l}]$, there are $(1 + d + d(d + 1))/2$ columns. The first $d + 1$ is composed of a column vector with all 1 and Z_i , and $Z_i^{k \times s} \cdot Z_i^{k \times l}$ is the s -th column and the l -th column Hessian product of columns. Perform an orthogonal transformation on M_i to obtain a column orthogonal matrix \tilde{M}_i , then the Hessian matrix is $H_i = \tilde{M}_i(:, d + 2 : 1 + d + d(d + 1)/2)^T$, where $H_i \in R^{k(d(d+1)/2)}$.

4) Using H_i of each neighborhood to construct a symmetric matrix H , whose element is $H_{ij} = \sum_{s=1}^N \sum_{l=1}^{d(d+1)/2} (H^s)_{l,i} (H^s)_{l,j}$.

5) Calculate the $d + 1$ minimum eigenvalues of H , and obtain the corresponding eigenvector as u_1, u_2, \dots, u_d , then $U = [u_2, u_2, \dots, u_{d+1}]$ is the required null space.

6) Denote the matrix $R_{ij} = \sum_{l \in J_l} U_{l,i} U_{l,j}$, where $i, j = 1, 2, \dots, d, J_l$ is the neighborhoods of specific samples. $U_{l,r}$ is the element value of the l -th row and the r -th column of U corresponding to H , then $T = R^{-1/2} U^T$ is a low-dimensional embedding.

C. FEATURE EXTRACTION BASED ON SVMPE-HLLE

SVMPE-HLLE feature extraction method are as follows.

1) For time series $x_i = \{x_1, x_2, \dots, x_N\}$, the parameters i.e. scale factor t , embedding dimension m , time delay t , and time series length N need to be used.

2) Calculate SVMWPE for S samples, and construct a high-dimensional sample matrix V of $S \times \tau$.

$$V = \begin{bmatrix} H_{11} & H_{12} & \dots & H_{1\tau} \\ H_{21} & H_{22} & \dots & H_{2\tau} \\ \vdots & \vdots & \dots & \vdots \\ H_{S1} & H_{S2} & \dots & H_{S\tau} \end{bmatrix} \quad (13)$$

3) Reduce the dimension of the high-dimensional matrix V through HLLE to obtain the low-dimensional feature vector $f = [f_1, f_2, \dots, f_d]$.

D. SIMULATION EXPERIMENT

For the fault signal of rotating machinery, which is non-stationary and non-linear, two frequency modulation and amplitude modulation signals $x_1(t)$ and $x_2(t)$ were constructed according to the vibration signal to study the SVMPE-HLLE feature extraction method, $x_1(t)$ is the frequency modulation and amplitude modulation signal with fundamental frequency of 30Hz and frequency modulation of 10Hz, $x_1(t) = (1+0.5\sin(10\pi t)) \times \cos(60\pi t+2\cos(20\pi t))$, $x_2(t)$ is the frequency modulation and amplitude modulation signal with fundamental frequency of 25 Hz and frequency modulation of 8Hz, $x_2(t) = (1+0.5\sin(8\pi t)) \times \cos(50\pi t+2\cos(20\pi t))$. Time $t \in [0, 1]$, sampling frequency is 1000Hz.

Random generation 3×2 mixing matrix E is used to simulate three signals, and three WGN signals $WGN_1(t)$, $WGN_2(t)$ and $WGN_3(t)$ are added to construct new signals $a(t)$, $b(t)$, and $c(t)$. The time-domain waveform is shown in Fig.11.

$$\begin{cases} a(t) = E(1, 1) \times x_1(t) + E(1, 2) \times x_2(t) + WGN_1(t) \\ b(t) = E(2, 1) \times x_1(t) + E(2, 2) \times x_2(t) + WGN_2(t) \\ c(t) = E(3, 1) \times x_1(t) + E(3, 2) \times x_2(t) + WGN_3(t) \end{cases} \quad (14)$$

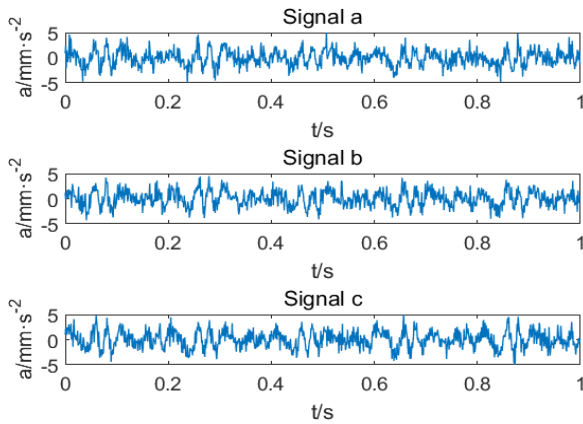


FIGURE 11. Time domain waveforms of three analog signals.

A total of 180 samples are taken 60 samples from each of the three signals. Time series length of each sample was $N = 4000$, the maximum scale factor $\tau_{max} = 20$, embedding dimension $m = 6$, and time delay $t = 1$. Adjacent points k is 12 and the low-dimensional space dimension d is 3 in HLLE. All samples were calculated by SVMWPE, and the high-dimensional feature matrix V of 180×20 was constructed. Then, the dimension of V was reduced by HLLE to obtain a low-dimensional feature matrix of 180×3 . The results after dimension reduction are shown in Fig. 12. We observed that the three analog signals can be completely separated by SVMWPE-HLLE. The three-dimensional features of different signals were relatively concentrated, and the clustering effect was better. In order to highlight the superiority of HLLE, LLE was used to reduce the dimension of the high-dimensional feature matrix constructed by SVMWPE calculation. The parameter selection was the same as that of HLLE, the result after dimension reduction is shown in Fig. 13. SVMWPE-LLE cannot completely separate the three analog signals, and the three-dimensional features were scattered, indicating that the SVMWPE-HLLE feature extraction method can accurately identify the three analog signals and the clustering effect was good.

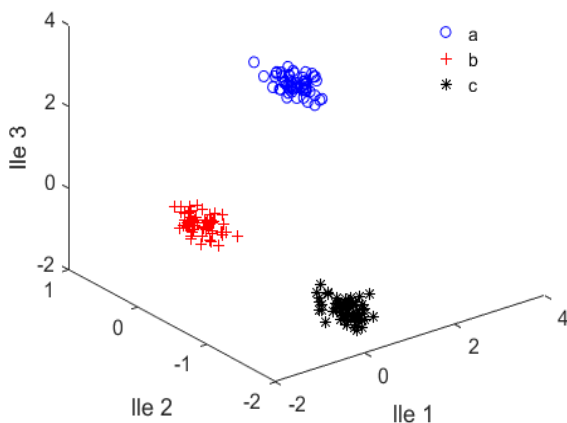


FIGURE 12. Dimensionality reduction results of SVMWPE-HLLE.

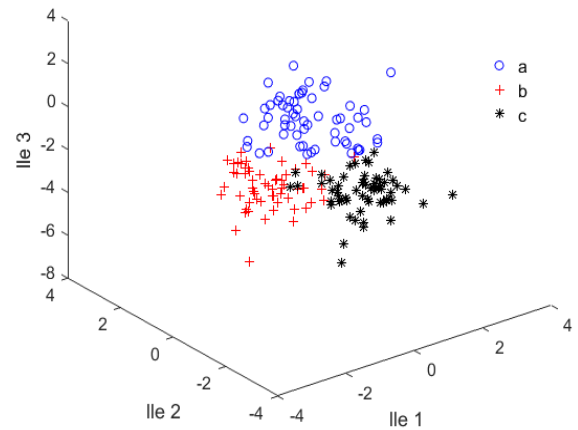


FIGURE 13. Dimensionality reduction results of SVMWPE-LLE.

IV. SVMWPE-HLLE BASED FAULT DIAGNOSIS APPROACH FOR ROLLING BEARING

A. ELM

ELM is proposed based on single hidden layer feedforward neural network, which has simple algorithm and fast training speed [33]. The network structure of ELM is shown in Fig. 14.

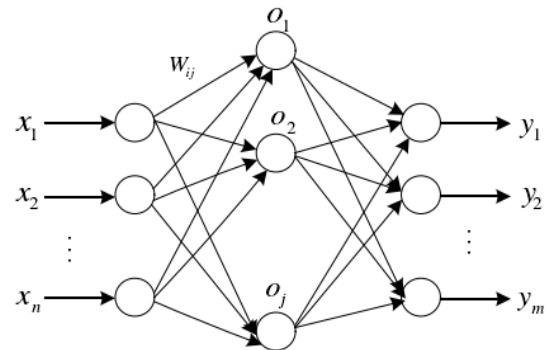


FIGURE 14. Network structure of ELM.

For a given training set (x_i, t_i) , the mathematical function of the above structure graph can be expressed by (15).

$$\sum_{i=1}^M \beta_i g(w_i \cdot x_i + b_i) = \sigma_j, \quad j = 1, 2, \dots, N \quad (15)$$

where M is the number of hidden layer nodes, β_i is the output weight, w_i is the input weight, b_i is the unit bias, and σ_i is the classification result.

B. FAULT DIAGNOSIS PROCESS

Fault diagnosis method for rolling bearings based on SVMWPE-HLLE and ELM is proposed through the analysis of the previous chapters. The diagnosis steps are as follows.

1) Set k types of rolling bearing fault states, collect n samples for each state, calculate the SVMWPE at τ scales for all samples, and construct a high-dimensional sample matrix $V \in R^{kn \times \tau}$.

2) The dimension of the high-dimensional matrix V is reduced by HLLE, and the low-dimensional matrix $P \in R^{kn \times d}$ with dimension d is obtained as the feature vector of the sample.

3) Randomly selected i training samples from n low dimensional matrices of each state as classifiers to form a training set of $ki \times d$ and the remaining $n-i$ as test samples to form a test set of $k(n-i) \times d$.

4) The feature vector of each state in the training set is input into ELM to train, and the ELM training model is obtained. Then the feature vector of the test set is input into ELM to test.

5) According to the output of the ELM classifier, the fault type and severity of the rolling bearing can be judged.

C. EXPERIMENT ANALYSIS OF ROLLING BEARING FAULT DIAGNOSIS

1) EXPERIMENTAL CASE 1

The experimental data of Western Reserve University were used to verify the applicability and superiority of the proposed method in rolling bearing fault diagnosis [34]. The rolling bearing test rig system and its schematic diagram are shown in Fig. 15 and Fig. 16. The drive end bearing selected for the test bearing was 6205-2RS JEM SKF deep groove ball bearing. Bearing Damage was single point damage in electric discharge machining. Under the experimental conditions of motor load was 1471W, rotational speed was 1750 r/min

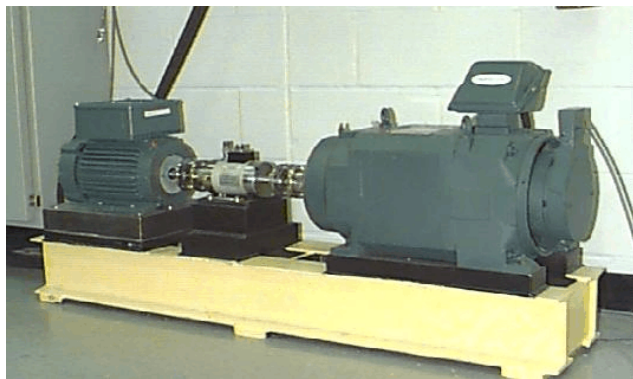


FIGURE 15. Rolling bearing test bench.

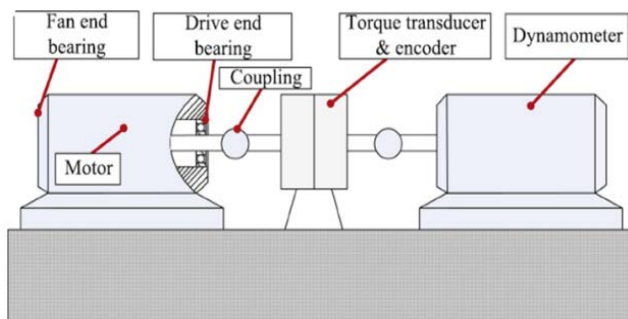


FIGURE 16. Rolling bearing test bench system sketch.

and sampling frequency was 12kHz. It was set into seven states: Normal (Norm), Inner Race (IR1, IR2), Ball Element (BE1, BE2) and Outer Race (OR1, OR2). Each state takes 20 samples with a data length of 4096 points, 5 samples were randomly selected as training data, and the remaining 15 samples were used as test data. The detailed description of the experimental data of the rolling bearing is shown in Table 1. The time-domain waveforms of the seven states are shown in Fig. 17.

TABLE 1. Detailed description of experimental data of rolling bearings.

Fault Location	Fault Diameter(mm)	Train set	Test set	Number
Inner Race(IR1)	0.1778	5	15	1
Inner Race(IR2)	0.5334	5	15	2
Ball element (BE1)	0.1778	5	15	3
Ball element (BE2)	0.5334	5	15	4
Outer Race (OR1)	0.1778	5	15	5
Outer Race (OR2)	0.5334	5	15	6
Normal(Norm)	0	5	15	7

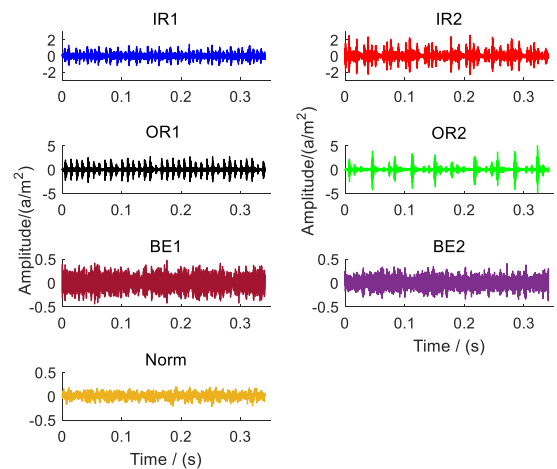


FIGURE 17. Time domain waveforms of seven states of rolling bearing.

MPE, MWPE, SVMPE, and SVMWPE at 20 scales were calculated for all samples by selecting parameters $t = 1$ and $m = 6$. The results of MPE, MWPE, SVMPE and SVMWPE in seven states were shown in Fig. 18-21. We observed that the entropy values of MPE and MWPE were obviously more volatile compared with SVMPE and SVMWPE, especially for OR1 and OR2, indicating that SVMPE and SVMWPE had better stability than MPE and MWPE. It is difficult to judge which method was superior observing the entropy curve of SVMWPE and SVMPE.

MPE, MWPE, SVMPE, and SVMWPE of all samples were constructed by 140×20 high-dimensional feature matrix, and the low-dimensional feature matrix was obtained by HLLE. For HLLE, we set $k = 12$ and $d = 3$. The results after dimension reduction are shown in Fig. 22-25.

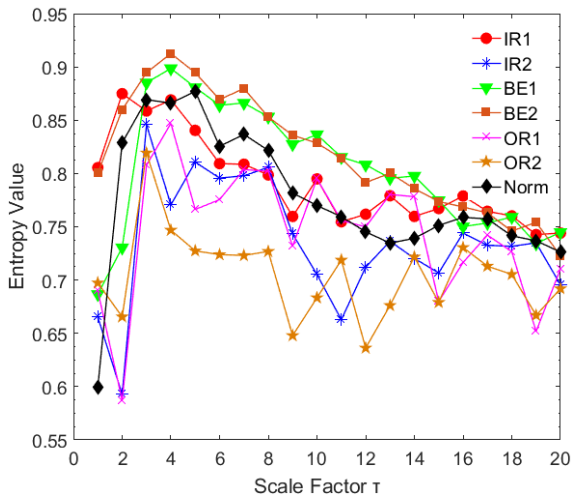


FIGURE 18. MPE of vibration signal of rolling bearing in different states.

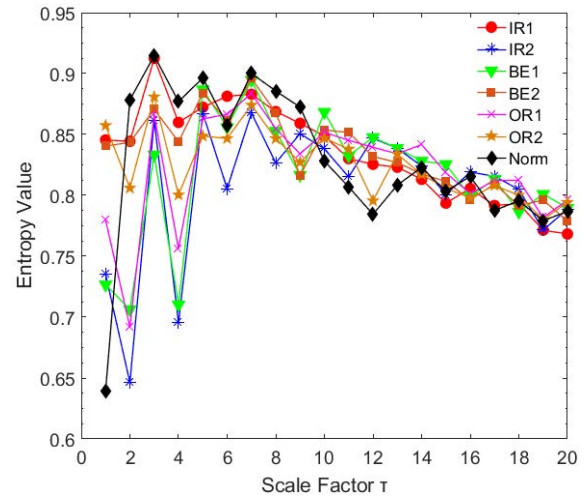


FIGURE 20. SVMPE of vibration signal of rolling bearing in different states.

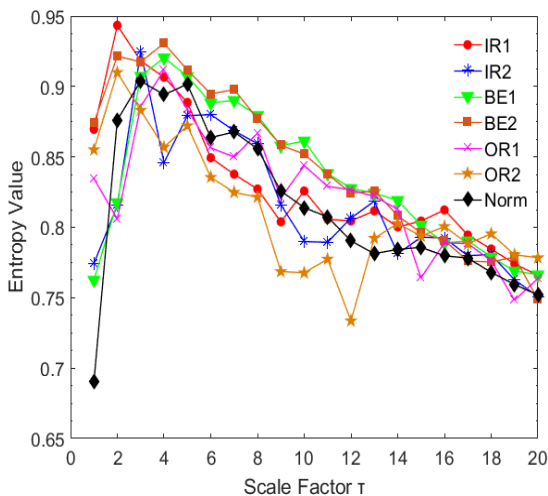


FIGURE 19. MVPE of vibration signal of rolling bearing in different states.

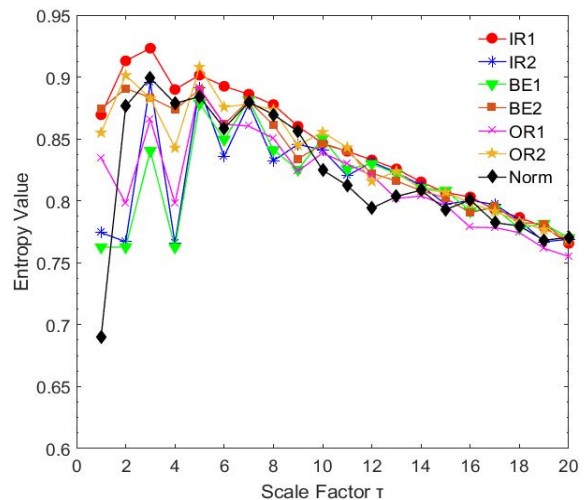


FIGURE 21. SVMVPE of vibration signal of rolling bearing in different states.

We observed that MPE-HLLE, MWPE-HLLE, and SVMPE-HLLE cannot completely separate all seven states, and there were cases of crossover. Whereas SVMWPE-HLLE can completely separate the seven states, and the three-dimensional features of different states of SVMWPE-HLLE were more concentrated than that those of MPE-HLLE, MWPE-HLLE, and SVMPE-HLLE, indicating that the SVMWPE-HLLE feature extraction method can accurately identify seven states of rolling bearings, which is superior to the other three methods. In order to illustrate the superiority of HLLE dimensionality reduction method. SVMWPE of all samples was constructed high-dimensional feature matrix, and the low-dimensional feature matrix was obtained by LLE method. The result of dimensionality reduction is shown in Fig. 26. We observed that whether the separation effect or the concentration of three-dimensional features, SVMWPE-HLLE had better clustering effect than SVMWPE-LLE, and can be used for the evaluation of the degree of fault defects of rolling bearings.

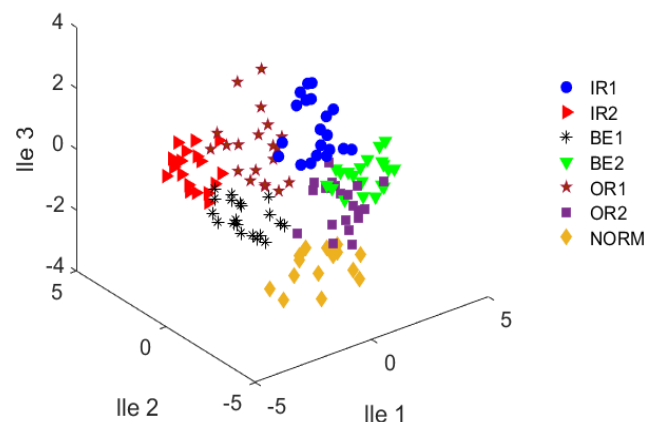


FIGURE 22. Clustering results based on MPE-HLLE.

In order to evaluate the fault state of the bearing more accurately, ELM was introduced to identify seven states. Firstly,

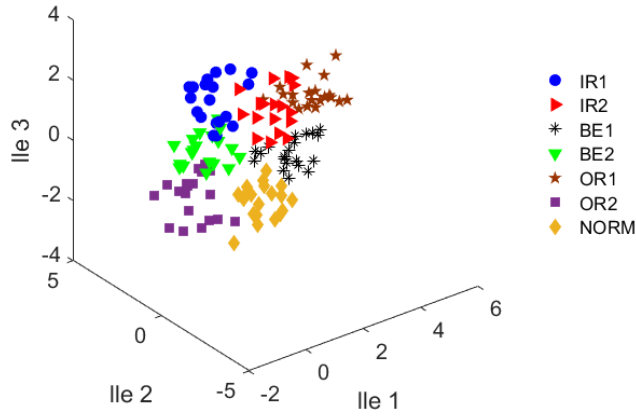


FIGURE 23. Clustering results based on MVPE-HLLE.

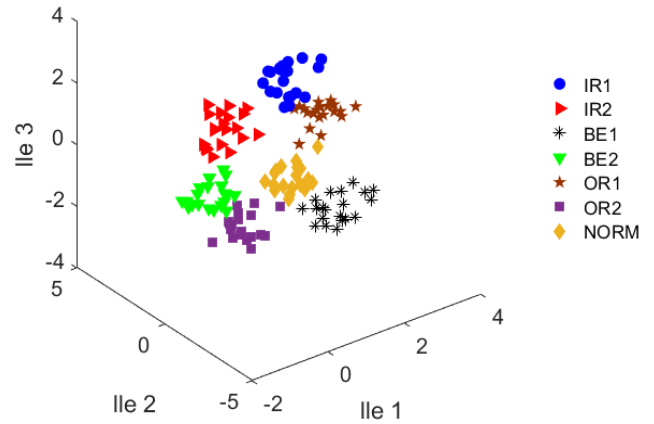


FIGURE 26. Clustering results based on SVMWPE-LLE.

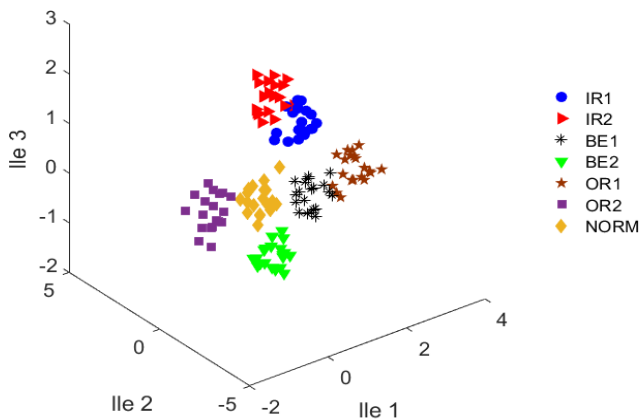


FIGURE 24. Clustering results based on SVMPE-HLLE.

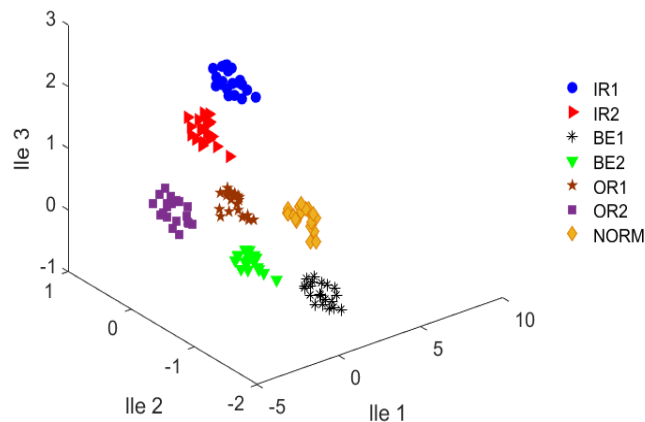


FIGURE 25. Clustering results based on SVMWPE-HLLE.

the samples of each state were extracted by SVMWPE-HLLE feature to obtain 20 low-dimensional matrices, 5 samples were randomly selected as training data, and the remaining 15 samples were used as test data. Among them, the activation function of ELM selects the Sigmoid function [35], and the number of neurons in the hidden layer was set to 25. Then, the training samples were inputted into the ELM for

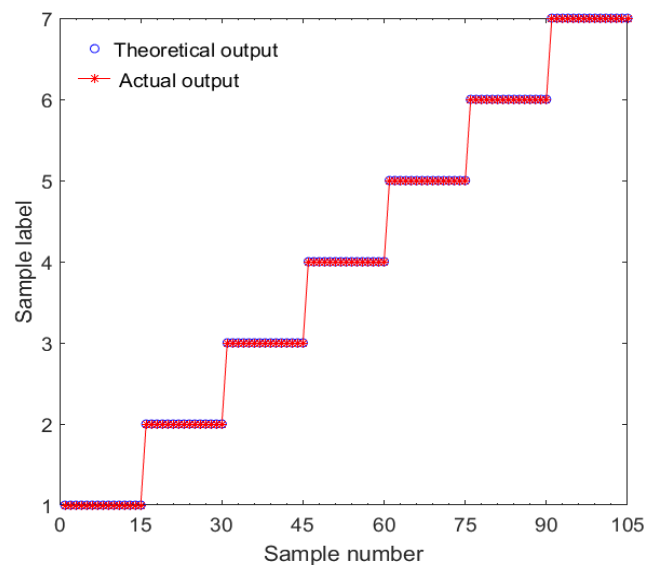


FIGURE 27. Recognition results based on SVMWPE-HLLE and ELM.

training. Finally, the test samples were inputted into the trained ELM for identification. The results are shown in Fig. 27. We observed that the actual output results were exactly the same as the theoretical results, and the fault identification rate was 100% (105/105), indicating that the method proposed in this paper can effectively identify the fault state of the rolling bearing.

In order to illustrate the superiority of SVMWPE-HLLE, all samples of each state were also extracted by MPE-HLLE, MWPE-HLLE, SVMPE-HLLE, and SVMWPE-LLE, and input to the ELM classifier for training and testing. The fault identification results are shown in Table 2. We observed that when the rolling bearing was in normal state, the recognition rate of the five methods is 100%, indicating that the five methods can accurately identify the normal state. When the rolling bearing was in a normal state, the recognition rates of the five methods were all 100%, indicating that the five methods can accurately identify the normal state. For several other states, SVMWPE-HLLE had no misclassification, whereas the other

TABLE 2. Fault identification results of five methods.

Rolling bearing condition	MFE-HLLE		MWPE-HLLE		SVMPE-HLLE		SVMWPE-HLLE		SVMWPE-LLE	
	Wrong Score	Recognition rate (%)	Wrong Score	Recognition rate (%)	Wrong Score	Recognition rate (%)	Wrong Score	Recognition rate (%)	Wrong Score	Recognition rate (%)
IR1	1	93.33	2	86.67	0	100	0	100	1	93.33
IR2	2	86.67	1	93.33	1	93.33	0	100	0	100
BE1	2	86.67	2	86.67	2	86.67	0	100	1	93.33
BE2	1	93.33	0	100	0	100	0	100	0	100
OR1	1	93.33	1	93.33	0	100	0	100	1	93.33
OR2	3	80.00	2	86.67	1	93.33	0	100	1	93.33
NORM	0	100	0	100	0	100	0	100	0	100
Total	10	90.48	8	92.38	4	96.19	0	100	4	96.19

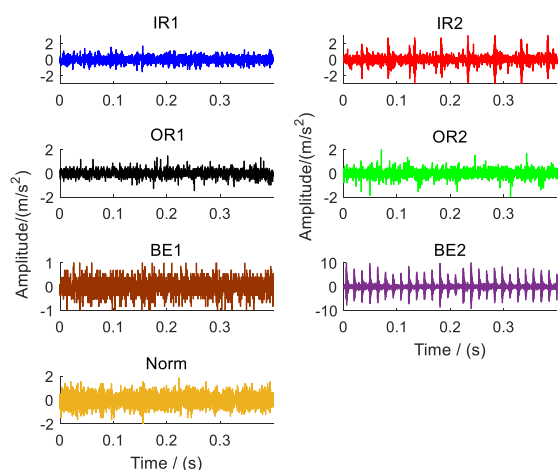


FIGURE 28. Time domain waveforms of seven states of rolling bearing.

four had misidentified. It showed that the recognition rate of SVMWPE-HLLE was better than the other four methods in both the overall and individual states, which verified the superiority of SVMWPE-HLLE in feature extraction.

2) EXPERIMENTAL CASE 2

The experimental data of Western Reserve University were collected under almost pure working conditions and not generic. In order to verify the applicability and superiority of this method under actual working conditions, the bearing selects 6308 deep groove ball bearing. Under the conditions of rotational speed was 1309 r/min and sampling frequency was 10240 Hz, seven states of Normal (Norm), Inner Race (IR1, IR2), Ball Element (BE1, BE2) and Outer Race (OR1, OR2) were collected. Each state took 30 samples with a data length of 4096 points, 10 samples were randomly selected as training data, and the remaining 20 samples were used as test data. The detailed description of the experimental data of the rolling bearing is shown in Table3. The time-domain waveforms of the seven states are shown in Fig. 28.

MPE, MWPE, SVMPE, and SVMWPE at 20 scales were calculated for all samples by selecting parameters $t = 1$ and

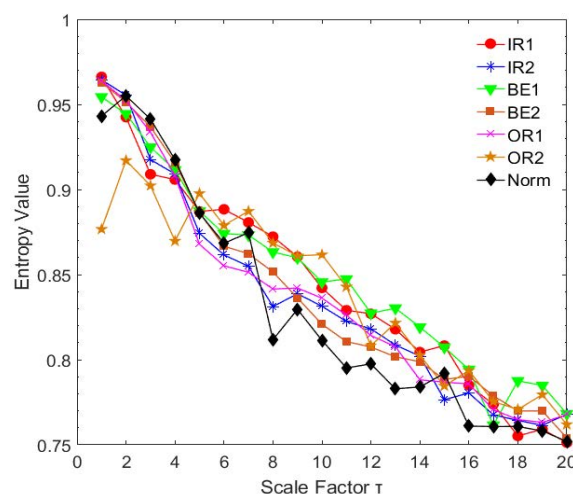


FIGURE 29. MPE of vibration signal of rolling bearing in different states.

TABLE 3. Detailed description of experimental data of rolling bearings.

Fault Location	Fault Size(mm ²)	Train set	Test set	Number
Inner Race(IR1)	1	10	20	1
Inner Race(IR2)	3.8	10	20	2
Ball element (BE1)	1	10	20	3
Ball element (BE2)	7	10	20	4
Outer Race (OR1)	3.8	10	20	5
Outer Race (OR2)	7	10	20	6
Normal(Norm)	0	10	20	7

$m = 6$. The results of MPE, MWPE, SVMPE, and SVMWPE in seven states are shown in Fig. 29-32. We observed that the entropy value of a single sample deviated greatly from the mean value of MPE and MWPE compared with SVMPE and SVMWPE, indicating that the computational stability of SVMPE and SVMWPE were better than MPE and SVMPE.

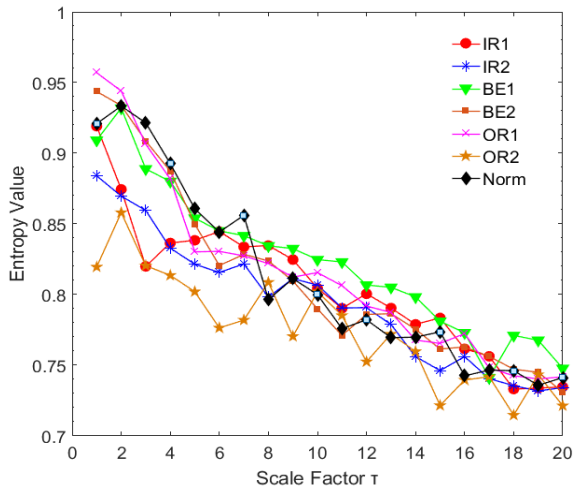


FIGURE 30. MVPE of vibration signal of rolling bearing in different states.

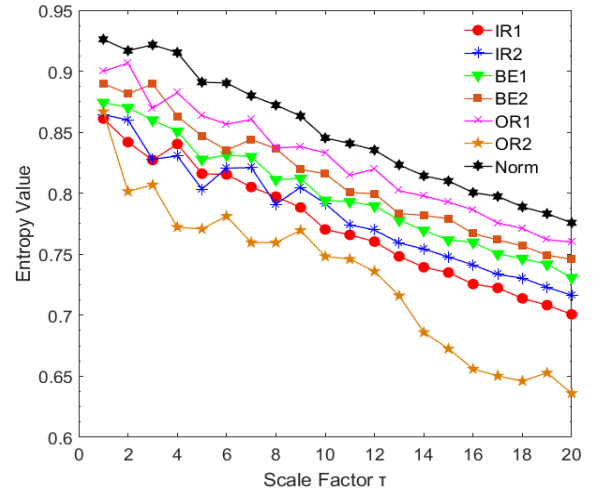


FIGURE 32. SVMWPE of vibration signal of rolling bearing in different states.

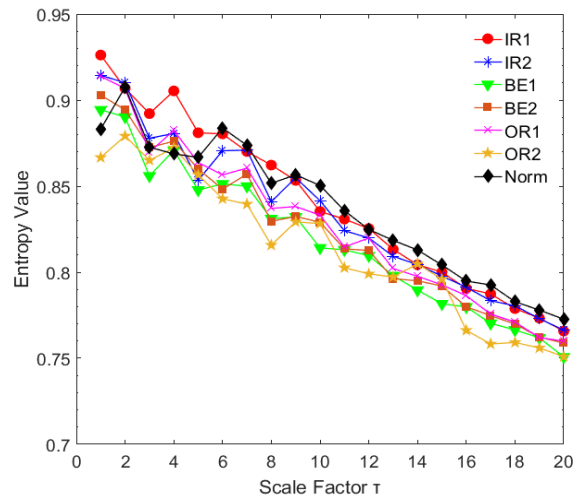


FIGURE 31. SVMPE of vibration signal of rolling bearing in different states.

The entropy values of each state of SVMWPE were distributed more dispersedly than SVMPE, at most scales, the entropy values were sorted from large to small as Norm, OR1, BE2, BE1, IR2, IR1, OR2, which indicated that SVMWPE has obvious advantages in feature extraction and has good stability.

MPE, MWPE, SVMPE, and SVMWPE of all samples were constructed by 210×20 high-dimensional feature matrix, and the low-dimensional feature matrix was obtained by HLLE. In order to verify the superiority of HLLE, SVMWPE feature matrix was additionally reduced by LLE algorithm, the output were shown in Fig. 33-37. The separation of MPE-HLLE and MWPE-HLLE were obviously poor. SVMPE-HLLE can roughly separated the seven states, but the separation effect was average, and the three-dimensional features of each state were scattered. SVMWPE-LLE cannot completely separate the seven states, and there were multiple crossovers. SVMWPE-HLLE can completely separate the

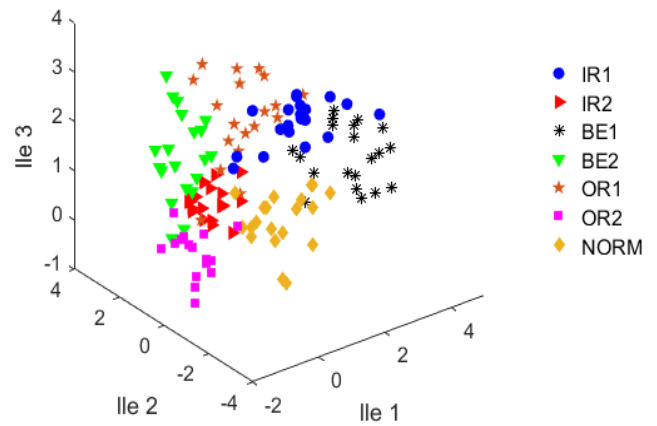


FIGURE 33. Clustering results based on MPE-HLLE.

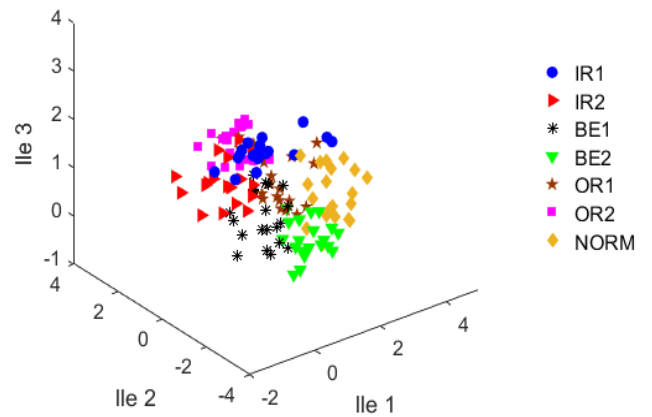


FIGURE 34. Clustering results based on MVPE-HLLE.

seven states, and the three-dimensional features of each state were relatively concentrated, and the clustering effect was obvious, which indicated that SVMWPE-HLLE exhibits

TABLE 4. Fault identification results of five methods.

Rolling bearing condition	MFE-HLLE		MWPE-HLLE		SVMPE-HLLE		SVMWPE-HLLE		SVMWPE-LLE	
	Wrong Score	Recognition rate (%)	Wrong Score	Recognition rate (%)	Wrong Score	Recognition rate (%)	Wrong Score	Recognition rate (%)	Wrong Score	Recognition rate (%)
IR1	2	90	3	85	0	100	0	100	0	100
IR2	2	90	1	95	1	95	0	100	2	90
BE1	1	95	2	90	1	95	0	100	1	95
BE2	1	95	0	100	0	100	0	100	1	95
OR1	2	90	1	95	1	95	0	100	0	100
OR2	3	85	2	90	1	95	0	100	1	95
NORM	1	95	1	95	0	100	0	100	0	100
Total	12	91.43	10	92.86	4	97.14	0	100	5	96.43

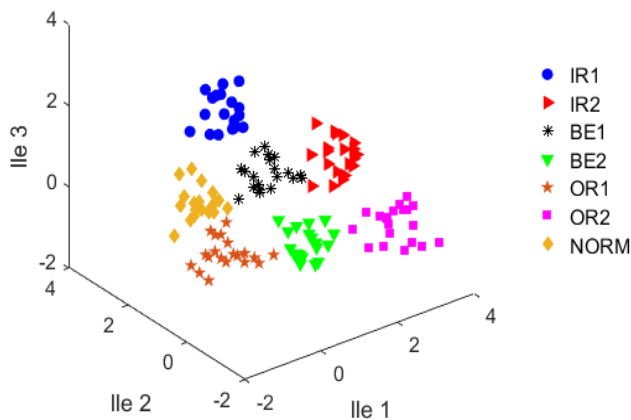


FIGURE 35. Clustering results based on SVMPE-HLLE.

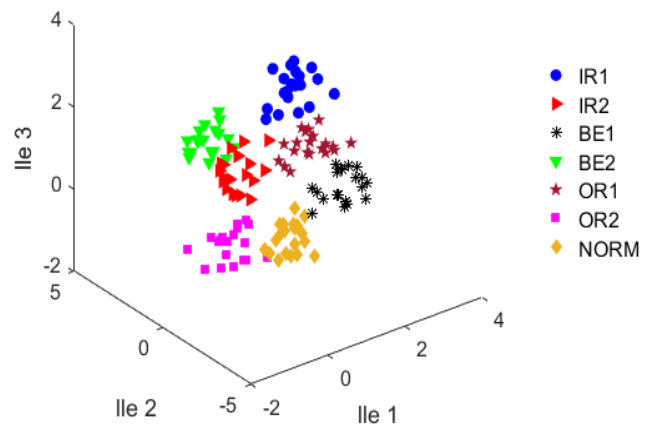


FIGURE 37. Clustering results based on SVMWPE-LLE.

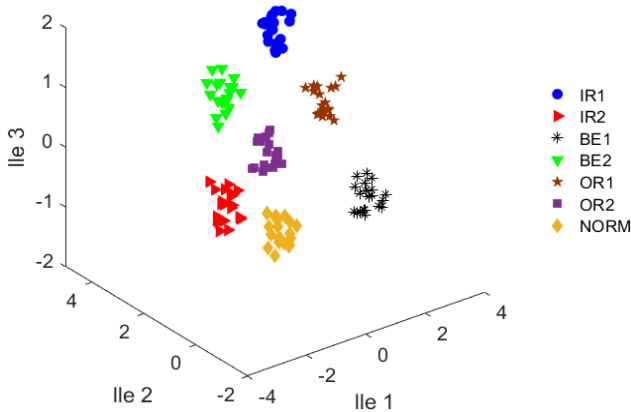


FIGURE 36. Clustering results based on SVMWPE-HLLE.

good ability to evaluate the fault degree of rolling bearing faults

The samples of each state were extracted by SVMWPE-HLLE feature to obtain 30 low-dimensional matrices, 10 samples were randomly selected as training data, and the remaining 20 samples were used as test data. Then, the training samples were inputted into the ELM for training. Finally, the test samples were inputted into the trained ELM

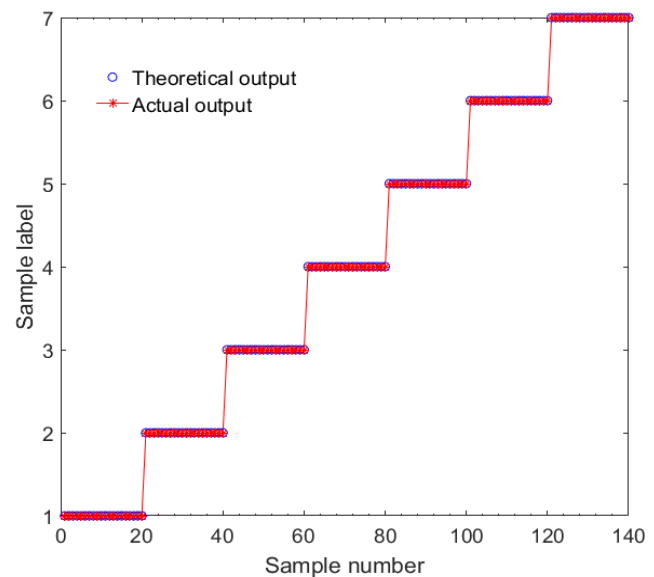


FIGURE 38. Recognition results based on SVMWPE-HLLE and ELM.

for identification. The results are shown in Fig. 38. The actual output results were exactly the same as the theoretical results, and the fault identification rate was 100% (140/140),

indicating that the fault diagnosis of rolling bearing can be effectively realized based on SVMWPE-HLLE and ELM.

V. CONCLUSION

1) In this paper, SVMWPE was proposed to measure the complexity of non-linear time series, which optimizes the coarse-grained sequence and fully considers the difference of amplitudes in the same symbol pattern in the permutation entropy calculation. Through the analysis of WGN and 1/f noise, the selection of various parameters in SVMWPE was determined it was verified that the superior of SVMWPE to MPE, MWPE and SVMPE.

All samples of each state were extracted by MPE-HLLE, MWPE-HLLE, SVMPE-HLLE and SVMWPE-LLE, and input to the ELM classifier for training and testing. The fault identification results are shown in Table 4. We observed that SVMWPE-HLLE has no misclassification in all states, whereas the other four are misidentified. It shows that the recognition rate of SVMWPE-HLLE is better than other four methods, which verifies the superiority of SVMWPE-HLLE in fault diagnosis.

2) SVMWPE-HLLE feature extraction method was proposed based on HLLE. The clustering effect was studied by comparing SVMWPE-HLLE with SVMWPE-LLE through the analysis of the three analog signals. SVMWPE-HLLE had better clustering effect than SVMWPE-LLE in feature extraction.

3) A new fault diagnosis method for rolling bearings was proposed and analyzed through two experimental cases of rolling bearings based on the respective advantages of SVMWPE-HLLE and ELM. The SVMWPE was compared with MPE, MWPE and SVMPE, and the results showed that the stability of SVMWPE was better than that those of MPE, MWPE and SVMPE. The clustering effect of the feature extraction method based on SVMWPE-HLLE was obvious, and it had a better ability to evaluate the fault degree of rolling bearing faults than SVMWPE-LLE, which verified the effectiveness and superiority of HLLE. SVMWPE-HLLE and ELM based fault diagnosis method had a higher fault recognition rate than the MPE-HLLE and ELM based method, MWPE-HLLE and ELM based method, SVMPE-HLLE and ELM based method, together with SVMWPE-LLE and ELM based method.

SVMWPE-HLLE can provide a new solution for rolling bearing fault diagnosis. SVMWPE-HLLE has certain advantages in characterizing fault characteristics, and could be extended to the fault diagnosis of other mechanical equipment.

ACKNOWLEDGMENT

The authors would like to thank the Case Western Reserve University for allowing them to use their data. This work has been supported by the “Three Vertical” Scientific Research Support Program (ZRCYP202207).

CONFLICTS OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

- [1] L. L. Li, “Research on rolling bearing fault diagnosis and life prediction based on deep learning,” M.S. thesis, Dept. Comput. Sci., Hunan Tech. Univ., Hunan, China, 2022.
- [2] Y. Qin, C. Li, F. Cao, and H. Chen, “A fault dynamic model of high-speed angular contact ball bearings,” *Mechanism Mach. Theory*, vol. 143, Jan. 2020, Art. no. 103627, doi: [10.1016/J.MECHMACHTHEORY.2019.103627](https://doi.org/10.1016/J.MECHMACHTHEORY.2019.103627).
- [3] M. Zhang, Z. Jiang, and K. Feng, “Research on variational mode decomposition in rolling bearings fault diagnosis of the multistage centrifugal pump,” *Mech. Syst. Signal Process.*, vol. 93, pp. 460–493, Sep. 2017, doi: [10.1016/J.YMSSP.2017.02.013](https://doi.org/10.1016/J.YMSSP.2017.02.013).
- [4] S. He, Y. Wang, and C. Liu, “Modified partial least square for diagnosing key-performance-indicator-related faults,” *Can. J. Chem. Eng.*, vol. 96, no. 2, pp. 444–454, 2018, doi: [10.1002/CJCE.23002](https://doi.org/10.1002/CJCE.23002).
- [5] L. N. Hao, “FPGA implementation of short-time Fourier transform and its application in fault diagnosis of rolling bearings,” M.S. thesis, Dept. Mech. Elect. Eng., Xian Univ., Xian, China, 2021.
- [6] X. J. Li, N. S. He, and K. F. He, “Cylindrical roller bearing diagnosis based on wavelet packet approximate entropy and SVM,” *J. Vib. Meas. Diagnosis*, vol. 35, no. 6, pp. 1031–1037, Dec. 2015, doi: [10.16450/J.CNKI.ISSN.1004-6801.2015.06.005](https://doi.org/10.16450/J.CNKI.ISSN.1004-6801.2015.06.005).
- [7] Y. H. Gu, T. T. Zhu, and W. J. Rao, “Fault diagnosis of rolling bearing based on EMD binary image and CNN,” *J. Vib. Meas. Diag.*, vol. 41, no. 1, pp. 105–113, 2021, doi: [10.16450/J.CNKI.ISSN.1004-6801.2021.01.015](https://doi.org/10.16450/J.CNKI.ISSN.1004-6801.2021.01.015).
- [8] H. Li, Q. Zhang, and X. R. Qin, “Fault diagnosis method for rolling bearings based on short-time Fourier transform and convolution neural network,” *J. Vib. Shock*, vol. 39, no. 19, pp. 124–131, Oct. 2018.
- [9] Y. P. Wu, J. F. Wu, and X. L. Xu, “Application of FastICA joint denoising method based on wavelet analysis in rolling bearing fault diagnosis,” *Chin. J. Sci. Inst.*, vol. 28, no. 18, pp. 2183–2188, Sep. 2017, doi: [10.3969/J.ISSN.1004-132X.2017.18.006](https://doi.org/10.3969/J.ISSN.1004-132X.2017.18.006).
- [10] Y. Yang, D. J. Yu, and J. S. Cheng, “Fault diagnosis method of rolling bearing based on EMD and neural network,” *J. Vib. Shock*, vol. 28, no. 1, pp. 85–89, Feb. 2005, doi: [10.13465/J.CNKI.JVS.2005.01.024](https://doi.org/10.13465/J.CNKI.JVS.2005.01.024).
- [11] S. M. Pincus, “Approximate entropy as a measure of system complexity,” *Proc. Nat. Acad. Sci. USA*, vol. 88, pp. 2297–2301, Mar. 1991, doi: [10.1073/PNAS.88.6.2297](https://doi.org/10.1073/PNAS.88.6.2297).
- [12] J. S. Richman and J. R. Moorman, “Physiological time series analysis using approximate entropy and sample entropy,” *Amer. J. Physiol.-Heart Circulatory Physiol.*, vol. 278, no. 6, pp. 2039–2049, Jun. 2000, doi: [10.1152/AJPHEART.2000.278.6.H2039](https://doi.org/10.1152/AJPHEART.2000.278.6.H2039).
- [13] W. T. Chen, “Research on feature extraction of surface EMG signals based on entropy,” Ph.D. dissertation, Dept. Biomed. Eng., Shanghai Jiao Tong Univ., Shanghai, China, 2008.
- [14] M. Rostaghi and H. Azami, “Dispersion entropy: A measure for time-series analysis,” *IEEE Signal Process. Lett.*, vol. 23, no. 5, pp. 610–614, May 2016, doi: [10.1109/LSP.2016.2542881](https://doi.org/10.1109/LSP.2016.2542881).
- [15] C. Bandt and B. Pompe, “Permutation entropy: A natural complexity measure for time series,” *Phys. Rev. Lett.*, vol. 88, no. 17, Apr. 2002, Art. no. 174102, doi: [10.1103/PHYSREVLETT.88.174102](https://doi.org/10.1103/PHYSREVLETT.88.174102).
- [16] H. F. Zhao, S. L. Liu, and X. S. Gu, “Fault complexity measure analysis of reciprocating compressor valve based on approximate entropy,” *Chem. Eng. Mach.*, vol. 37, no. 3, pp. 296–298, Jun. 2010.
- [17] L. G. Zhang, P. Li, M. M. Li, S. Q. Zhang, and Z. F. Zhang, “Fault diagnosis of rolling bearing based on ITD fuzzy entropy and GG clustering,” *Chin. J. Sci. Inst.*, vol. 35, no. 11, pp. 2624–2632, Nov. 2014, doi: [10.19650/J.CNKI.CJSI.2014.11.027](https://doi.org/10.19650/J.CNKI.CJSI.2014.11.027).
- [18] F. Z. Feng, A. W. Si, G. Q. Rao, and P. C. Jiang, “Early fault diagnosis technology of bearings based on wavelet correlation permutation entropy,” *J. Mech. Eng.*, vol. 48, no. 13, pp. 73–79, Jul. 2012, doi: [10.3901/JME.2012.13.073](https://doi.org/10.3901/JME.2012.13.073).
- [19] S. J. Wu, F. Z. Feng, and C. Z. Wu, “Research on fault diagnosis method of tank planetary gearbox based on VMD-DE,” *J. Vib. Shock*, vol. 39, no. 10, pp. 170–179, May 2020, doi: [10.13465/J.CNKI.JVS.2020.10.023](https://doi.org/10.13465/J.CNKI.JVS.2020.10.023).
- [20] Y. Wu, P. Shang, and Y. Li, “Multiscale sample entropy and cross-sample entropy based on symbolic representation and similarity of stock markets,” *Commun. Nonlinear Sci. Numer. Simul.*, vol. 56, pp. 49–61, Mar. 2018, doi: [10.1016/J.CNSNS.2017.07.021](https://doi.org/10.1016/J.CNSNS.2017.07.021).
- [21] L. Tang, H. Lv, and L. Yu, “An EEMD-based multi-scale fuzzy entropy for complexity analysis in clean energy markets,” *Appl. Soft Comput.*, vol. 56, pp. 124–133, Jul. 2017, doi: [10.1016/J.ASOC.2017.03.008](https://doi.org/10.1016/J.ASOC.2017.03.008).

- [22] V. T. Tran and B.-S. Yang, "An intelligent condition-based maintenance platform for rotating machinery," *Expert Syst. Appl.*, vol. 39, no. 3, pp. 2977–2988, Feb. 2012, doi: [10.1016/J.ESWA.2011.08.159](https://doi.org/10.1016/J.ESWA.2011.08.159).
- [23] A. Humeau-Heurtier, "The multiscale entropy algorithm and its variants: A review," *Entropy*, vol. 17, no. 5, pp. 3110–3123, May 2015, doi: [10.3390/E17053110](https://doi.org/10.3390/E17053110).
- [24] S.-D. Wu, P.-H. Wu, C.-W. Wu, J.-J. Ding, and C.-C. Wang, "Bearing fault diagnosis based on multiscale permutation entropy and support vector machine," *Entropy*, vol. 14, no. 8, pp. 1343–1356, Jul. 2012, doi: [10.3390/E14081343](https://doi.org/10.3390/E14081343).
- [25] S. Yin, X. Zhu, and O. Kaynak, "Improved PLS focused on key-performance-indicator-related fault diagnosis," *IEEE Trans. Ind. Electron.*, vol. 62, no. 3, pp. 1651–1658, Mar. 2015, doi: [10.1109/TIE.2014.2345331](https://doi.org/10.1109/TIE.2014.2345331).
- [26] L. Ma, M. M. Crawford, X. Yang, and Y. Guo, "Local-manifold-learning-based graph construction for semisupervised hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, May 2015, doi: [10.1109/TGRS.2014.2365676](https://doi.org/10.1109/TGRS.2014.2365676).
- [27] Y. H. Wei, W. Liu, Y. J. Yang, and J. J. Su, "Bearing fault diagnosis model based on LLE and its improved distance algorithm," *Mod. Mach. Tool. Auto. Manuf. Tech.*, no. 7, pp. 73–77, Jul. 2018, doi: [10.13462/J.CNKI.MMTAMT.2018.07.019](https://doi.org/10.13462/J.CNKI.MMTAMT.2018.07.019).
- [28] D. L. Donoho and C. Grimes, "Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data," *Proc. Nat. Acad. Sci. USA*, vol. 100, no. 10, pp. 5591–5596, May 2003, doi: [10.1073/PNAS.1031596100](https://doi.org/10.1073/PNAS.1031596100).
- [29] W. Aziz and M. Arif, "Multiscale permutation entropy of physiological time series," in *Proc. Pakistan Sect. Multitopic Conf.*, Dec. 2005, pp. 1–6.
- [30] J. D. Zheng, J. S. Cheng, and Y. Yang, "Multi-scale permutation entropy and its application in fault diagnosis of rolling bearings," *Chin. Mech. Eng.*, vol. 24, no. 19, pp. 2641–2646, May 2013, doi: [10.3969/J.ISSN.1004-132X.2013.19.017](https://doi.org/10.3969/J.ISSN.1004-132X.2013.19.017).
- [31] Z. L. Dong and J. D. Zheng, "Time-shift multi-scale weighted permutation entropy and GWO-SVM based fault diagnosis approach for rolling bearing," *Entropy*, vol. 21, no. 6, pp. 601–621, Jun. 2019, doi: [10.3390/E21060621](https://doi.org/10.3390/E21060621).
- [32] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, Dec. 2000, doi: [10.1126/SCIENCE.290.5500.2323](https://doi.org/10.1126/SCIENCE.290.5500.2323).
- [33] J. Zheng, Z. Dong, H. Pan, Q. Ni, T. Liu, and J. Zhang, "Composite multi-scale weighted permutation entropy and extreme learning machine based intelligent fault diagnosis for rolling bearing," *Measurement*, vol. 143, pp. 69–80, Sep. 2019, doi: [10.1016/J.MEASUREMENT.2019.05.002](https://doi.org/10.1016/J.MEASUREMENT.2019.05.002).
- [34] Case Western Reserve University. *Bearing Data Center*. Accessed: Feb. 20, 2022. [Online]. Available: <http://csegroups.case.edu/bearingdatacenter/pages/download-data-file>
- [35] W. Chen, Z. Wang, H. Xie, and W. Yu, "Characterization of surface EMG signal based on fuzzy entropy," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 15, no. 2, pp. 266–272, Jun. 2007, doi: [10.1109/TNSRE.2007.897025](https://doi.org/10.1109/TNSRE.2007.897025).



CHUNHUI LI received the bachelor's degree from the Environmental and Chemical Engineering College, Dalian University, China, in 2017, and the master's degree from the Mechanical Science and Engineering College, Northeast Petroleum University, China, in 2020. He is currently a Teaching Assistant with the Engineering College, Heilongjiang Bayi Agricultural University. His current research interests include mechanical fault diagnosis, noise, and vibration control.



YOUFU TANG received the master's degree from the Mechanical Science and Engineering College, Northeast Petroleum University, China, in 2006, and the Ph.D. degree in mechanical and electronic engineering from the School of Mechatronic Engineering and Automation, Shanghai University, China, in 2013. He is currently a Professor in engineering mechanics with the Mechanical Science and Engineering College, Northeast Petroleum University. His current research interests include mechanical fault diagnosis, noise, and vibration control.



GUOLIN XU received the bachelor's degree from the Power and Nuclear Engineering College, Harbin Engineering University, China, in 2004, and the master's degree from the Energy Science and Engineering, Harbin Institute of Technology, China, in 2019. He has been engaged in turbine design at Harbin Turbine Company Ltd., for more than 15 years. He is currently a full-time Teacher with the Engineering College, Heilongjiang Bayi Agricultural University. His current research interests include mechanical structure design and failure analysis of power machinery.

• • •