Hindawi
Shock and Vibration
Volume 2022, Article ID 5859155, 11 pages
https://doi.org/10.1155/2022/5859155

Research Article

Rolling Bearing Fault Diagnosis Method Based on MCMF and SAIMFE

Dejun Meng, Changyun. Miao, Xianguo Li, Jia Shi, Yi Liu, and Jie Li

Tianjin Key Laboratory of Optoelectronic Detection Technology and System, School of Electronics and Information Engineering, Tiangong University, Tianjin, China

Correspondence should be addressed to Changyun. Miao; mengdejun@tiangong.edu.cn

Received 12 August 2022; Revised 2 November 2022; Accepted 7 November 2022; Published 12 December 2022

Academic Editor: Roger Serra

Fault diagnosis of rolling bearing is important for ensuring the safe operation of industrial machinery. In order to improve diagnosis accuracy of bearing fault, a rolling bearing fault diagnosis method based on multiscale combined morphological filter (MCMF) and self-adaption improved multiscale fuzzy entropy (SAIMFE) is proposed in this paper. First, the MCMF is designed to eliminate noise and preserve fault information more effectively. Second, SAIMFE is proposed to extract bearing fault features, and the optimized scale factor of SAIMFE is determined based on the absolute skewness. Third, some experiments are completed to demonstrate the effectiveness and superiority of the proposed method. The experimental results show that the proposed method not only has high diagnosis accuracy but also less dependent on the diagnosis model.

1. Introduction

Rolling bearings are the most critical and easily damaged components in rotating machinery; the availability, reliability, and productivity of rotating machinery depends on the health state of rolling bearings; therefore, the rolling bearing fault diagnosis is very vital to the stable, reliable, and efficient operation of rotating machinery [1, 2]. Vibration signals of rolling bearing fault can reflect the fault features [3, 4], so the time domain, frequency domain, and time-frequency analysis methods based on vibration signals are used to diagnose bearing fault frequently [5–7]. In recent years, with the rapid development of machine learning (ML) algorithms and deep learning algorithms, data-driven fault diagnosis methods have been given more and more attention [8]. Li et al. [9] used hierarchical symbolic dynamic entropy and binary tree support vector machine for rolling bearing fault diagnosis. Hou et al. [10] proposed a rolling bearing clustering fault diagnosis method based on ensemble empirical mode decomposition, permutation entropy (PE), linear discriminant analysis (LDA), and the Gath–Geva clustering algorithm. He et al. [11] combined multiscale permutation entropy (CMPE) and a reverse cognitive fruit fly optimization algorithm optimized for extreme learning machine (RCFOA-ELM) to diagnose rolling bearing faults. Wan et al. [12] combined a back-propagation neural network (BPNN) optimized by the quantum particle swarm optimization (QPSO) algorithm and the Dempster–Shafer evidence theory to diagnose rolling bearing faults. Xie et al. [13] developed a rolling bearing fault diagnosis method based on a deep belief network optimized by Nesterov momentum.

Extraction of fault features is critical for achieving fault diagnosis of bearing based on machine learning algorithms and deep learning algorithms [14]. Entropy is a vague concept, which is used to measure the uncertainty of a piece of information. Approximate entropy [15], sample entropy [16], and fuzzy entropy [17, 18] are used to extract the faulty features of bearing. However, approximate entropy often produces illegal values. Sample entropy lacks stability. Fuzzy entropy only describes the characteristics of the signal on a single scale [19]. Costa et al. [20] proposed the concept of multiscale entropy in 2002, which is used to measure the complexity and self-similarity of the time series under different scale factors. This assumes the characteristics of fast computation and anti-interference under multiscale factors.

If one sequence of entropy values is more important than the other sequence of entropy values under more scales, the sequence is more complex than the other sequence. If one sequence of entropy values monotonously decreases with increasing scale factors, then the sequence structure is relatively simple. It contains more information on the smallest scales. If one sequence of entropy values monotonously increases with increasing scale factors, then the sequence contains more information on the multiscale. Zheng et al. [21] proposed multiscale fuzzy entropy (MFE), which is used to feature the complexity and irregularity of rolling bearing vibration signals. But MFE has the disadvantage of information loss. Feng et al. [22] proposed a novel cyclic-correntropy based health indicator to monitor the gear surface degradation induced by gear wear progression. With the help of the novel indicator developed, the health status of the gearbox can be well evaluated. In addition, the signal noise reduction effect also directly affects the feature extraction of fault features. Feng et al. [23] proposed a novel bandwidth selection methodology for the Vold–Kalman filter. Through comprehensively depicting the targeted harmonic response using features in multiple domains, the rational bandwidth be selected for Vold–Kalman filtering. The effectiveness and superiority of the proposed adaptive Vold–Kalman filtration for wind turbine planetary gearbox diagnostics are demonstrated and validated experimentally.

In this paper, a rolling bearing fault diagnosis method based on multiscale combined morphological filter (MCMF) and self-adaption improved multiscale fuzzy entropy (SAIMFE) is proposed. Firstly, the MCMF based on morphological filter (MF) is developed to eliminate noise and preserve fault information more effectively. Secondly, based on multiscale fuzzy entropy (MFE) and improved multiscale fuzzy entropy (IMFE), SAIMFE is proposed to extract bearing fault features, and the optimized scale factor of SAIMFE is determined based on the absolute skewness. Thirdly, some experiments are completed to demonstrate the effectiveness and superiority of the proposed method. The experimental results show that the proposed method not only has high diagnosis accuracy but also less dependent on the diagnosis model.

2. Principle and Verification

2.1. Multiscale Combined Morphological Filter. MF was initially used for image processing [24–27]. Recently, MF has been used as a denoting method in machinery fault diagnosis. The four bases which include erosion, dilation, opening, and closing morphological operations are often used in various applications. Let $f(n)$ be the signal that is the function over a domain $F = (0, 1, \ldots\ldots, N-1)$. Let $g(m)$ be the structure element (SE) over a domain $G = (0, 1, \ldots\ldots, M-1, M \le N)$. The four operations are defined [24] by equations (1)–(4)

Erosion:

$$(f \oplus g)(n) = \min \{f(n+m) - g(m)\}. \quad (1)$$

Dilation:

$$(f \otimes g)(n) = \min \{f(n-m) + g(m)\}. \quad (2)$$

Opening:

$$(f \odot g)(n) = (f \oplus g \otimes g)(n). \quad (3)$$

Closing:

$$(f \ominus g)(n) = (f \otimes g \oplus g)(n), \quad (4)$$

where $\oplus$, $\otimes$, $\odot$, and $\ominus$ denote the erosion, dilation, opening, and closing operations, respectively. They can be capable of picking up positive or negative impulses from vibration signals. Based on the four basic operations, difference filtering (DF) [24], opening-closing and closing-opening (COCO) [27–29], and average operations (AVG) [30] are developed for machine fault diagnosis.

Considering the different effects of the four basic operations on impulse, they can be divided into two categories. One can retain the positive impulses and suppress the negative impulses through dilation and closing operation (DC). The other can retain the negative impulse and remove the positive impulses through erosion and opening operation (EO). Thus, these operations can be integrated to enhance the effect of positive or negative impulses in vibration signals. A combined morphological filter (CMF), which is expressed by equations (5)–(7), is proposed.

$$F_{\mathrm{DC}}(f(n)) = (f \otimes g \ominus g)(n), \quad (5)$$

$$F_{\mathrm{EO}}(f(n)) = (f \oplus g \odot g)(n), \quad (6)$$

$$F_{\mathrm{CMF}}(f(n)) = F_{\mathrm{DC}} + F_{\mathrm{EO}}. \quad (7)$$

The SE selection affects the performance of MF significantly. The MF with single-scale SE has defects in signal reconstruction for complex signals. The MF with different scales SE has different noise reduction effects. The MF with small-scale SE has good performance in detail of the original signal preserving and poor noise suppression effect. The MF with big-scale SE has good performance in noise suppression, but its smoothing effect of local waves became rough. In this study, an MCMF is proposed to reduce signal noise. According to [25], the flat SE is selected as the unit SE (USE), which is expressed by

$$g = \frac{\{A_1, A_2, A_3 \ldots\ldots, A_i\}}{N}, \quad i = 1, 2, 3 \ldots\ldots N, \quad (8)$$

where A_i is the amplitude of the i th faulty feature frequency, N is the length of USE, and g_λ expressed by (9) which is multiscale SE obtained by $(\lambda - 1)$ times dilation operation of USE.

$$g_\lambda = \underbrace{g \otimes g \otimes \ldots\ldots \otimes g}_{(\lambda-1)\text{ times}}. \quad (9)$$

g_λ as SE is brought into equations (5)–(7), and the final output signal under λ scale can be expressed by

$$y = \sum_{j=1}^{\lambda} w_j y_j, \quad j = 1, 2, 3 \ldots\ldots \lambda, \tag{10}$$

where y_j is the output signal under j scale and w_j is the weight coefficient of j scale. In this study, correlation between the filtered signal under j scale and the original signal is selected to determine w_j. By equation (11), w_j can be calculated as follows:

$$w_j = \frac{c_j}{c_1 + c_2 \ldots\ldots c_\lambda}, \tag{11}$$

where λ is the scale of SE, and c_j is the correlation between the filtered signal under j scale and the original signal.

2.2. Self-Adaption Improved Multiscale Fuzzy Entropy

2.2.1. MFE and IMFE. MFE technology contains two steps [21]. Firstly, the coarse-grained procedure is used to obtain a time series with multiple scales from the original time series. Secondly, fuzzy entropy at each coarse-grained time series is calculated. Two procedures of MFE technology are described as follows [31]:

(1) We obtain the coarse-grained time series at a scale factor of τ (τ is a positive integer). The original time series is divided into disjointed windows with length τ, and the data points are averaged inside each window, namely, according to the equation (11), the coarse-grained time series at a scale factor of τ, y_τ can be constructed.

$$y_j^\tau = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} u(i), 1 \le j \le \frac{N}{\tau}, \tag{12}$$

where N is the data length and τ is the scale factor.

(2) In MFE technology, the fuzzy entropy of each coarse-grained time series is calculated and then plotted as the function of the scale factor τ, which can be expressed as

$$\mathrm{MFE}(x, \tau, m, r) = \mathrm{FuzzyEn}\left(y_j^{\tau,m,r}\right), \tag{13}$$

where x, $\{x_1, x_2, x_3, \ldots, x_N\}$, is the original time series, m is the embedding dimension, and r is the similarity tolerance.

Note that the r in the calculation for different scales is the same, which is obtained by $r = \lambda \times \mathrm{SD}$, and SD is the standard deviation of the original time series.

Depending on equation (11), the larger the scale factor, the shorter the coarse-grained sequence length, and the deviation of the entropy value gradually increases as the coarse-grained sequence length decreases. Moreover, effective information may be missed in coarse-grained procedure. For example, when τ is two, coarse-grained procedure is illustrated in Figure 1.

As shown in Figure 1, $y_j^{(2)}$, which is calculated using the relationship between X_i and X_{i+1} ($i = 1, 3, 5 \ldots\ldots$), is a coarse-grained sequence. But uncoarse-grained sequence ($z_j^{(2)}$), which is calculated using the relationship between X_i and X_{i+1} ($i = 2, 4, 6 \ldots\ldots$), is missed, whose disadvantages will affect the bearing fault diagnosis. Therefore, IMFE is proposed based on the moving average in this paper.

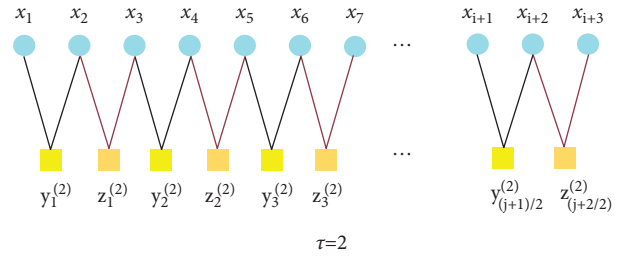

Figure 1: Coarse-grained procedure with scale factor to 2.

For the disadvantages of traditional MFE, IMFE is proposed. The main difference between IMFE and MFE is the coarse-grained procedure. In this study, an improved coarse-grained procedure is proposed depending on equation (14) [32].

$$y_j^\tau = \frac{1}{\tau} \sum_{i=j}^{j+\tau-1} u(i), 1 \le j \le \frac{N}{\tau}. \tag{14}$$

To illustrate the difference between IMFE and MFE, the IMFE algorithm with scaling factor of 2 is shown in Figure 2.

As shown in Figure 2, feature information of adjacent data is used in the improved coarse-grained procedure, which avoids information omission.

Values of IMEF and MEF are influenced by the data length (N), embedding dimension (m), and similar tolerance (r). Gaussian white noise is used to analyze the effects of different parameters on IMFE and MFE below.

(1) Influence of Similar Tolerance. In this section, we select six different r, which are 1.0 R_{SD}, 1.2 R_{SD}, 1.4 R_{SD}, 1.6 R_{SD}, 1.8 R_{SD}, 2.0 R_{SD}, respectively, to quantify the effects of r on IMFE and MFE, where R_{SD} is standard deviation of the time series. When m is set to 2, N is 2000 and τ is 20, MFE and IMFE of Gaussian white noise with different r are shown in Figure 3.

As shown in Figure 3, r has great influence on the values of MFE and IMFE. The larger r is, the smaller entropy is. The smaller r is, the larger entropy is. If r is too large, a lot of statistical information will be lost. If r is too small, the effect of the statistical characteristics estimated is not ideal. In addition, the entropy curve of IMFE is smoother than that of MFE, and the volatility of IMFE with different r is smaller than the volatility of MFE. Obviously, IMFE has smaller dependence on r. In this paper, r is set to 1.5 R_{SD}.

(2) Influence of Embedding Dimension. In this section, we select six different m, which are 2, 3, 4, 5, 6, and 7, respectively, to quantify the effects of m on IMFE and MFE. When r is set to 1.5 R_{SD}, N is 2000 and τ is 20, MFE and IMFE of Gaussian white noise with different m are shown in Figure 4.

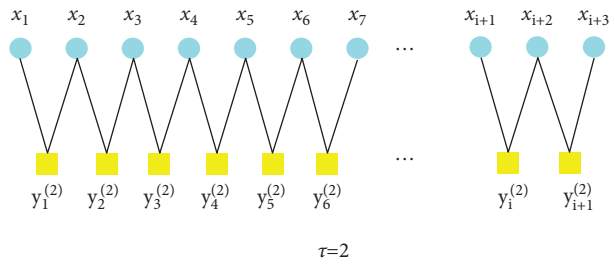

Figure 2: Improved coarse-grained procedure with scale factor of 2.

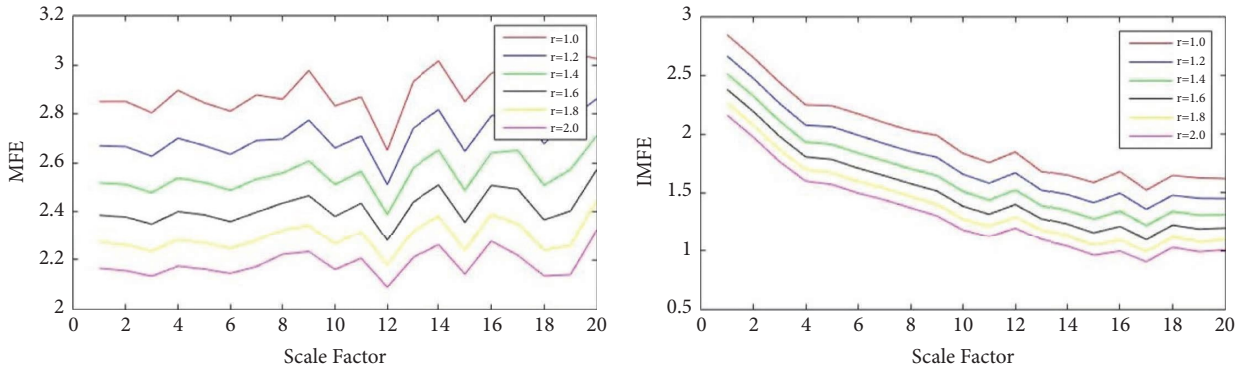

Figure 3: MFE and IMFE of Gaussian white noise with different r. (a) MFE of Gaussian white noise with different r and (b) IMFE of Gaussian white noise with different r.

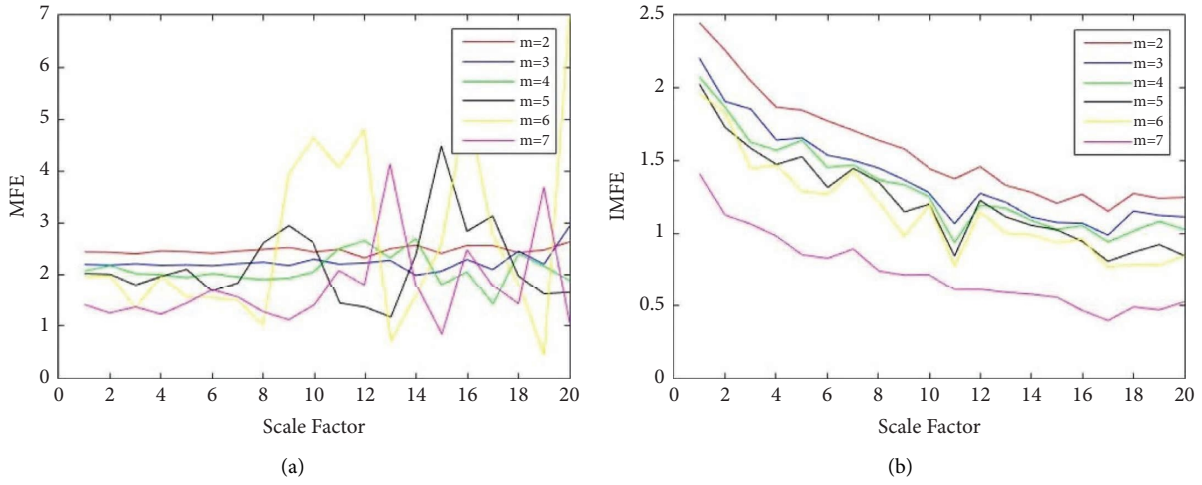

Figure 4: MFE and IMFE of Gaussian white noise with different m. (a) MFE of Gaussian white noise with different m. (b) IMFE of Gaussian white noise with different m.

As shown in Figure 4, when m is 2, the entropy curve of IMFE and MFE is smoother and the volatility of the entropy is smaller. Therefore, m is set to 2 in this paper.

(3) Influence of Data Length. In this section, we select six different N, which are 2000, 2500, 3000, 3500, 4000, and 4500, respectively, to quantify the effects of N on IMFE and MFE. When r is set to 1.5 R_{SD}, m is 2 and τ is 20, MFE and IMFE of Gaussian white noise with different N are shown in Figure 5.

As shown in Figure 5, the volatility of IMFE with different N is small than the volatility of MFE. Obviously, IMFE has smaller dependence on data length.

2.2.2. SAIMFE. SAIMFE is used to select optimal r of IMFE automatically, and r has great influence of diagnosis accuracy. The large-scale factor has redundancy which affects accuracy of diagnosis. The small-scale factor cannot fully reflect fault feature. So, absolute skewness of IMFE, which is

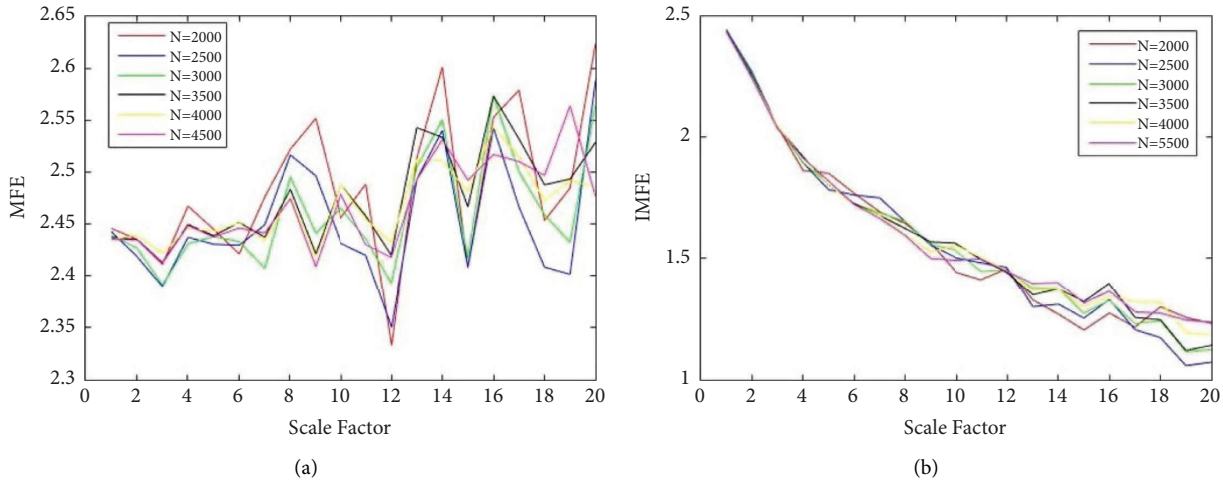

Figure 5: MFE and IMFE of Gaussian white noise with different N. (a) MFE of Gaussian white noise with different N. (b) IMFE of Gaussian white noise with different N.

utilized to select optimal r, is proposed in this paper. The smaller the absolute skewness of IMFE, the more accurate the fault feature. The skewness of IMFE can be expressed by equation (15).

$$S_{abke}(\text{IMFE}) = \left| \frac{3\left(V_{\text{IMFE}} - M_{\text{IMFE}}\right)}{S_{\text{IMFE}}} \right|, \tag{15}$$

where $S_{abke}(\text{IMFE})$ is the absolute skewness of IMFE, V_{IMFE} is the mean of IMFE, M_{IMFE} is the median of IMFE, and S_{IMFE} is the standard deviation of IMFE.

3. Experiment

3.1. Experiment Platform and Sample Dataset. In order to verify the effectiveness of the proposed method, a series of experiments is carried out on the rolling bearing sample dataset provided by Case Western Reserve University bearing data center. The sample dataset used in this experiment includes normal status (NS) data and three kinds fault data, namely, inner race fault (IRF) data under three kinds fault degree, outer race fault (ORF) data under three kinds fault degree, and rolling element fault (REF) data under three kinds fault degree. All the data are collected by sensors deployed at the drive end of bearing test-bed [33] shown in Figure 6 at 12 kHz sampling frequency when the bearing speed is 1730 r/min. Since the original rolling bearing dataset is small, it is difficult to effectively evaluate the performance of the proposed rolling bearing fault diagnosis method. Therefore, the overlapping sampling [8] method shown in Figure 7 is used to enhance the original vibration data. After preprocessing the enhanced vibration data, sample dataset is obtained, as listed in Table 1. As shown in Table 1, the sample dataset includes ten kinds of subsamples under different bearing status, namely, NS, IRF1, ORF1, REF1, IRF2, ORF2, REF2, IRF3, ORF3, and REF3. Each subsample contains 100 data with data length of 2,048. The sample dataset is divided into training sample and testing sample according to the ratio of 3 : 2.

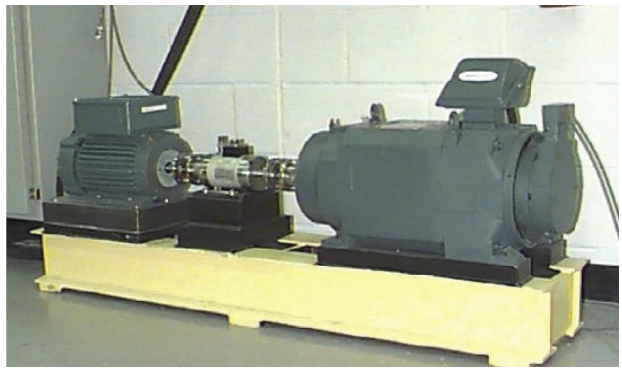

Figure 6: Bearing test-bed.

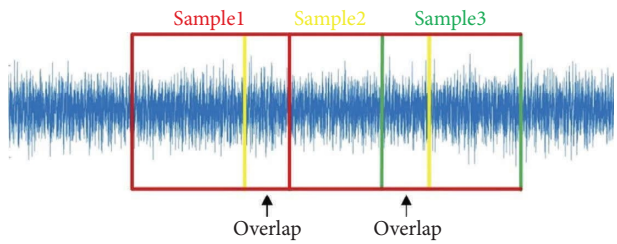

Figure 7: The schematic diagram of overlapping sampling.

3.2. Experiment and Result

3.2.1. Data Noise Reduction. Before the fault feature extraction, all the data in the sample dataset need to be processed by MCMF to eliminate noise. In this section, the subsample of IRF1 is selected to show the filtering effect, where the length of USE is 10. Figure 8 shows envelope spectra of the original signal and denoised signals using different methods. From Figure 8(a), we can see noise marked by red circles. The filtering effect of single-scale CMF is shown in Figure 8(b). As shown in Figure 8(b), filtering effect is not obvious and noise marked by green circles still exists in filtered signals. The filtering effect of MCMF is

Table 1: Sample dataset.

Bearing status	Fault diameter (mm)	Training sample number	Testing sample number	Sample length (N)	Class label
NS	0	60	40	2048	NS
IRF	0.1778	60	40	2048	IRF1
ORF	0.1778	60	40	2048	ORF1
REF	0.1778	60	40	2048	REF1
IRF	0.3556	60	40	2048	IRF3
ORF	0.3556	60	40	2048	ORF3
REF	0.3556	60	40	2048	REF3
IRF	0.5334	60	40	2048	IRF5
ORF	0.5334	60	40	2048	ORF5
REF	0.5334	60	40	2048	REF5

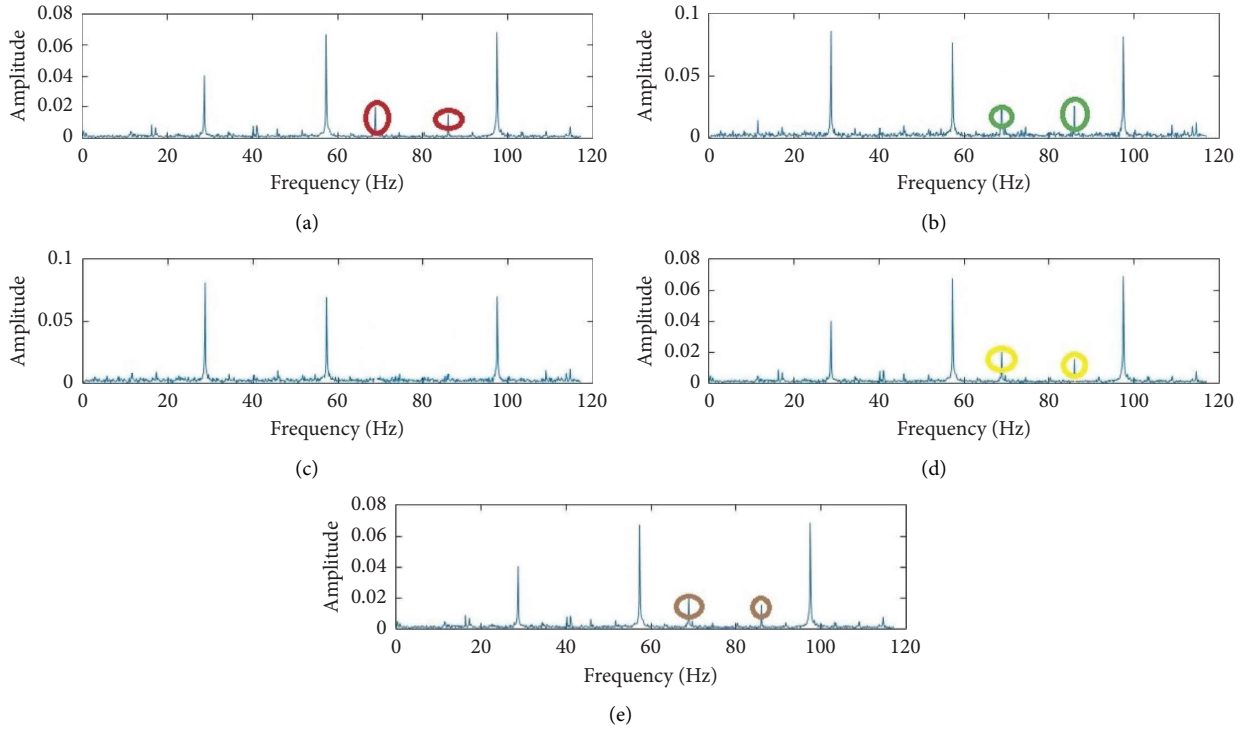

Figure 8: Envelope spectrums of original signal and denoise signal using different methods. (a) Envelope spectrum of original signal. (b) Envelope spectrum of denoise signal by single-scale CMF. (c) Envelope spectrum of denoise signal by multi-scale CMF. (d) Envelope spectrum of denoise signal by EMD. (e) Envelope spectrum of denoise signal by WD.

shown in Figure 8(c) when λ is 6. As shown in Figure 8(c), filtering effect is obvious and noise is eliminated effectively. Figures 8(d) and 8(e) show the filtering effects of empirical mode decomposition (EMD) and wavelet denoising (WD). As shown in Figures 8(d) and 8(e), noise marked by yellow circles and brown circles still exists in filtered two signals. According to the experimental results, filtering effect of MCMF is obvious.

3.2.2. Fault Feature Extraction. After noise reduction processing, the sample dataset is split into the original training sample and the original testing sample based on an overlapping sampling method. According to the principle of SAIMFE proposed above, optimal r needs to be selected based on the absolute skewness of IMFE when fault features are extracted. One training sample of IRF1 is selected to show the selection process of optimal r. The computation result of absolute skewness of IMFE is shown in Figure 9 under different r.

Other parameters of IMFE are listed in Table 2. As shown in Figure 9, the absolute skewness of IMFE is minimum when r is 4. Considering that the sample dataset used comes from the same platform and working condition, r of all samples is set to 4. After feature extraction, each fault feature dataset is a 4-dimensional feature vector.

Figure 10 is a 2-dimensional (2-D) fault feature extracted by the SAIMFE cluster diagram under different fault diameters. It can also be seen from Figure 10 that the fault features of NS, ORF1, ORF3, ORF5, REF1, and REF5 have good clustering, but the fault features of IRF1, IRF3, IRF5, and REF3 have poor clustering, and fault features cannot be

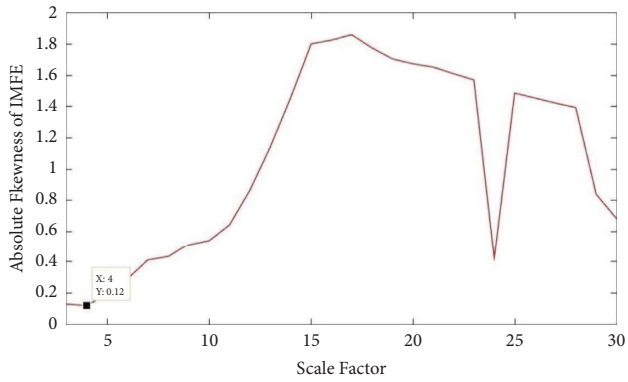

Figure 9: Computation result of absolute skewness of IMFE.

Table 2: The parameter settings of IMFE algorithm.

Parameter name	Parameter value
Similar tolerance	0.15
Embedding dimension	2
Data length	2048

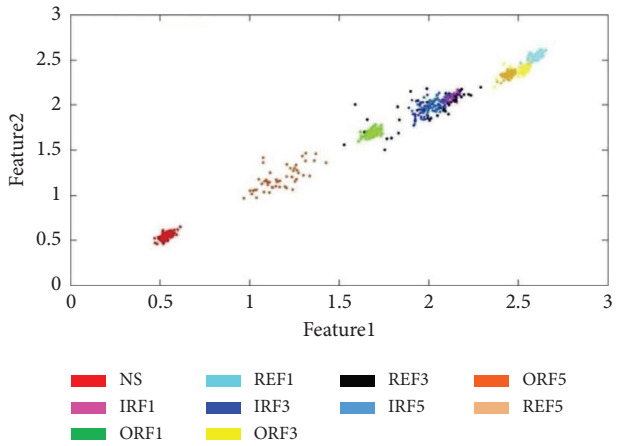

Figure 10: 2-D feature cluster diagram of different fault degrees.

distinguished. Considering that the clustering effect shown in Figure 10 is based on the first two fault features, the proposed method of fault feature extraction is helpful for the fault diagnosis of rolling bearing.

Figure 11 is a 3-dimensional (3-D) feature cluster diagram under different fault diameters (the first three features). It can be seen from the first three features that the clustering effect of 3D is more obvious than the clustering effect of 2D. According to Figures 10 and 11, the feature extraction method proposed in this paper has good performance in describing fault classification. The effect of different fault classifications is obvious.

3.2.3. Fault Diagnosis. In the first experiment, the subsamples of NS, IRF1, ORF1, and REF1 are used to train and test the fault diagnosis model of rolling bearing under different r. The purpose of the experiment is to verify diagnosis accuracy under the same fault degree and different

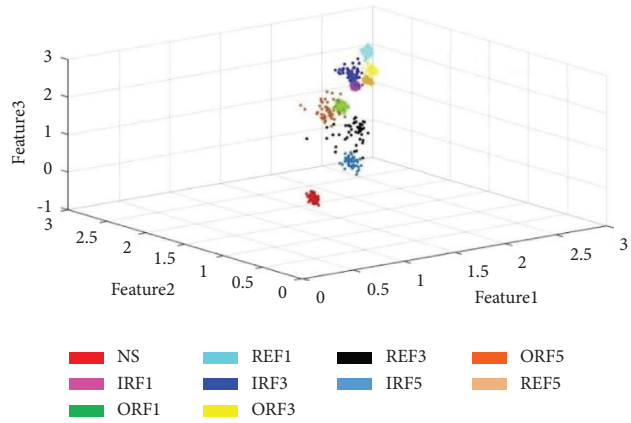

Figure 11: 3-D feature cluster diagram of different fault degrees.

damage. In this section, random forest (RF) [34] is selected as fault diagnosis model and the optimized parameters of the RF algorithm by grid search are listed in Table 3. Figure 12 presents the comparison of fault diagnosis results under different r. As can be seen from Figure 12, the highest diagnosis accuracy is 96.88% when r is 4, and diagnosis accuracy is less than 96.88% under other r. This experiment demonstrates that the proposed method is helpful for improving fault accuracy.

In the second experiment, the diagnosis models of RF, support vector machine (SVM), gradient boosting decision tree (GBDT), eXtreme gradient boosting (XGBoost), light gradient boosting machine (LightGBM), and decision tree (DT) [35–39] are selected to better explain the performance of the proposed method. Figure 13 presents the comparison of fault diagnosis result under different diagnosis models. As can be seen from Figure 13, there is no obvious difference in testing accuracy and training accuracy between different diagnosis models. This result shows that the proposed denoising and feature extracting methods have low dependence on the diagnosis model.

Figure 14 shows the confusion matrix of diagnosis results between different diagnosis models. The diagonal elements of the matrix represent the recall rate for each fault mode. It can be observed in Figure 14 that all the diagnosis models can fully identify NS. The models of GBDT, LightGBM, and DT can fully identify IRF. The number of misdiagnoses of IRF, ORF, and REF is very low.

In order to further evaluate the diagnostic effect of the proposed method, the sample dataset with different fault degrees listed in Table 2 is utilized to finish the third experiment. In the third experiment, the diagnosis models of RF, SVM, GBDT, XGBoost, LightGBM, and DT are selected as the fault diagnosis model. The results are shown in Table 4. From Table 4, diagnosis accuracies of different diagnosis model are very proximal. The results further illustrate that the proposed denoising and feature extraction method has low dependence on the diagnosis model.

Figure 15 shows the confusion matrix of the diagnosis result of RF under ten classification faults. As can be seen from Figure 15, the proposed method can fully identify five

Table 3: The parameters setting of RF algorithm.

Parameter name	Parameter value
n_estimators	101
max_depth	16
max_features	3
min_samples_leaf	1
min_samples_split	15
Criterion	Entropy

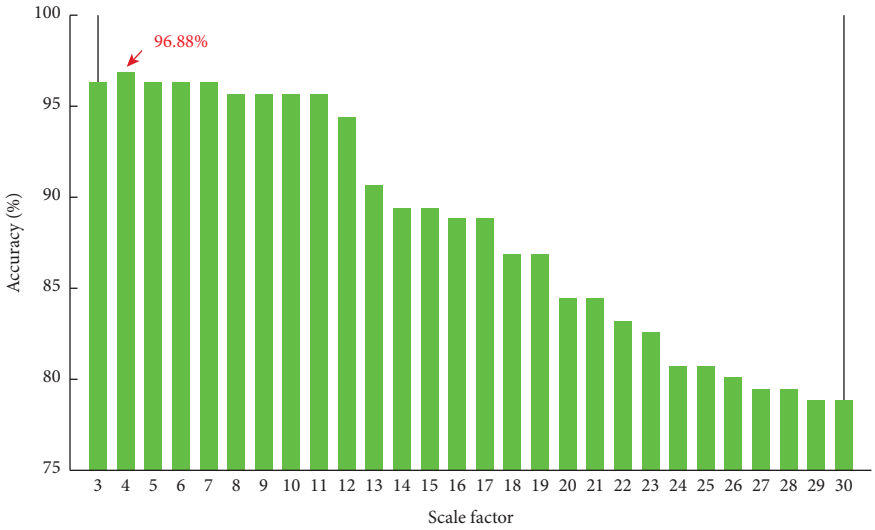

Figure 12: Comparison of fault diagnosis result under different r.

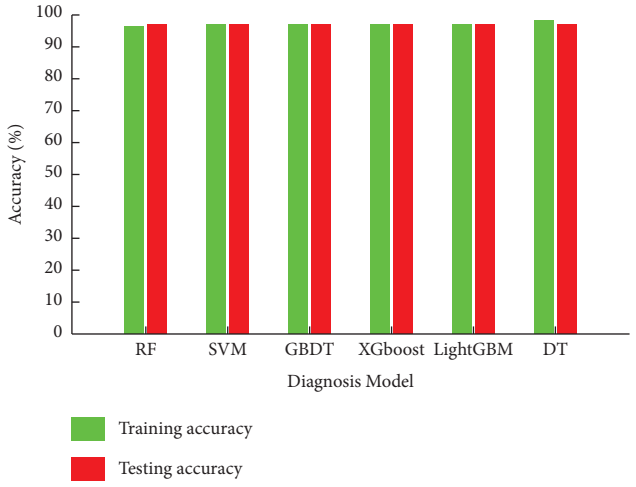

Figure 13: Comparison of fault diagnosis result under different diagnosis models.

categories, which are NS, IRF2, and ORF2 with a diameter of 0.3556 mm, ORF3 and REF3 with a diameter of 0.5334 mm. The rest of the five categories of precision are no less than 95%. The experimental results show that the proposed method can not only identify fault type but also analyze bearing fault degree.

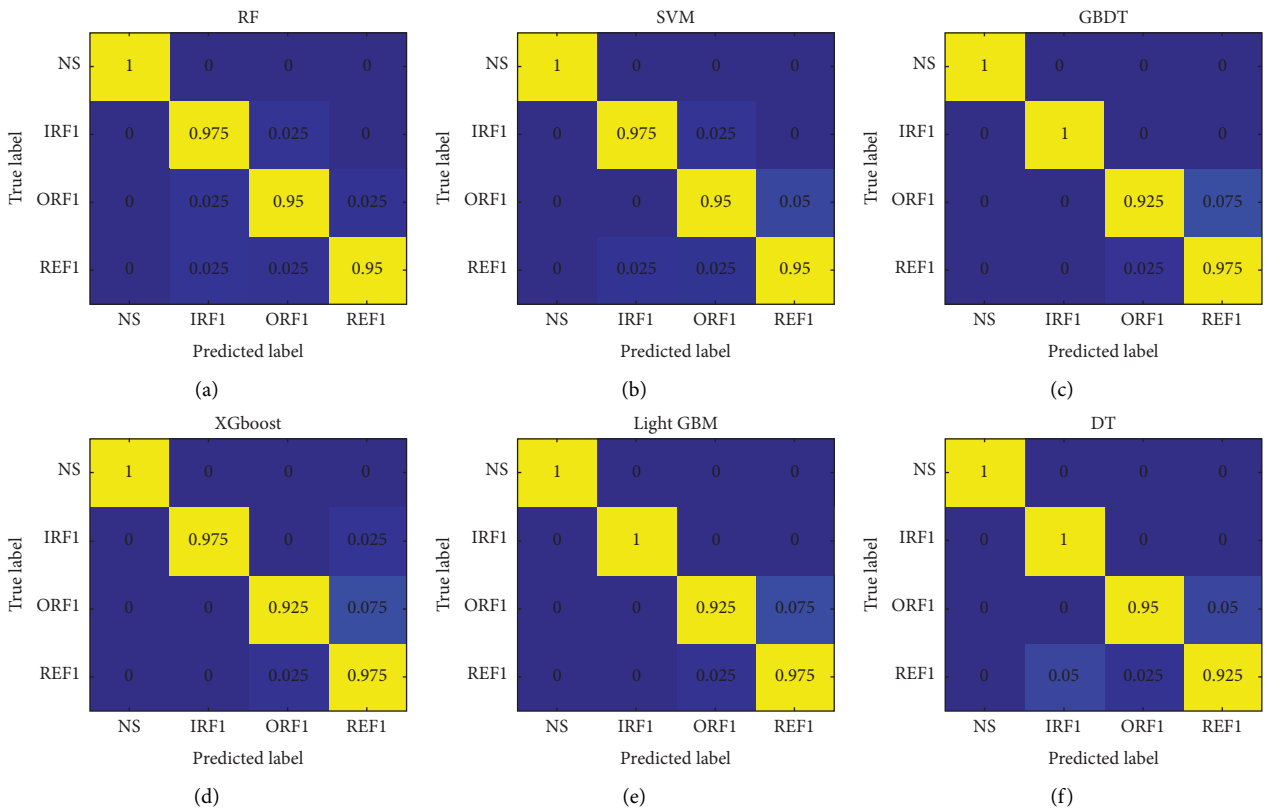

Figure 14: Fault diagnosis result under different number of training sample.

Table 4: Diagnosis accuracy of different diagnosis model.

Diagnosis model	Diagnosis accuracy
RF	97.5%
SVM	98%
GBDT	97.75
XGboost	97.25
LightGBM	97.25
DT	97.5

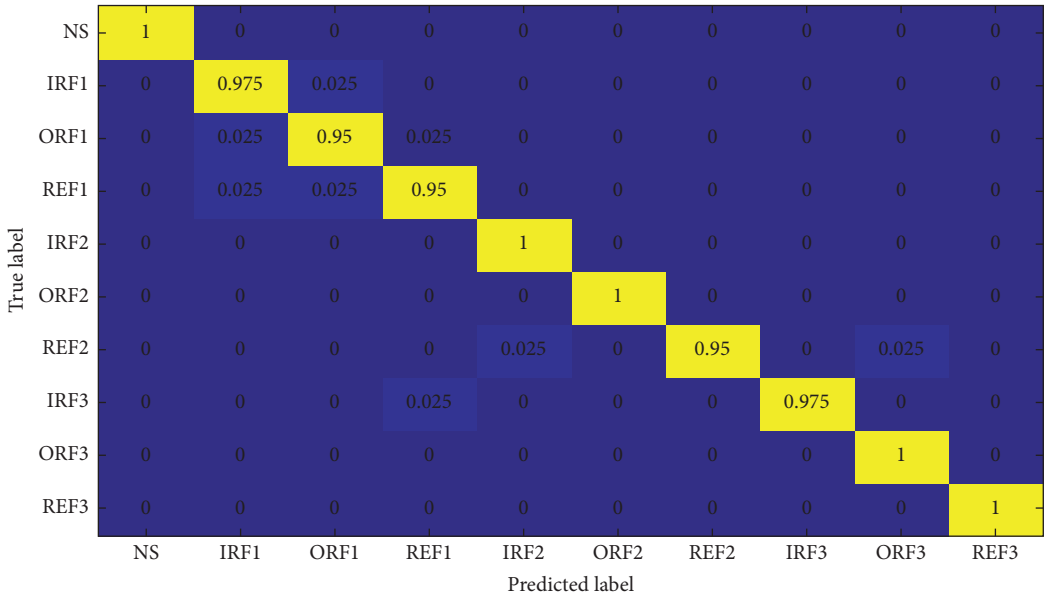

Figure 15: The confusion matrix of diagnosis result under ten classification faults.

4. Conclusion

In this paper, rolling bearing fault diagnosis method based on MCMF and SAIMFE is proposed. Three experiments are completed. The first experimental results show that r has great influence of diagnosis accuracy; the second experimental results show that the proposed denoising and feature extraction method has low dependency on the diagnosis model; the third experimental results show that the proposed method has high diagnosis accuracy. All the experimental results show that the proposed method is useful for improving fault accuracy.

Data Availability

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

Conflicts of Interest

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

Acknowledgments

This work was supported by the Key Program of Science and Technology Support of Tianjin (grant no. 18YFZCGX00930). Relay Projects of Key R&D Program Achievements Conversion of Tianjin (grant no. 18YFJLCG00060). The Key Technologies R&D Program of Tianjin (grant no. 18YFZNGX0).

References

[1] I. El-Thalji and E. Jantunen, "A summary of fault modelling and predictive health monitoring of rolling element bearings," *Mechanical Systems and Signal Processing*, vol. 60-61, pp. 252–272, 2015.

[2] Y. Qin, "A new family of model-based impulsive wavelets and their sparse representation for rolling bearing Fault Diagnosis," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 3, pp. 2716–2726, 2018.

[3] L. Cui, D. Mo, H. wang, and P. Chen, "Resonance-based nonlinear demodulation analysis method of rolling bearing fault," *Advances in Mechanical Engineering*, vol. 5, Article ID 420694, 2013.

[4] M. Xiao, K. Wen, C. Zhang, X. Zhao, W. Wei, and D. Wu, "Research on fault feature extraction method of rolling bearing based on NMD and wavelet threshold denoising," *Shock and Vibration*, vol. 2018, Article ID 9495265, 11 pages, 2018.

[5] G. Dong and J. Chen, "Noise resistant time frequency analysis and application in fault diagnosis of rolling element bearings," *Mechanical Systems and Signal Processing*, vol. 33, no. 11, pp. 212–236, 2012.

[6] H. Yuan, J. Chen, and G. Dong, "Machinery fault diagnosis based on time-frequency images and label consistent K-SVD," *Proceedings of the Institution of Mechanical Engineers - Part C: Journal of Mechanical Engineering Science*, vol. 232, no. 7, pp. 1317–1330, 2017.

[7] H. Helmi and A. Forouzantabar, "Rolling bearing fault detection of electric motor using time domain and frequency domain features extraction and ANFIS," *IET Electric Power Applications*, vol. 13, no. 5, pp. 662–669, 2019.

[8] L. Wan, K. Gong, G. Zhang, X. Yuan, C. Li, and X. Deng, "An efficient rolling bearing Fault Diagnosis method based on spark and improved random forest algorithm," *IEEE Access*, vol. 9, pp. 37866–37882, 2021.

[9] Y. Li, Y. Yang, X. Wang, B. Liu, and X. Liang, "Early fault diagnosis of rolling bearings based on hierarchical symbol dynamic entropy and binary tree support vector machine," *Journal of Sound and Vibration*, vol. 428, pp. 72–86, 2018.

[10] J. Hou, Y. Wu, H. Gong, A. S. Ahmad, and L. Liu, "A novel intelligent method for bearing Fault Diagnosis based on EEMD permutation entropy and GG clustering," *Applied Sciences*, vol. 10, no. 1, p. 386, 2020.

[11] C. He, T. Wu, R. Gu, Z. Jin, R. Ma, and H. Qu, "Rolling bearing fault diagnosis based on composite multiscale permutation entropy and reverse cognitive fruit fly optimization algorithm-Extreme learning machine," *Measurement*, vol. 173, Article ID 108636, 2021.

[12] L. Wan, H. Li, Y. Chen, and C. Li, "Rolling bearing fault prediction method based on QPSO-BP neural network and dempster–shafer evidence theory," *Energies*, vol. 13, no. 5, p. 1094, 2020.

[13] J. Xie, G. Du, C. Shen, N. Chen, L. Chen, and Z. Zhu, "An end-to-end model based on improved adaptive deep belief network and its application to bearing Fault Diagnosis," *IEEE Access*, vol. 6, pp. 63584–63596, 2018.

[14] H. Li, T. Liu, X. Wu, and Q. Chen, "A bearing fault diagnosis method based on enhanced singular value decomposition," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 5, pp. 3220–3230, 2021.

[15] R. Yan and R. X. Gao, "Approximate Entropy as a diagnostic tool for machine health monitoring," *Mechanical Systems and Signal Processing*, vol. 21, no. 2, pp. 824–839, 2007.

[16] F. Yang, Z. Kou, J. Wu, and T. Li, "Application of mutual information-sample entropy based MED-ICEEMDAN denoising scheme for weak Fault Diagnosis of hoist bearing," *Entropy*, vol. 20, no. 9, p. 667, 2018.

[17] H. B. Xie, W. X. He, and H. X. Liu, "Measuring time series regularity using nonlinear similarity-based sample entropy," *Physics Letters A*, vol. 372, no. 48, pp. 7140–7146, 2008.

[18] W. Chen, J. Zhuang, W. Yu, and Z. Wang, "Measuring complexity using FuzzyEn, ApEn, and SampEn," *Medical Engineering & Physics*, vol. 31, no. 1, pp. 61–68, 2009.

[19] L. Wang, H. Qiu, P. Yang, and L. Mu, "Arc Fault detection algorithm based on variational mode decomposition and improved multi-scale fuzzy entropy," *Energies*, vol. 14, no. 14, p. 4137, 2021.

[20] M. Costa, A. L. Goldberger, and C. K. Peng, "Multiscale entropy analysis of complex physiologic time series," *Physical Review Letters*, vol. 89, no. 6, Article ID 068102, 2002.

[21] J. Zheng, J. Cheng, Y. Yang, and S. Luo, "A rolling bearing fault diagnosis method based on multi-scale fuzzy entropy and variable predictive model-based class discrimination," *Mechanism and Machine Theory*, vol. 78, pp. 187–200, 2014.

[22] K. Feng, J. Ji, Y. Li, Q. Ni, H. Wu, and J. Zheng, "A novel cyclic-correntropy based indicator for gear wear monitoring," *Tribology International*, vol. 171, Article ID 107528, 2022.

[23] K. Feng, J. Ji, and Q. Ni, "A novel adaptive bandwidth selection method for Vold–Kalman filtering and its application in wind turbine planetary gearbox diagnostics," *Structural Health Monitoring*, Article ID 147592172210999, 2022.

[24] J. Yu, T. Hu, and H. Liu, "A new morphological filter for fault feature extraction of vibration signals," *IEEE Access*, vol. 7, pp. 53743–53753, 2019.
[25] R. Lerallut, E. Decenciere, and F. Meyer, "Image filtering using morphological amoebas," *Image and Vision Computing*, vol. 25, no. 4, pp. 395–404, 2007.
[26] G. Boato, D. T. Dang-Nguyen, and F. G. B. De Natale, "Morphological filter detector for image forensics applications," *IEEE Access*, vol. 8, pp. 13549–13560, 2020.
[27] B. Li, P. l. Zhang, Z. j. Wang, S. s. Mi, and D. s. Liu, "A weighted multi-scale morphological gradient filter for rolling element bearing fault detection," *ISA Transactions*, vol. 50, no. 4, pp. 599–608, 2011.
[28] W. Jiang, Z. Zheng, Y. Zhu, and Y. Li, "Demodulation for hydraulic pump fault signals based on local mean decomposition and improved adaptive multiscale morphology analysis," *Mechanical Systems and Signal Processing*, vol. 58-59, pp. 179–205, 2015.
[29] J. Wang, G. Xu, Q. Zhang, and L. Liang, "Application of improved morphological filter to the extraction of impulsive attenuation signals," *Mechanical Systems and Signal Processing*, vol. 23, no. 1, pp. 236–245, 2009.
[30] Y. Dong, M. Liao, X. Zhang, and F. Wang, "Faults diagnosis of rolling element bearings based on modified morphological method," *Mechanical Systems and Signal Processing*, vol. 25, no. 4, pp. 1276–1286, 2011.
[31] H. Zhao, J. Wang, H. Han, and Y. Gao, "A feature extraction method based on HLMD and MFE for bearing clearance fault of reciprocating compressor," *Measurement*, vol. 89, pp. 34–43, 2016.
[32] H. Zhao, M. Sun, W. Deng, and X. Yang, "A new feature extraction method based on EEMD and multi-scale fuzzy entropy for motor bearing," *Entropy*, vol. 19, no. 1, p. 14, 2016.
[33] "Case Western Reserve University Bearing Data Center Website," 2011, http://http//csegroups.case.edu/bearingdatacenter/pages/download-data-file.
[34] S. S. Roy, S. Dey, and S. Chatterjee, "Autocorrelation aided random forest classifier-based bearing fault detection framework," *IEEE Sensors Journal*, vol. 20, no. 18, pp. 10792–10800, 2020.
[35] A. Widodo and B. S. Yang, "Machine health prognostics using survival probability and support vector machine," *Expert Systems with Applications*, vol. 38, no. 7, pp. 8430–8437, 2011.
[36] H. Yang, W. Li, K. Hu, Y. Liang, and Y. Lv, "Deep ensemble learning with non-equivalent costs of fault severities for rolling bearing diagnostics," *Journal of Manufacturing Systems*, vol. 61, pp. 249–264, 2021.
[37] C. Xiang, Z. Ren, P. Shi, and H. Zhao, "Data-driven Fault Diagnosis for rolling bearing based on DIT-FFT and XGBoost," *Complexity*, vol. 2021, Article ID 4941966, 13 pages, 2021.
[38] D. Zhang and Y. Gong, "The comparison of LightGBM and XGBoost coupling factor analysis and prediagnosis of acute liver failure," *IEEE Access*, vol. 8, pp. 220990–221003, 2020.
[39] A. Krishnakumari, A. Elayaperumal, M. Saravanan, and C. Arvindan, "Fault diagnostics of spur gear using decision tree and fuzzy classifier," *International Journal of Advanced Manufacturing Technology*, vol. 89, no. 9-12, pp. 3487–3494, 2017.