Research Article

A New Bearing Fault Diagnosis Method Based on Refined Composite Multiscale Global Fuzzy Entropy and Self-Organizing Fuzzy Logic Classifier

Zhang Ziying 1,2 and Zhang Xi1

1 School of Mechanical Electronic & Information Engineering, China University of Mining and Technology-Beijing, Beijing 100083, China
2 Shanxi Institute of Energy, Taiyuan 030006, Shanxi, China

Correspondence should be addressed to Zhang Ziying; tbp1600401010@student.cumtb.edu.cn

Received 17 September 2020; Revised 10 October 2020; Accepted 2 December 2020; Published 2 September 2021

In this paper, a new feature extraction method called refined composite multiscale global fuzzy entropy (RCMGFE) is proposed. Based on the proposed RCMGFE and self-organizing fuzzy logic classifier (SOF), a new method for bearing fault diagnosis is proposed. Firstly, the fault features of the original bearing signal are extracted by using the proposed refined composite multiscale global fuzzy entropy, and the fault feature set of RCMGFE is constructed on this basis. Secondly, the extracted RCMGFE fault feature set is divided into an offline training sample set, an online training sample set, and a testing sample set. The offline training sample set and the online training sample set are, respectively, input into the offline training stage and the online training stage of the SOF for selecting representative samples and constructing fuzzy rules. Then, the testing sample set is input to the testing stage of the SOF for classification. Finally, the data of drive end bearing and fan end bearing provided by Case Western Reserve University are used to verify the validity of the proposed fault diagnosis method. The experimental results show that, compared with other methods, the proposed fault diagnosis method has a higher classification effect.

1. Introduction

Rotating machinery is an essential part of large and important equipment, which is applied in many engineering fields, such as electricity, chemical industry, and machinery manufacturing. The working conditions of rotating machinery directly affect the safety and stability of mechanical operations. At the same time, rolling bearings are not only a vital part of machinery and equipment but also an indispensable part of modern industrial processing enterprises. [1–4]. However, rolling bearings are very sensitive to the operating conditions of mechanical equipment. Bearings are one of the general parts prone to failure, and its failure accounts for a large proportion of mechanical equipment failures. Therefore, it is very important to carry out health management and fault diagnosis of bearings [5–8].

Traditional signal analysis includes time-domain analysis and spectrum analysis, and the signal of the bearing often shows nonstationarity during operation, and so traditional methods cannot well characterize the characteristics of this signal [9–13]. Therefore, scholars at home and abroad have developed time-frequency analysis methods, such as the variational mode decomposition (VMD) proposed in 2014, and the highly adaptive empirical modal decomposition (EMD) proposed by Huang et al. in 1988 and singular spectrum decomposition (SSD). Among them, VMD has high decomposition efficiency and fast decomposition speed, and there is no problem with modal aliasing. However, the effect of VMD has a lot to do with setting parameters, and it has no self-adaptability. In addition to the need to set parameters for EMD and SSD, there are also problems of modal aliasing and boundary effects [14–17]. Besides, due to
changes in operating conditions, analysis methods such as VMD cannot accurately find fault characteristics of non-linear signals, and so they cannot effectively reflect fault information [18–20]. The development of artificial intelligence technology has promoted the fault diagnosis discipline and its related research to enter the stage of intelligent diagnosis. Generally, the first step of intelligent diagnosis technology for bearing components is to obtain sufficient fault characteristic information from the collected raw signals. At the same time, since the entropy-based method can effectively measure the disorder of the signal, it can be used as an effective means to distinguish signals, and it has gradually become an important method for extracting fault characteristics, for example, sample entropy (SE), fuzzy entropy (FE), permutation entropy (PE), and dispersion entropy (DE) [21–23]. The common drawback of these methods is that only a single scale of signal irregularity is considered, which will lose a lot of fault information, so it may lead to misdiagnosis of the fault. In order to make up for such defects, researchers use multiscale coarse-graining operations to improve them and then apply them to signal feature extraction. The relevant research in recent years is as follows: Costa et al. [24] proposed multiscale sample entropy (MSE) and applied it to the evaluation of heart rate health in the medical field. Zhu et al. [25] used MFE in bearing fault diagnosis to extract fault characteristics. In 2005, Aziz and Arif [26] combined coarse-graining operation with permutation entropy to propose multiscale replacement entropy (MPE) and used it in the detection of congestive heart failure. In 2017, Azami proposed multiscale discrete entropy (MDE) and applied it to the determination of Senile dementia, which was subsequently used by many scholars in bearing fault diagnosis [27–29]. A large number of studies have proved that fuzzy entropy (FE) is advantageous in the computation of signal disorder. It obtains the initial vector by subtracting the local mean, which can only concentrate the local features of the signal and is not conducive to the accurate extraction of fault information. Therefore, global fuzzy entropy (GFE) is proposed [30,31]. Since GFE can only extract fault feature information from one scale, it will be difficult to accurately identify fault types if it is used to obtain the information contained in the signal directly. In order to make up for this defect in the calculation of GFE, we combine the coarse-grained processing technology with the global fuzzy entropy and propose the refine composite multiscale global fuzzy entropy (RCMGFE). The next step after fault feature extraction is to use the feature set as the input vector of the classifier for classification. As a research direction of machine learning, many classification algorithms are used in fault diagnosis. For example, kernel function extreme learning machine (KELM), least-square support vector machine (LSSVM), and extreme learning machine (ELM) [32].

The self-organizing fuzzy logic classifier (SOF) has low computational cost, transparent calculation process, easy to understand, and it does not need to set parameters in advance [33,34]. We propose a novel method of fault diagnosis based on RCMGFE and SOF. Firstly, we use RCMGFE to obtain signal features and then input the obtained feature set into the SOF classifier. Two benchmark data sets from Case Western Reserve University (CWRU) are used to verify the rationality of the proposed fault diagnosis method and to highlight the advantages of our proposed method by comparing it with other methods.

2. The Proposed RCMGFE and SOF

2.1. The Proposed Refined Composite Multiscale Global Fuzzy Entropy

2.1.1. Fuzzy Entropy. FE is a parameter that calculates the probability of generating a new model when the dimension changes [35]. It uses the fuzzy function $e^{-d_{ij}^m}$ to estimate the similarity of two vectors, and the calculation steps of FE are as follows:

1. Formula (1) is used to process the collected vibration signal \{u(i) \mid 1 \leq i \leq N\} to obtain an m-dimensional vector:
   \[ X_i^m = \{u(i), u(i+1), \ldots, u(i+m-1)\} - u_0(i), \]
   \[ \text{where} \ i = 1, 2, \ldots, N - m + 1, u_0(i) = \frac{1}{m} \sum_{j=0}^{m-1} u(i+j) \]
   (1)
   (2) The Chebyshev distance between $X_i^m$ and $X_j^m$ is calculated:
   \[ d_{ij}^m = \max_{k \in \{0,m-1\}} \left\{|(u(i+l) - u_0(i)) - (u(j+l) - u_0(j))|\right\}, \]
   \[ \text{where} \ i = 1, 2, \ldots, N - m + 1, i \neq j. \]
   (3) The similarity between $X_i^m$ and $X_j^m$ is calculated:
   \[ D_{ij}^m = \mu(d_{ij}^m, n, r) = e^{-d_{ij}^m/r^n}, \]
   \[ \text{where} \ n = 2 \text{ in general.} \]
   (4) The function $\phi^m(n,r)$ is defined:
   \[ \phi^m(n,r) = \frac{1}{N-m} \sum_{i=1}^{N-m} \left(1 - \frac{1}{N-m-1} \sum_{j=1}^{N-m} D_{ij}^m\right). \]
   (5) $m + 1$ vector $\phi^{m+1}(n,r)$ is constructed according to steps (1) to (4).
   (6) The fuzzy entropy is calculated:
   \[ \text{fuzzy entropy} (m,n,r,N) = \ln\left(\frac{\phi^m(n,r)}{\phi^{m+1}(n,r)}\right). \]

2.1.2. Refined Composite Multiscale Global Fuzzy Entropy. As an important method to calculate the disorder degree of signals, FE is suitable for fault feature extraction. However, it obtains the initial vector by subtracting the local mean, which only concentrates the local features, ignoring global characteristics of time series [36]. Therefore, some scholars
have proposed global fuzzy entropy (GFE) [37]. The operation of subtracting the local mean from the FE is changed to subtracting the mean of the whole series to calculate the global fuzzy entropy. In other words, \( u_0(i) \) in equation (1) is changed to \( 1/N \sum_{i=1}^{N} u(i) \), and then, the calculation of FE to get GFE is followed.

The overall characteristics of the bearing signal may change with its operating state, and the global fuzzy entropy was originally designed to characterize local fluctuations in a short period of time. Because the global fuzzy entropy can only calculate the characteristics from one scale, the fault information obtained is very simple, so it may not be ideal to directly use GFE to calculate the change of signal disorder degree. Based on the global fuzzy entropy, we propose a refined composite multiscale global fuzzy entropy (RCMGFE). Calculating the RCMGFE first requires multiscale coarse-graining operation. The formula for multiscale coarse-graining operation is as follows:

\[
    x_{k,j}^{(r)} = \frac{1}{\tau} \sum_{(i-1)\tau+1}^{\tau j} x_i, \quad 1 \leq j \leq \left\lfloor \frac{N}{\tau} \right\rfloor, \quad 1 \leq k \leq \tau,
\]

where \( \tau = 1, 2, \ldots, n \) refers to the scale factor, \( x_{k,j}^{(r)} \) is the \( k \)th coarse-graining series of \( x \).

After performing the coarse-graining operation, the RCMGFE value of each scale can be obtained by the step of calculating the global fuzzy entropy. For each scale of RCMGFE, its calculation can also be expressed by the following formula:

\[
    \text{RCMGFE}(\mathbf{x}, m, r, \tau) = -\ln \left( \frac{\pi_{m+1}^{n+1}}{\pi_{m,r}^{n+1}} \right).
\]

where \( \pi_{m+1}^{n+1} = (1/\tau) \sum_{k=1}^{\tau} a_{k,r}^{m+1} \) and \( \pi_{m,r}^{n+1} = (1/\tau) \sum_{k=1}^{\tau} a_{k,r}^{m} \), \( a_{k,r}^{m+1} \) and \( a_{k,r}^{m} \) represent the total number of pairs of \( m \) and \( m+1 \) dimension matching vectors.

Figure 1 is the calculation process of RCMGFE.

2.2. Self-Organizing Fuzzy Logic Classifier. SOF is a classifier with high efficiency. Its calculation has three processes [34]. The first two stages iteratively update the meta-parameters to construct fuzzy rules for each category. The testing samples are classified during the testing stage. Its specific process is as follows.

2.2.1. Offline Training Stage. The purpose of this process is to select representative samples (prototypes) in different kinds of samples and form fuzzy rules based on the selected prototypes. It is assumed that there are \( K \) samples in the sample set (the samples here refer to the multidimensional vector about RCMGFE), while the sample set in category \( c \) is expressed as \( \{x_i^c\} = \{x_{c,1}, x_{c,2}, x_{c,3}, \ldots, x_{c,C}^c\} \) and \( c \in \{1, 2, \ldots, C\} \). In both experiments in this paper, \( C = 4 \). The set containing all different samples and the number of repetitions of a sample in the set are \( \{u_i^c\} = \{u_{i,1}^c, u_{i,2}^c, \ldots, u_{i,U_k}^c\} \) and \( \{f_i^c\} = \{f_{i,1}^c, f_{i,2}^c, \ldots, f_{i,U_k}^c\} \), respectively. \( K^c \) is the number of elements in \( \{x_i^c\}_{K^c} \). So, \( \sum_{c=1}^{C} K^c = \sum_{c=1}^{C} U_k^c = U_k \), \( U_k \) is a unique sample set for all categories. It should be noted that the search for prototypes is done in the same category. The specific process of selecting a prototype is as follows:

1. The multimodal density of the sample is calculated using formula (8), where \( d \) represents the straight-line distance between two samples:

\[
    D_{K^c}^{MM}\left(u_i^c\right) = \frac{\sum_{j=1}^{K^c} d^2 \left(x_j^c, x_i^c\right)}{2K^c \sum_{j=1}^{K^c} d^2 \left(u_j^c, x_i^c\right)}, \quad i = 1, 2, \ldots, U_k^c.
\]

(8)

2. The samples were sorted based on the calculated \( D_{K^c}^{MM} \) and mutual distance. The set after sorting processing is \( \{r\} = \{r_1, r_2, \ldots, r_{U_k^c}\} \), where \( r_1 = \text{arg max} \{D_{K^c}^{MM}\left(u_i^c\right)\} \). \( r_2 \) is the closest sample to \( r_1 \), \( r_3 \) is the closest sample to \( r_2 \), and then, this rule is followed. The multimodal density set of the set \( \{r\} \) is denoted as \( \{DMM_{K^c}(r)\} \). After sorting the multimodal density, the initial prototype is selected by formula (10). If the \( D_{K^c}^{MM} \) of any sample in the sorted sample set \( \{r\} \) is greater than the multimodal density of two adjacent samples, it will be used as an initial prototype. What needs to be explained here is that \( r_1 \) is definitely an initial prototype. And the collection containing the initial prototype is \( \{p\}_0 \).

3. The initial prototype attracts nearby samples to build the data cloud.

4. The \( \{p\}_0 \) obtained in step (2) is defined as \( \{\varphi\}_0 \), and \( D_{K^c}^{MM} \) is recalculated through the following formula:

\[
    D_{K^c}^{MM}\left(\varphi_i\right) = \frac{S \sum_{j=1}^{n} d^2 \left(\varphi_i, \varphi_j\right)}{2K^c \sum_{j=1}^{n} d^2 \left(\varphi_i, \varphi_j\right)}, \quad i = 1, 2, \ldots, U_k^c;
\]

where \( \varphi_i \in \{\varphi\}_0 \). \( S \) is the total number of elements in the \( i \)th data cloud constructed in the above steps, and \( n \) is the sum of the number of samples contained in all data clouds \( \{\varphi\}_0 \).

5. If the square of the distance between the sample that does not meet the initial prototype condition in step (2) and an initial prototype is not greater than the average radius \( G_{K^c}^L \) of the affected area of the sample related to the granularity, this sample is taken as the adjacent center of the initial prototype and this rule construct adjacent center set \( \{\varphi\}_0^\text{neighbor} \) is followed. In this paper, we set the granularity as \( L = 12 \), and its calculation process is shown in following formula:

\[
    G_{K^c}^{L} = \frac{\sum d^2 (x, y)}{Q_{K^c}^{L}}, \quad (10)
\]
where \( x \) and \( y \) are different samples in the same category, the square of the distance between \( x \) and \( y \) should not be higher than \( G_{Kc,L}^{L-1} \). \( G_{Kc,L}^{L-1} \) is the average radius corresponding to \( L-1 \). \( D_{Kc}^{c} \) is total number of which the distance between any two samples in the sample set is less than \( G_{Kc,L}^{L-1} \). The total number of sample pairs with a distance less than \( d_{Kc}^{-} \)in the set \( \{x\}^{c}_{Kc} \) is denoted as \( Q_{Kc}^{c} \).

\[
\bar{d}_{Kc}^{-} = \frac{1}{(Kc)^2} \sum_{i=1}^{Kc} \sum_{j=1}^{Kc} d^2(x_i^c, x_j^c).
\]

(6) If the multimode density of the initial prototype is higher than the multimode density of the sample attracted by the prototype, the initial prototype is regarded as a final prototype, and the final prototype set is defined as \( \{p\}^c \).

\[
\text{if} \left( D_{Kc}^{c}(\varphi) > D_{Kc}^{c}(\varphi') \right), \quad \text{then} \quad \varphi_i \in \{p\}^c,
\]

where \( \varphi \in \{\varphi\}^c_{\text{neighbor}} \).

(7) After step (6), fuzzy rules is constructed according to the following formula. If a sample is similar to any prototype in a certain category, the sample is divided into this category.

\[
\text{if} \left( x \sim p_i^c \right) \text{or} \left( x \sim p_j^c \right) \ldots \text{or} \left( x \sim p_N^c \right), \quad \text{then} \quad x \in \text{(category c)},
\]

where \( \sim \) stands for similarity, \( N^c \) is the number of \( c \)-type samples that satisfy the prototype discrimination condition in formula (12); in other words, \( \sum_{c=1}^{C} N^c \) is the total number of all prototypes.

2.2.2. Online Training Stage. In order to further select prototypes, the online training stage is proposed. As an important stage of SOF classifier, the process is established on the basis of treating samples as streaming data. If the newly input sample of category \( x_i \) is \( x_{Kc+1}^c \), the new sample set is \( \{x\}^{c}_{Kc+1} \). The steps in the online training stage are as follows:

(1) Based on the consideration of reducing the calculation cost, \( G_{Kc}^{c} \) in formula (10) will be calculated according to following formula:

\[
G_{Kc+1}^{c} = \frac{d_{Kc+1}^{-}}{d_{Kc}^{-}} G_{Kc}^{c}. \quad (14)
\]

(2) \( x_{Kc+1}^c \) is defined as a prototype if the multimode density of \( x_{Kc+1}^c \) is greater than the maximum multimode density of the prototype in the same class or less than the minimum multimode density of the prototype in the same class. That is, if \( D_{Kc+1}^{c}(x_{Kc+1}^c) \) is a maximum or minimum, then \( x_{Kc+1}^c \) is a prototype.

(3) If \( x_{Kc+1}^c \) does not satisfy the judgment condition in step (2), another condition can be used to determine whether \( x_{Kc+1}^c \) is a prototype. If the square of the minimum Euclidean distance between the newly input sample \( x_{Kc+1}^c \) and prototype is greater than \( G_{Kc+1}^{c} \), we still regard \( x_{Kc+1}^c \) as a prototype.

(4) If the new input sample \( x_{Kc+1}^c \) satisfies any one of the conditions of step (2) or step (3), the meta-parameters are calculated according to the following formula:

\[
N^c + 1 \rightarrow N^c, \\
\text{if} \left( x_{Kc+1}^c \rightarrow p_{N^c}^c \right), \\
1 \rightarrow S_{N^c}^c, \\
\{p\}^c + p_{N^c}^c \rightarrow \{p\}^c.
\]

If neither step is satisfied, then \( p_{N^c}^c = \arg \min (d(x_{Kc+1}^c, p)) \). And the corresponding meta-parameters are updated according to formula (17):
of SOF classifies the input testing samples and outputs the corresponding category labels.

4. Experiments

4.1. Experiment 1. In order to test the classification effect of the proposed method, the benchmark data of the drive end collected by CWRU was used in Experiment 1. The motor runs 1797 revolutions per minute. The number of acceleration signals obtained per second is 12000, and the signal contains a normal working state and three fault working states. The three fault states are ball fault, inner ring fault, and outer ring fault. There are three different fault sizes for each fault state, 0.007 inches and its double size and its triple size, respectively. The first 102,400 points of rolling bearing data in each state are evenly divided into 50 samples without the same data points, that is, each sample contains 2048 data points. It can be concluded that there are 50 samples in the normal state, 150 samples in each of the three failure states. The experiment was performed on a computer equipped with a 2.4 GHz CPU and 4 GB RAM, and it was run in the MATLAB 2019a version. Table 1 statistics the details of the bearing’s working status and corresponding labels. Figure 3 plots the time-domain waveforms corresponding to various state signals. We can know that the amplitude of the signal under normal working conditions is the smallest, which is easy to distinguish from the signal in a fault state. We can also see that it is not easy to clearly identify these three fault states from the waveform. The RCMGFE diagram corresponding to Figure 3 is given in Figure 4. We can see that, in the first 13 scales, the entropy value of the outer ring faulty bearing is the smallest. This is because, the vibration of this kind of bearing is very regular, and its complexity is relatively small, so its entropy value is relatively low.

After calculating RCMGFE, a 500×25 fault feature set was formed. This set will be divided into three different sets. Firstly, 30% samples are randomly selected in each state to use in the SOF offline training phase. Then, according to the above steps, another 30% of the total sample is selected from the remaining samples for online training. Finally, all other samples are used in the test phase of SOF. In general, the offline training sample set contains 150 samples, including 15 samples in normal state and 45 samples in each failure state. The online training sample set used in the SOF online training stage has a total of 150 samples, of which there are 15 samples in the normal state and 45 samples in each of the three fault states. So there are 200 test samples in the experiment, including 20 samples in the normal state and 60 samples in each of the three fault states. In this paper, by changing the input of the classifier and using different classification methods, they are compared with the methods proposed, and the advantages of our proposed method are quantitatively demonstrated through common indicators. Different classification inputs are multiscale fuzzy entropy (MFE), multiscale global fuzzy entropy (MGFE), and refined composite multiscale fuzzy entropy (RCMFE). The parameter settings of these three fault extraction methods are the same as those of RCMGFE. The classification methods used for the comparison are K-nearest neighbor (KNN),
ELM, LSSVM, and KELM. The parameter for all the classification methods mentioned above for comparison is the same as those in reference 34 [28,29,34,38]. Based on the perspective of reducing the influence of accidental factors on the experiment and improving persuasiveness, we did ten experiments. According to the four indicators shown in Figure 2, the fault diagnosis results of various methods were evaluated.

As can be seen from Table 2, in these 10 experiments, the fault diagnosis method of RCMF-SOF has the most stable classification performance among all methods. Contrary to RCMF-SOF, it has the largest standard deviation (Std) in

### Table 1: Working states and label.

| Working state        | Fault size (inch) | Code name | Category label | Number of samples |
|----------------------|-------------------|-----------|----------------|-------------------|
| Normal               | 0                 | Normal    | 1              | 50                |
|                      | 0.007             |           |                |                   |
|                      | 0.014             |           |                |                   |
| Inner ring failure   | 0.021             | IR        | 2              | 150               |
|                      | 0.007             |           |                |                   |
|                      | 0.014             |           |                |                   |
| Outer ring failure   | 0.021             | OR        | 3              | 150               |
|                      | 0.007             |           |                |                   |
|                      | 0.014             |           |                |                   |
| Ball failure         | 0.021             | B         | 4              | 150               |

Figure 2: Flowchart of the proposed fault diagnosis method.
the classification results of MFE-SOF. From the maximum and minimum values of classification accuracy, MFE-SOF has the worst classification effect. According to the analysis of the results of the first four-fault diagnosis methods, when the classification method is SOF, using RCMGFE to extract fault features can obtain higher classification accuracy, that is to say, the proposed RCMGFE is more suitable for extracting fault features. The maximum classification accuracy and minimum classification accuracy of RCMGFE-KNN and RCMGFE-ELM are the same, which are 98.5% and 95%, respectively. The maximum classification accuracy of RCMGFE-LSSVM and RCMGFE-KELM simultaneously reached 99%. It can also be seen that the minimum value of classification accuracy of the proposed method is 96.5%, which is 0.5% lower than that of RCMFE-SOF. Moreover, from the perspective of standard deviation, RCMGF-SOF is not as stable as RCMFE-SOF. However, the maximum classification accuracy of RCMFE-SOF proposed in this paper has reached 100%, and the average value of its classification accuracy has reached 98.4%, which is satisfactory. Based on the point of displaying the experimental results more intuitively, the confusion matrix of the results of various fault diagnosis methods used in the third experiment is given in Figure 5. The horizontal coordinate is the prediction label of a certain category of samples, the vertical coordinate is the real label of samples, and the number in the grid indicates the proportion of the prediction label that is the actual label or the wrong label. The classification result of MFE-SOF is shown in Figure 5(a). A total of eight samples in label 3 and label 4 were misclassified, and its total accuracy was 96%. The experimental results of the method named MGFE-SOF are shown in Figure 5(b). There are misclassified...
Figure 5: Continued.
samples in labels 3 and 4. Two of the samples labeled 3 were incorrectly classified as category 4, and two of the samples labeled 4 were incorrectly classified as category 3. It can be concluded that the overall classification accuracy of MGFE-SOF has reached 98%. The classification result of RCMFE-SOF is shown in Figure 5(c). The sample with misclassification is the sample with label 2, and three of the samples are misclassified into the third category shown in Figure 5(d) is the classification result of the proposed method. Each sample in the set is identified, and the classification accuracy rate reaches 100%. The classification result of RCMGFE-KNN is shown in Figure 5(e). Among the samples labeled 4, two samples were erroneously classified into the third category, one sample was erroneously classified into the second category, and the classification accuracy of the fourth category sample reached 95%. The classification result of RCMGFE-ELM is shown in Figure 5(f). Among the samples labeled 3 and 4, one and two samples were misclassified, respectively. The overall classification accuracy of the RCMGFE-ELM reached 96%. The classification result of RCMGFE-LSSVM is shown in Figure 5(g). Among the samples labeled 3 and 4, one and two samples were misclassified, respectively, and the total sample recognition rate is 98.5%. The classification result of RCMGFE-KELM is shown in Figure 5(h). Two of the samples with label 4 were incorrectly classified as the third category. The classification accuracy of the fourth category reached 97%, and its overall classification accuracy reached 99.

4.2. Experiment 2. The fan end data provided by CWRU is another commonly used benchmark set, which is regarded by many scholars as a more challenging database [23]. We used this data set in Experiment 2. It should be noted that all the parameter settings of this experiment are not different from the parameter settings in Section 4.1. Table 3 provides the final classification effect of each method. From the perspective of standard deviation, the stability of the proposed method is not as good as the stability of RCMGFE-KNN, RCMGFE-ELM, and RCMGFE-KELM. However, no matter the maximum value, minimum value, or average value of classification recognition rate, our method has the best classification effect. Besides, our method achieved a classification recognition rate of up to 99.5% in these ten experiments; this value is 1.5% greater than the maximum classification accuracy of RCMGFE-KNN, RCMGFE-ELM, and RCMGFE-KELM. From the minimum classification accuracy, the minimum classification accuracy of the proposed method reaches 97.5%, which is 1.5% higher than the minimum classification accuracy of RCMGFE-KNN, RCMGFE-ELM, and RCMGFE-KELM. The average classification accuracy of the proposed method reached 98.4%. All in all, the result of Experiment 2 is satisfactory.

5. Conclusion

Based on RCMGFE and SOF, we propose a new method for fault diagnosis of rolling bearing components. First, RCMGFE is calculated according to the acceleration signal; on this basis, the overall feature set is formed. Then, according to the three stages of SOF’s calculation process, the above overall feature set is randomly assigned to three corresponding feature sample sets. Finally, the three sets are input into the corresponding SOF process to obtain the final classification results. Two bearing data sets of CWRU were

| Method          | Maximum | Minimum | Average | Std  |
|-----------------|---------|---------|---------|------|
| MFE-SOF         | 93.5    | 89      | 91.3    | 1.45 |
| MGFE-SOF        | 98.5    | 95      | 96.55   | 1.13 |
| RCMFE-SOF       | 98.5    | 94      | 96.15   | 1.32 |
| Proposed method | 99.5    | 97.5    | 98.4    | 0.7  |
| RCMGFE-KNN      | 98      | 96      | 97.05   | 0.65 |
| RCMGFE-ELM      | 98      | 96      | 96.85   | 0.67 |
| RCMGFE-LSSVM    | 99      | 96.5    | 97.9    | 0.73 |
| RCMGFE-KELM     | 98      | 96      | 97.15   | 0.67 |
used to verify the rationality of the proposed fault diagnosis method. Through comparative analysis with other methods, we draw the following conclusions:

1. The final classification result obtained by changing the input of SOF indicates the proposed RCMGFE can obtain the fault information of bearing more comprehensively than MFE, MGFE, and RCMFE. In other words, RCMGFE is an effective way to measure signal complexity.

2. Compared with different classification methods, self-organizing fuzzy logic classifier is a more efficient classifier, which is very suitable for fault feature classification.

3. In the two experiments we have done, the maximum classification recognition rate of the proposed method is 100% and 99.5%, respectively, which makes its effectiveness and versatility reflected.

Data Availability

The data used in the manuscript can be downloaded from the open-source database: https://csegroups.case.edu/bearingdatacenter/pages/welcome-case-western-reserve-university-bearing-data-center-website.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors’ Contributions

Xi Zhang and performed the experiments; Ziying Zhang wrote the paper.

Acknowledgments

This research was funded by the Shanxi Natural Science Foundation (201901D111298).

Supplementary Materials

The data files DE97.txt, DE105.txt, DE118.txt, and DE130.txt correspond to Figure 4 in the paper; the data files DE118.txt, DE130.txt, DE169.txt, DE185.txt, DE197.txt, DE209.txt, DE234.txt and DE222.txt correspond to Figure 5 in the paper. (Supplementary Materials)

References

[1] C. Shen, J. Y. Qi, and Z. G. ZhuCai, “An automatic and robust features learning method for rotating machinery fault diagnosis based on contractive autoencoder,” Engineering Applications of Artificial Intelligence, vol. 76, pp. 170–184, 2018.

[2] X. Jiang, J. Wang, J. Shi, C. Shen, W. Huang, and Z. Zhu, “A coarse-to-fine decomposing strategy of VMD for extraction of weak repetitive transients in fault diagnosis of rotating machines,” Mechanical Systems and Signal Processing, vol. 116, pp. 668–692, 2019.

[3] J. Liu and Y. Shao, “Dynamic modeling for rigid rotor bearing systems with a localized defect considering additional deformations at the sharp edges,” Journal of Sound and Vibration, vol. 398, pp. 84–102, 2017.

[4] Y. Lei, J. Lin, Z. He, and M. J. Zuo, “A review on empirical mode decomposition in fault diagnosis of rotating machinery,” Mechanical Systems and Signal Processing, vol. 35, no. 1-2, pp. 108–126, 2013.

[5] Z. Wang, J. Zhou, J. Wang et al., “A novel fault diagnosis method of gearbox based on maximum kurtosis spectral entropy deconvolution,” IEEE Access, vol. 7, pp. 29520–29532, 2019.

[6] L. Wang, Z. Liu, H. Cao, and X. Zhang, “Subband averaging kurtogram with dual-tree complex wavelet packet transform for rotating machinery fault diagnosis,” Mechanical Systems and Signal Processing, vol. 142, Article ID 106755, 2020.

[7] Y. Zhang, Q. Gao, Y. Lu, D. Sun, Y. Xia, and X. Peng, “A novel intelligent method for bearing fault diagnosis based on hermitian scale-energy spectrum,” IEEE Sensors Journal, vol. 18, no. 16, pp. 6743–6755, 2018.

[8] Y. Wang, P. W. Tse, B. Tang et al., “Order spectrogram visualization for rolling bearing fault detection under speed variation conditions,” Mechanical Systems and Signal Processing, vol. 122, pp. 580–596, 2019.

[9] X. Jiang, C. Shen, J. Shi, and Z. Zhu, “Initial center frequency-guided VMD for fault diagnosis of rotating machines,” Journal of Sound and Vibration, vol. 435, pp. 36–55, 2018.

[10] Y. Li, G. Li, Y. Yang, X. Liang, and M. Xu, “A fault diagnosis scheme for planetary gearboxes using adaptive multi-scale morphology filter and modified hierarchical permutation entropy,” Mechanical Systems and Signal Processing, vol. 105, pp. 319–337, 2018.

[11] Y. Miao, M. Zhao, J. Lin, and Y. Lei, “Application of an improved maximum correlated kurtosis deconvolution method for fault diagnosis of rolling element bearings,” Mechanical Systems and Signal Processing, vol. 92, pp. 173–195, 2017.

[12] H. Wang, W. Song, E. Zio, A. Kudreisky, and Y. Zhang, “Remaining useful life prediction for Lithium-ion batteries using fractional Brownian motion and Fruit-fly Optimization Algorithm,” Measurement, vol. 161, Article ID 107904, 2020.

[13] W. Song, X. X. Chen, and C. Cattani, “Multifractal brownian motion and quantum-behaved partial swarm optimization for bearing degradation forecasting,” Complexity, vol. 2020, Article ID 8543131, 9 pages, 2020.

[14] Y. Gao, F. Villecco, M. Li, and W. Song, “Multi-scale permutation entropy based on improved LMD and HMM for rolling bearing diagnosis,” Entropy, vol. 19, no. 4, p. 176, 2017.

[15] W. Deng, H. Liu, S. Zhang, H. Liu, H. Zhao, and J. Wu, “Research on an adaptive variational mode decomposition with double thresholds for feature extraction,” Symmetry, vol. 10, no. 12, p. 684, 2018.

[16] Y. Ku, J. Cao, and J. Zhao, “Application of fast singular spectrum decomposition method based on order statistic filter in rolling bearing fault diagnosis,” Measurement Science and Technology, vol. 30, p. 125001, 2019.

[17] H. Li, T. Bao, C. Gu, and B. Chen, “Vibration feature extraction based on the improved variational mode decomposition and singular spectrum analysis combination algorithm,” Advances in Structural Engineering, vol. 22, no. 7, pp. 1519–1530, 2019.

[18] C. Li, Y. Tao, W. Ao, S. Yang, and Y. Bai, “Improving forecasting accuracy of daily enterprise electricity consumption using a random forest based on ensemble empirical mode decomposition,” Energy, vol. 165, pp. 1220–1227, 2018.

[19] K. D. Seger, M. H. Al-Badrawi, J. L. Miksis-Olds, N. J. Kirsch, and A. P. Lyons, “An empirical mode decomposition-based
detection and classification approach for marine mammal vocal signals,” *The Journal of the Acoustical Society of America*, vol. 144, no. 6, pp. 3181–3190, 2018.

[20] Y. Tian, J. Ma, C. Lu, and Z. Wang, “Rolling bearing fault diagnosis under variable conditions using LMD-SVD and extreme learning machine,” *Mechanism and Machine Theory*, vol. 90, pp. 175–186, 2015.

[21] 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.

[22] L. Zhang, G. Xiong, H. Liu, H. Zou, and W. Guo, “Bearing fault diagnosis using multi-scale entropy and adaptive neuro-fuzzy inference,” *Expert Systems with Applications*, vol. 37, no. 8, pp. 6077–6085, 2010.

[23] N. Rodriguez, P. Alvarez, L. Barba, and G. Cabrera-Guerrero, “Combining multi-scale wavelet entropy and kernelized classification for bearing multi-fault diagnosis,” *Entropy*, vol. 21, no. 2, p. 152, 2019.

[24] M. Costa, A. L. Goldberger, and C. K. Peng, “Multiscale entropy analysis of complex physiologic time series,” *Physical Review Letters*, vol. 89, Article ID 068102, 2002.

[25] K. H. Zhu, L. Chen, and X. Hu, “A multi-scale fuzzy measure entropy and infinite feature selection based approach for rolling bearing fault diagnosis,” *Journal of Nondestructive Evaluation*, vol. 38, p. 90, 2019.

[26] W. Aziz and M. Arif, “Multiscale permutation entropy of physiological time series,” in *Proceedings of the 9th International Multitopic Conference*, Karachi, Pakistan, December 2005.

[27] H. Azami and E. Kinney-Lang, “Multiscale dispersion entropy for the regional analysis of resting-state magnetoencephalogram complexity in alzheimer’s disease,” in *Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 3182–3185, Jeju Island, South Korea, July 2017.

[28] L. Zheng, Z. Wang, Z. Zhao, J. Wang, and W. Du, “Research of bearing fault diagnosis method based on multi-layer extreme learning machine optimized by novel ant lion algorithm,” *IEEE Access*, vol. 7, pp. 89845–89856, 2019.

[29] X. Yan and M. Jia, “Intelligent fault diagnosis of rotating machinery using improved multiscale dispersion entropy and mRMR feature selection,” *Knowledge-Based Systems*, vol. 163, pp. 450–471, 2019.

[30] G. Li, Q. Guan, and H. Yang, “Noise reduction method of underwater acoustic signals based on CEEMDAN, effort-to-compress complexity, refined composite multiscale dispersion entropy and wavelet threshold denoising,” *Entropy*, vol. 21, no. 1, p. 11, 2018.

[31] W. B. Zhang and J. Z. Zhou, “A comprehensive fault diagnosis method for rolling bearings based on refined composite multiscale dispersion entropy and fast ensemble empirical mode decomposition,” *Entropy*, vol. 21, no. 7, p. 680, 2019.

[32] Q. W. Gao, W. Y. Liu, B. P. Tang, and G. J. Li, “A novel wind turbine fault diagnosis method based on intergral extension load mean decomposition multiscale entropy and least squares support vector machine,” *Renewable Energy*, vol. 116, pp. 169–175, 2018.

[33] X. Gu and P. P. Angelov, “Self-organizing fuzzy logic classifier,” *Information Sciences*, vol. 447, pp. 36–51, 2018.

[34] W. Du, X. Guo, Z. Wang et al., “A new fuzzy logic classifier based on multiscale permutation entropy and its application in bearing fault diagnosis,” *Entropy*, vol. 22, no. 1, p. 27, 2019.

[35] C. Liu, K. Li, L. Zhao et al., “Analysis of heart rate variability using fuzzy measure entropy,” *Computers in Biology and Medicine*, vol. 43, no. 2, pp. 100–108, 2013.

[36] K. Zhu, L. Chen, and X. Hu, “Rolling element bearing fault diagnosis based on multi-scale global fuzzy entropy, multiple class feature selection and support vector machine,” *Transactions of the Institute of Measurement and Control*, vol. 41, no. 14, pp. 4013–4022, 2019.

[37] J. Zheng, Z. Jiang, and H. Pan, “Sigmoid-based refined composite multiscale fuzzy entropy and t-SNE based fault diagnosis approach for rolling bearing,” *Measurement*, vol. 129, pp. 332–342, 2018.

[38] X. J. Gao and H. F. Wei, “A rolling bearing fault diagnosis method based on LSSVM,” *Advances in Mechanical Engineering*, vol. 12, Article ID 1687814019899561, 2020.