

## Article

# Intelligent Diagnosis of Rolling Element Bearing Based on Refined Composite Multiscale Reverse Dispersion Entropy and Random Forest

Aiqiang Liu <sup>1</sup>, Zuye Yang <sup>2</sup>, Hongkun Li <sup>1,\*</sup>, Chaoge Wang <sup>3</sup> and Xuejun Liu <sup>1</sup>

<sup>1</sup> School of Mechanical Engineering, Dalian University of Technology, Dalian 116024, China; liuaq@mail.dlut.edu.cn (A.L.); liuxuejun@mail.dlut.edu.cn (X.L.)

<sup>2</sup> Microcyber Corporation, Shenyang 110179, China; yang.zuye@microcyber.cn

<sup>3</sup> School of Logistics Engineering, Shanghai Maritime University, Shanghai 201306, China; cgwang@shmtu.edu.cn

\* Correspondence: lihk@dlut.edu.cn

**Abstract:** Rolling bearings are the vital components of large electromechanical equipment, thus it is of great significance to develop intelligent fault diagnoses for them to improve equipment operation reliability. In this paper, a fault diagnosis method based on refined composite multiscale reverse dispersion entropy (RCMRDE) and random forest is developed. Firstly, rolling bearing vibration signals are adaptively decomposed by variational mode decomposition (VMD), and then the RCMRDE values of 25 scales are calculated for original signal and each decomposed component as the initial feature set. Secondly, based on the joint mutual information maximization (JMIM) algorithm, the top 15 sensitive features are selected as a new feature set and feed into random forest model to identify bearing health status. Finally, to verify the effectiveness and superiority of the presented method, actual data acquisition and analysis are performed on the bearing fault diagnosis experimental platform. These results indicate that the presented method can precisely diagnose bearing fault types and damage degree, and the average identification accuracy rate is 97.33%. Compared with the refined composite multiscale dispersion entropy (RCMDE) and multiscale dispersion entropy (MDE), the fault diagnosis accuracy is improved by 2.67% and 8.67%, respectively. Furthermore, compared with the RCMRDE method without VMD decomposition, the fault diagnosis accuracy is improved by 3.67%. Research results prove that a better feature extraction technique is proposed, which can effectively overcome the deficiency of existing entropy and significantly enhance the ability of fault identification.

**Keywords:** refined composite multiscale reverse dispersion entropy; random forest model; rolling bearing; intelligent diagnosis



**Citation:** Liu, A.; Yang, Z.; Li, H.; Wang, C.; Liu, X. Intelligent Diagnosis of Rolling Element Bearing Based on Refined Composite Multiscale Reverse Dispersion Entropy and Random Forest. *Sensors* **2022**, *22*, 2046. <https://doi.org/10.3390/s22052046>

Academic Editors: Steven Chatterton and Lang Xu

Received: 7 February 2022

Accepted: 3 March 2022

Published: 6 March 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

As an important component of large-scale electromechanical equipment, the rolling bearing health status is critical for the stable operation of equipment. Unfortunately, the rolling bearings are vulnerable components due to large load and unstable operation. Therefore, the intelligent diagnosis of rolling bearing is essential to improve the safety and stability of equipment, reduce maintenance costs, and avoid safety accidents [1–3].

The equipment will inevitably produce vibration during production operation, and the vibrations are the reflection of dynamic equipment characteristics. Therefore, the vibration signals contain a large amount of information characterizing the equipment's health status, which has a clear physical meaning and can be easily identified [4]. At present, scholars have proposed various time-frequency analysis methods based on vibration signals, which can also be used to decompose signals. Common time-frequency analysis methods include the wavelet transform (WT) [5], resonance-based sparse signal

decomposition (RSSD) [6], Hilbert vibration decomposition (HVD) [7], and empirical mode decomposition (EMD) [8], etc. However, the WT method needs to select wavelet base and decomposition layer beforehand. The EMD is prone to mode overlap and endpoint effects, and the RSSD method requires determining appropriate parameters in advance. The HVD is sensitive to additive noise and is not applicable for analyzing intermittent or non-oscillating signals, which makes them unable to obtain good application results [9–11]. To address these issues, Dragomiretskiy [12] presented the variational mode decomposition (VMD), which can achieve a perfect decomposition of non-linear, non-stationary signals. VMD has a solid theoretical background and analytical formulation has strong robustness and other advantages and has been shown to outperform EMD for intra-wave signal analysis [13,14]. Subsequently, VMD is widely used for rotating machinery fault diagnosis. In [15], VMD and energy entropy are combined for rolling bearing fault detection, which has a better effect than EMD and WT. In [16], the combination of VMD and Fourier synchro squeezed transform is used to extract bearing defect characteristics. In addition, VMD is also combined with deep learning methods such as convolution neural network [17], long short-term memory network [18], deep belief network [19] and plays a crucial role in fault diagnosis and remaining useful life prediction.

Due to the influence of complex operating conditions and environment, the collected vibration signals of the rolling bearing are usually non-stationary and non-linear [20]. Many non-linear dynamic methods, such as sample entropy [21], permutation entropy [22], fuzzy entropy [23], Rényi entropy [24], Wiener entropy [25], Instantaneous Spectral Entropy [26], dispersion entropy [27], are proposed, which can reflect the non-linear properties of vibration signal and characterize equipment health status. Sample Entropy is slow in calculating long time series, poor in real-time performance, and prone to a sudden change in similarity measurement [28]. Although the calculation of permutation entropy is simple, the amplitude information of the time series is ignored [29]. Fuzzy entropy is an improvement of sample entropy, but there are still some problems, such as slow calculation speed [30]. Dispersion entropy has the advantage of being less affected by sudden change signals and solves the problems existing in permutation entropy with better stability [31]. However, due to the complexity of vibration signals, single-scale entropy cannot fully reflect fault information, thus multiscale signals can fully excavate fault information. Subsequently, the methods of time series multiscale [32], composite multiscale [33], refine composite multiscale [34] are gradually put forward by scholars. Among them, refined composite multiscale dispersion entropy (RCMDE) has better stability and feature extraction ability than multiscale dispersion entropy (MDE) [35]. In [36], a new feature for analyzing time series complexity, called reverse dispersion entropy, is put forward. The reverse dispersion entropy combines the advantages of permutation and dispersion entropy and has stronger noise robustness and stability. Inspired by refined composite multiscale, the refined composite multiscale reverse dispersion entropy (RCMRDE) is proposed in this article, which can mine rolling bearing fault information comprehensively.

To realize the automatic recognition of equipment health status, many intelligent diagnostic models, such as backpropagation neural network (BPNN), support vector machine (SVM), k-nearest neighbor (KNN), were extensively used. However, BPNN is easy to fall into local minimum [37], SVM is unable to deal with large-scale data [38], and KNN classification has high spatial complexity and poor robustness [39]. As a classification and regression tool, random forest (RF) can usually obtain excellent diagnostic accuracy in mechanical fault detection [40]. In [41], the refined composite hierarchical fuzzy entropy (RCHFE) and RF are utilized for planetary gearbox fault diagnosis. In [42], a fault diagnosis method based on core principal component analysis and RF is proposed, which is applied to a wind energy conversion system and achieves good results. Therefore, considering the excellent performance of RF, in order to accurately mine fault information from bearing monitoring data and achieve high precision fault pattern recognition, an intelligent fault diagnosis method integrating VMD decomposition, refine composite multiscale reverse dispersion entropy (RCMRDE), JMIM feature selection and RF is proposed. The stability,

noise robustness, and signal discrimination ability of the proposed RCMRDE are verified by simulation signals. Furthermore, the effectiveness and superiority of the presented method are certificated by actual bearing vibration signals.

The main innovations of this research include the following two aspects:

- (1) RCMRDE is proposed for the first time, and its advantages in fault diagnosis are explored. Simulation and experimental results indicate that RCMRDE exhibits outstanding performance compared with several existing entropy.
- (2) There are few studies based on VMD and JMIM feature selection. JMIM feature selection can effectively calculate the resolution of each feature and select RCMRDE with high sensitivity to construct fault feature set. In this study, through JMIM feature selection, the original RCMRDE set is reduced by 91.4%, and the recognition accuracy is still 97.33%.

The remaining sections are organized as follows: Section 2 describes the implementation steps of RCMRDE and validates the superiority of RCMRDE by simulation case. Section 3 introduces the relevant theories and the proposed diagnostic framework. In Section 4, different rolling bearing health signals are collected, and then feature set construction based on VMD and RCMRDE and feature selection method based on JMIM algorithm are discussed. Next, some discussions and comparative analyses are carried out. Finally, the main contributions are summarized in Section 5.

## 2. Refined Composite Multiscale Reverse Dispersion Entropy

### 2.1. Reverse Dispersion Entropy

Reverse dispersion entropy [36] is used for detecting signal mutation, which combines the advantages of dispersion entropy and permutation entropy, and has better performance in mutation signal detection. Its calculation method is as follows:

Step1: Mapping time series to  $c$  classes.

- (1) Mapping by the normal distribution function.

For the time series  $X = \{x(i), i = 1, 2, \dots, T\}$  with  $T$  values, mapping  $X$  to  $Y = \{y(i), i = 1, 2, \dots, T\}$  by Equation (1):

$$y_i = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x_i} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt \quad (1)$$

where  $y(i) \in [0, 1]$ .  $\mu$  and  $\sigma$  denote expectation and variance, respectively.

- (2) Using a linear algorithm to map each  $y_i$  to integers in  $[1, c]$ .

We map  $Y$  to  $Z = \{z(i), i = 1, 2, \dots, T\}$  by using round  $(c * y_i + 0.5)$ , where  $c$  is the category number.

Step 2: Using Equation (2) to calculate the embedding vectors, we reconstruct  $Z$  into  $L$ :

$$\begin{bmatrix} \{z(1), z(1+\tau), \dots, z(1+(m+1)d)\} \\ \vdots \\ \{z(j), z(j+\tau), \dots, z(j+(m+1)d)\} \\ \vdots \\ \{z(L), z(L+\tau), \dots, z(L+(m+1)d)\} \end{bmatrix} \quad (2)$$

where  $d$  is the time delay and  $m$  is the embedding dimension,  $L$  is equal to  $T - (m - 1)d$ .

Step 3: Mapping the dispersion pattern of each embedded vector.

There exist  $c^m$  dispersion patterns, and each embedding vector can be mapped to a dispersion pattern  $\pi$ .

Step 4: Calculate the probability of each dispersion pattern.

The probability of  $i$ -th dispersion pattern can be written as:

$$P(\pi_i) = \frac{\text{Number}(\pi_i)}{N - (m-1)d} \quad (1 \leq i \leq c^m) \quad (3)$$

where  $\text{Number}(\pi_i)$  is the number of mapping from each dispersion pattern to  $\pi_i$ ,  $P(\pi_i)$  stands for the proportion of the number of  $i$ -th dispersion patterns to the number of embedding vectors.

Step 5: Calculating RDE.

$$H_{RDE}(X, m, c, d) = \sum_{i=1}^{c^m} \left( P(\pi_i) - \frac{1}{c^m} \right)^2 = \sum_{i=1}^{c^m} P(\pi_i)^2 - \frac{1}{c^m} \quad (4)$$

The normalized RDE is expressed as:

$$H_{RDE} = \frac{H_{RDE}(X, m, c, d)}{1 - 1/c^m} \quad (5)$$

## 2.2. Refined Composite Multiscale Reverse Dispersion Entropy

The refine processing of time series can mine statistical information more thoroughly than the coarse-grained process of multiscale algorithm, and the influence of the position of initial points on the calculation results can be effectively solved. RCMRDE includes the following steps:

For the time series  $X = \{x_1, x_2, \dots, x_T\}$  of length  $T$ , the  $k$ -th coarse-grained series for a given scale factor  $\tau$  can be given by:

$$x_{k,j}^\tau = \frac{1}{\tau} \sum_{i=(j-1)\tau+k}^{j\tau+k-1} x_i, 1 \leq j \leq [T/\tau], 1 \leq k \leq \tau \quad (6)$$

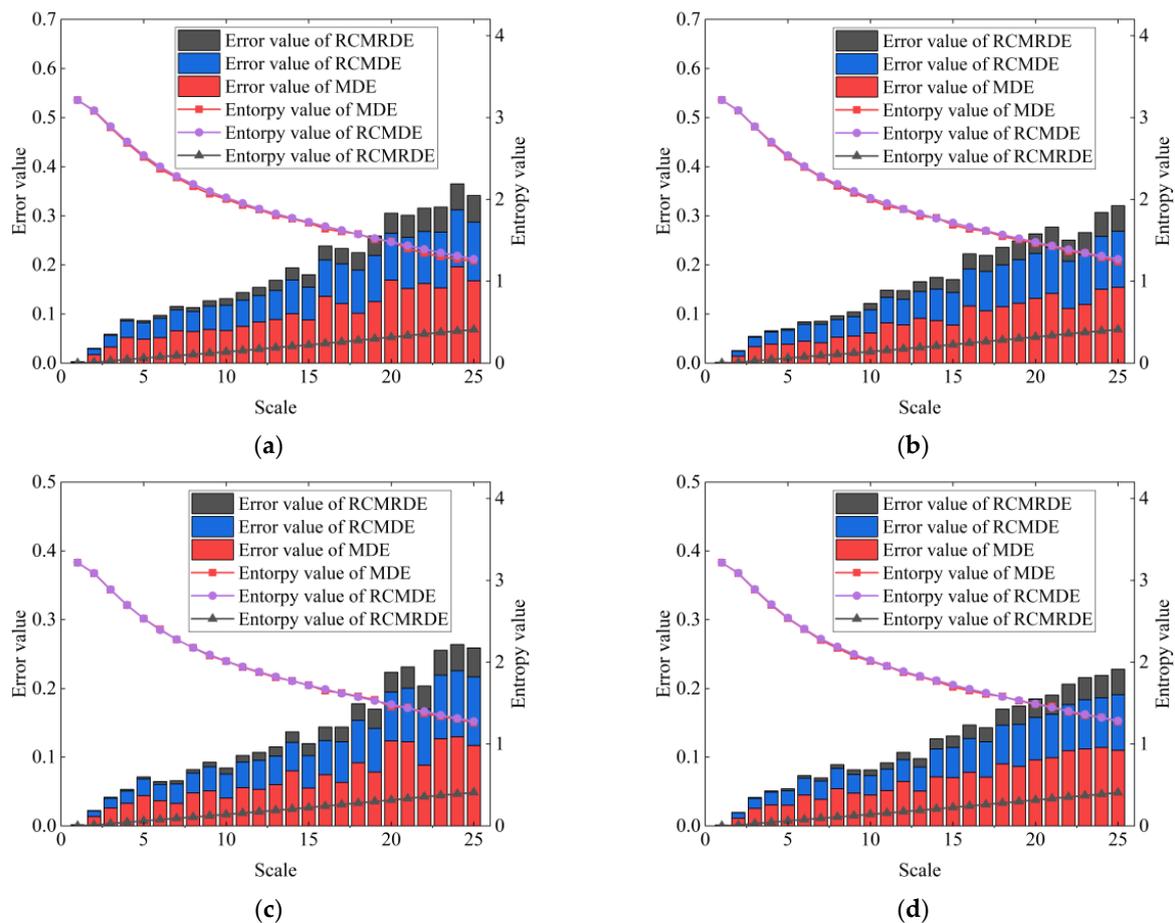
For each scale factor  $\tau$ , RCMRDE can be defined as:

$$\text{RCMRDE}(X, m, c, d, \tau) = - \sum_{i=1}^{c^m} \bar{P}(\pi_i)^2 - \frac{1}{d^m} \quad (7)$$

where  $\bar{P}(\pi_i) = \frac{1}{\tau} \sum_{k=1}^{\tau} P_k^{(\tau)}$ ,  $P_k^{(\tau)}$  indicates the dispersion mode probability corresponding to the  $k$ -th coarsening sequence under scale  $\tau$ .

## 2.3. Comparison between MDE, RCMDE, and RCMRDE Using Simulation Signals

In order to compare the entropy stability of RCMRDE, RCMDE, and MDE, we calculate the average entropy curves and error values (i.e., standard deviation) of RCMRDE, RCMDE, and MDE of 30 groups of white noise sequences under different data lengths and set  $m = 2$ ,  $d = 1$ ,  $c = 5$ ,  $\tau_{\max} = 25$ . The results are exhibited in Figure 1. It can be seen that the entropy curves of RCMRDE and RCMDE change more smoothly than MDE, which shows that the entropy estimation performance of RCMRDE and RCMDE is better than MDE. Meanwhile, compared with RCMDE and MDE, the standard deviation of RCMRDE is the smallest, which indicates that RCMRDE has outstanding stability.



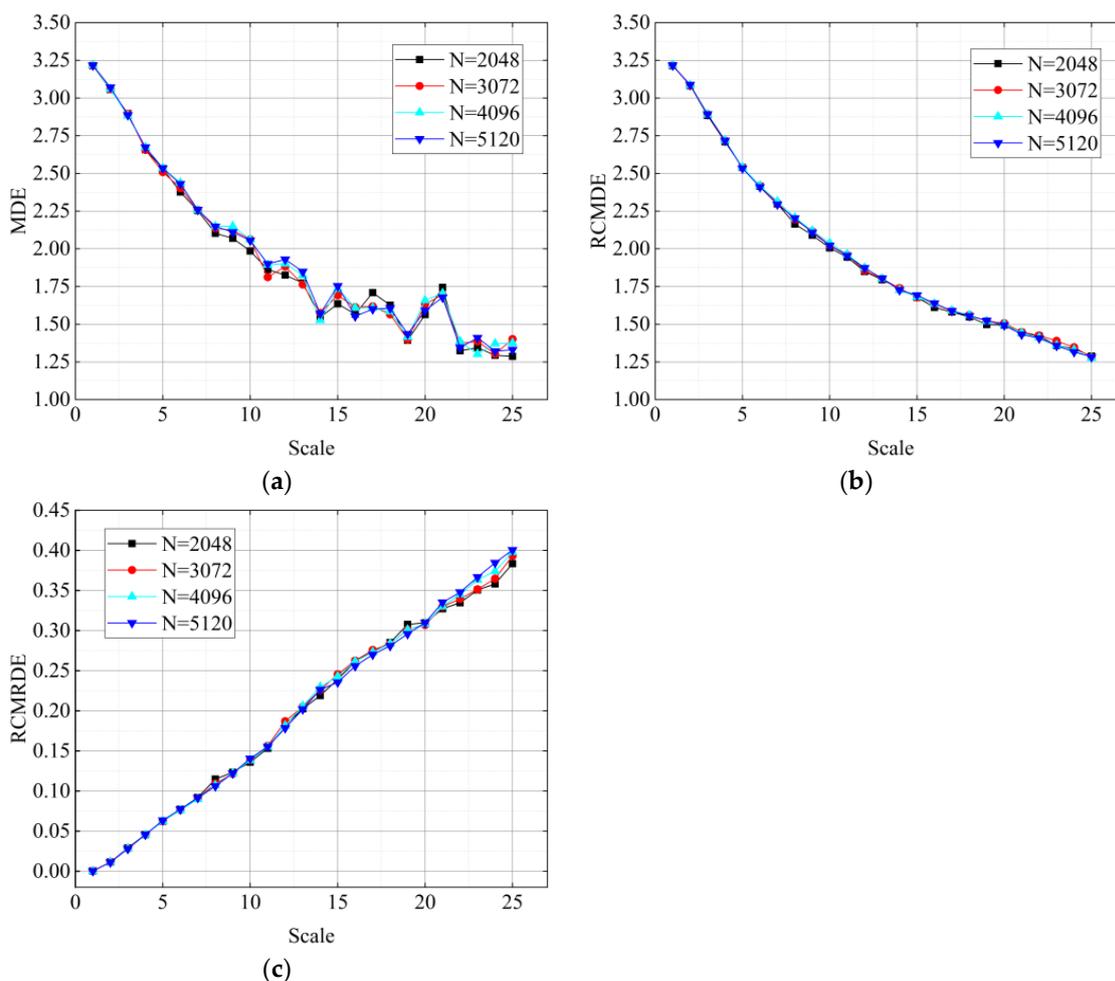
**Figure 1.** The mean curve and error value of MDE, RCMDE and RCMRDE for white noise with different  $N$ : (a)  $N = 2048$ ; (b)  $N = 3072$ ; (c)  $N = 4096$ ; (d)  $N = 5120$ .

To study the influences of data length on RCMRDE, three entropy values of white noise with different data lengths ( $N = 2048, 3072, 4096, 5120$ ) were calculated, respectively. The parameters of the three entropy values were set as  $m = 2, d = 1, c = 5, \tau_{\max} = 25$ . The results are shown in Figure 2. With the increase of  $\tau$ , the curves of RCMRDE and RCMDE were smoother and had a smaller fluctuation range than MDE, which shows that RCMRDE and RCMDE have better stability than MDE. Meanwhile, the proposed RCMRDE can exhibit good feature extraction ability under less data length.

Subsequently, in order to evaluate the calculation complexity of the three entropy values, we used WGN noise with different data lengths ( $N = 2048, 3072, 4096, 5120$ ) to calculate the computing times they spent. The results are listed in Table 1. The parameters of the three kinds of entropy were set as  $m = 2, d = 1, c = 5, \tau_{\max} = 25$ . Table 1 shows that the calculation time of RCMRDE was slightly longer than MDE but largely lower than RCMDE. With the increase of data length, the calculation time of three entropy values increased significantly. Therefore, selecting less data length can improve the computational efficiency when extracting features.

**Table 1.** The comparison of computation time of three entropy values under different data lengths.

Methods	Data Length			
	2048	3072	4096	5120
MDE	0.1059 s	0.1148 s	0.1235 s	0.1335 s
RCMDE	0.2061 s	0.2867 s	0.3704 s	0.4490 s
RCMRDE	0.1997 s	0.2818 s	0.3637 s	0.4431 s



**Figure 2.** The MDE, RCMDE, and RCMRDE of WGN noise under different data lengths: (a) MDE; (b) RCMDE; (c) RCMRDE.

Furthermore, in order to research the influences of category number  $c$  on RCMRDE, three entropy values of WGN (MDE, RCMDE, and RCMRDE) were calculated under different values. The results are displayed in Figure 3. The parameters of the three entropy values were set as  $m = 2$ ,  $d = 1$ ,  $\tau_{\max} = 25$ ,  $n = 2048$ , and  $c$  increased from 3 to 7. It can be found that comparing with the MDE curves, the RCMDE curves and the RCMRDE curves fluctuated slightly (except  $c = 3$ ), which indicates that RCMDE and RCMRDE can provide stable and reliable entropy values under different  $c$ .

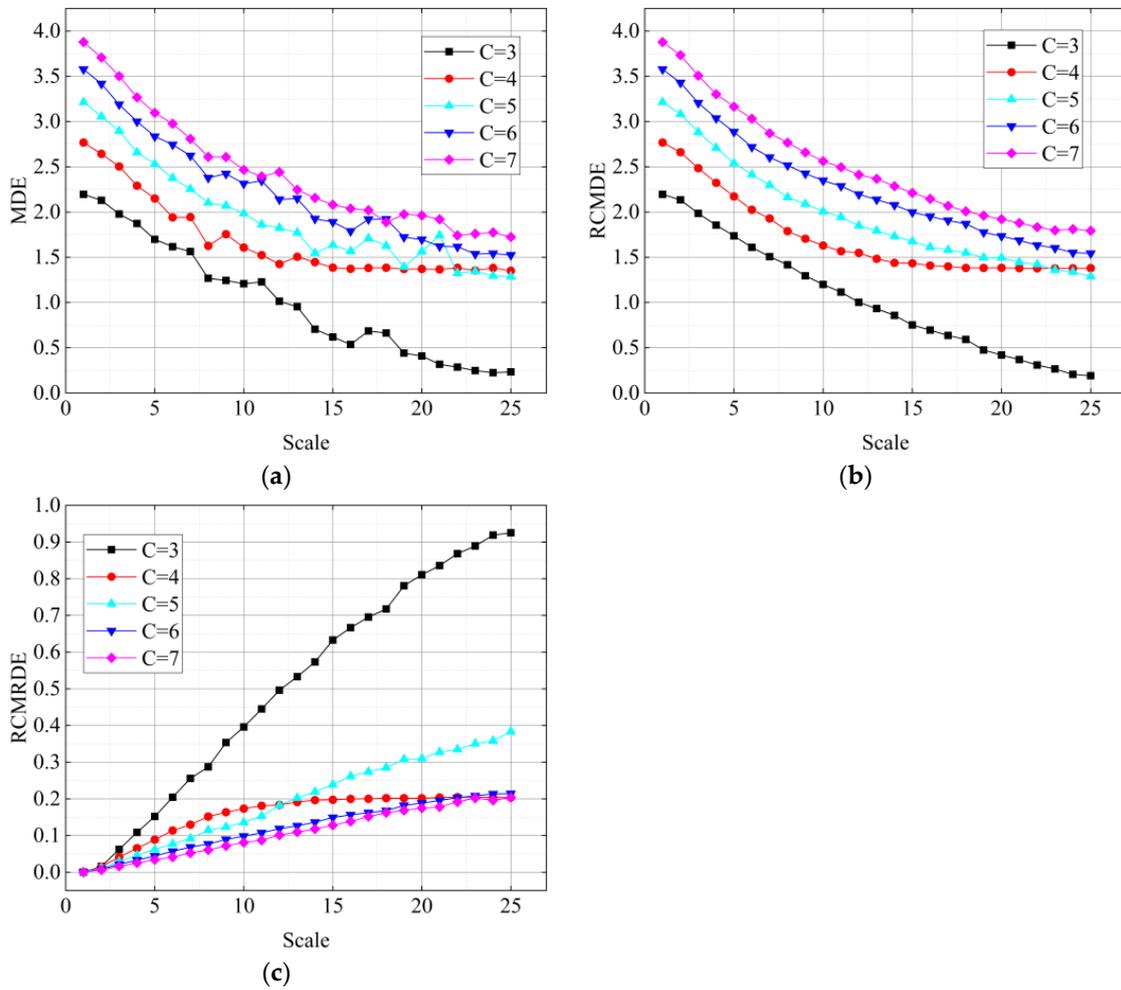
In addition, we can see from Figure 3c that with the increase of  $c$  value, the dispersion mode will increase, thus leading to the decrease of RCMRDE curves. Meanwhile, in practical application, if the  $c$  value was too small, the signal characteristic information cannot be fully extracted. Therefore, considering the reliability and calculation efficiency, the  $c$  value should be set as 5 or 6.

When local damage occurs to rolling bearings, the measured vibration signal contains much noise. Therefore, we used Equation (8) to study the anti-noise performance of RCMRDE.

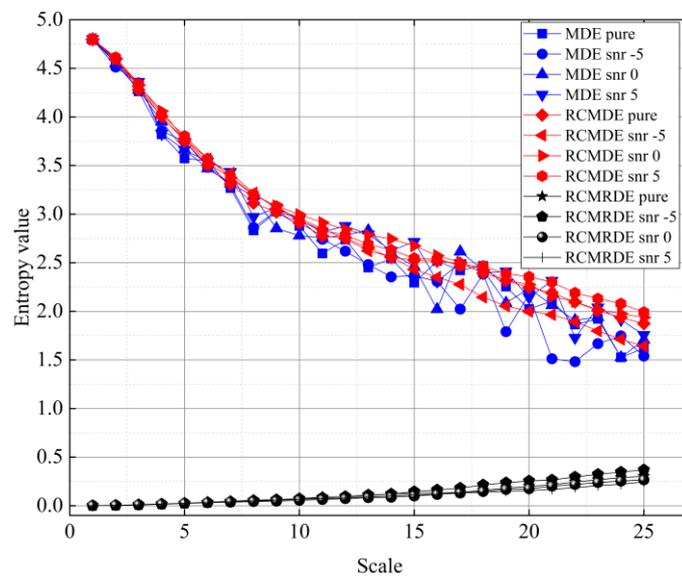
$$x(t) = [1 + 0.5 \cos(8\pi t)] \cos[200\pi t + 2 \cos(10\pi t)] + 0.8 \sin(\pi t) \sin(30\pi t) + n(t) \quad t \in (0, 1) \quad (8)$$

where  $n(t)$  represents the Gaussian white noise with a SNR of  $-5$ ,  $0$ , and  $5$  dB, respectively. In addition,  $m$ ,  $c$ ,  $d$  and  $\tau_{\max}$  of MDE, RCMDE, and RCMRDE are set to 3, 5, 1, and 25, respectively. The signal length was 2000 points. Figure 4 displays the calculation results of

three entropy values. With the increase of  $\tau$ , the RCMRDE value was more concentrated and the fluctuation was smaller, which proves that RCMRDE has excellent noise robustness.



**Figure 3.** MDE, RCMDE, RCMRDE of WGN noise under different numbers of  $c$  (a) MDE of WGN; (b) RCMDE of WGN; (c) RCMRDE of WGN.



**Figure 4.** The MDE, RCMDE, and RCMRDE of synthetic signal with different SNRs.

Herein, the bearing outer race fault model with four different fault degrees was established to validate the superiority of RCMRDE in signal discrimination ability. The mathematical expression of the fault model is as follows:

$$x(t) = e(t) + r(t) + p(t) + n(t) \quad (9)$$

where  $e(t)$  denotes a repetitive pulse caused by a local defect:

$$e(t) = \sum_{i=1}^{M_1} A(t) e^{-\zeta_\alpha(t-iT_\alpha-\delta_i)} \cos[2\pi f_\alpha(t-iT_\alpha-\delta_i) + \varphi_\alpha] \quad (10)$$

where  $M_1$  represents the number of fault pulses and  $A(t)$  stands for the amplitude of fault impulses.  $T_\alpha$  is the time interval between two adjacent fault impulse.  $\delta_i$  is a random time lag between two repetitive impulses.  $f_\alpha$ ,  $\zeta_\alpha$  and  $\varphi_\alpha$  are the resonant frequency, damping coefficient, and phase excited by fault impacts, respectively.

In Equation (9),  $r(t)$  stands for random impulses generated by external excitation:

$$r(t) = \sum_{s=1}^{M_2} B_s e^{-\zeta_b(t-T_s)} \cos[2\pi f_b(t-T_s) + \varphi_b] \quad (11)$$

where  $M_2$ ,  $B_s$  and  $T_s$  represents the number, amplitude, and occurrence time of random impulses, respectively. The meanings of  $f_b$ ,  $\zeta_b$  and  $\varphi_b$  the same as those of the corresponding symbols in Equation (10).

In Equation (9),  $p(t)$  is the pure periodic component generated by axis rotation:

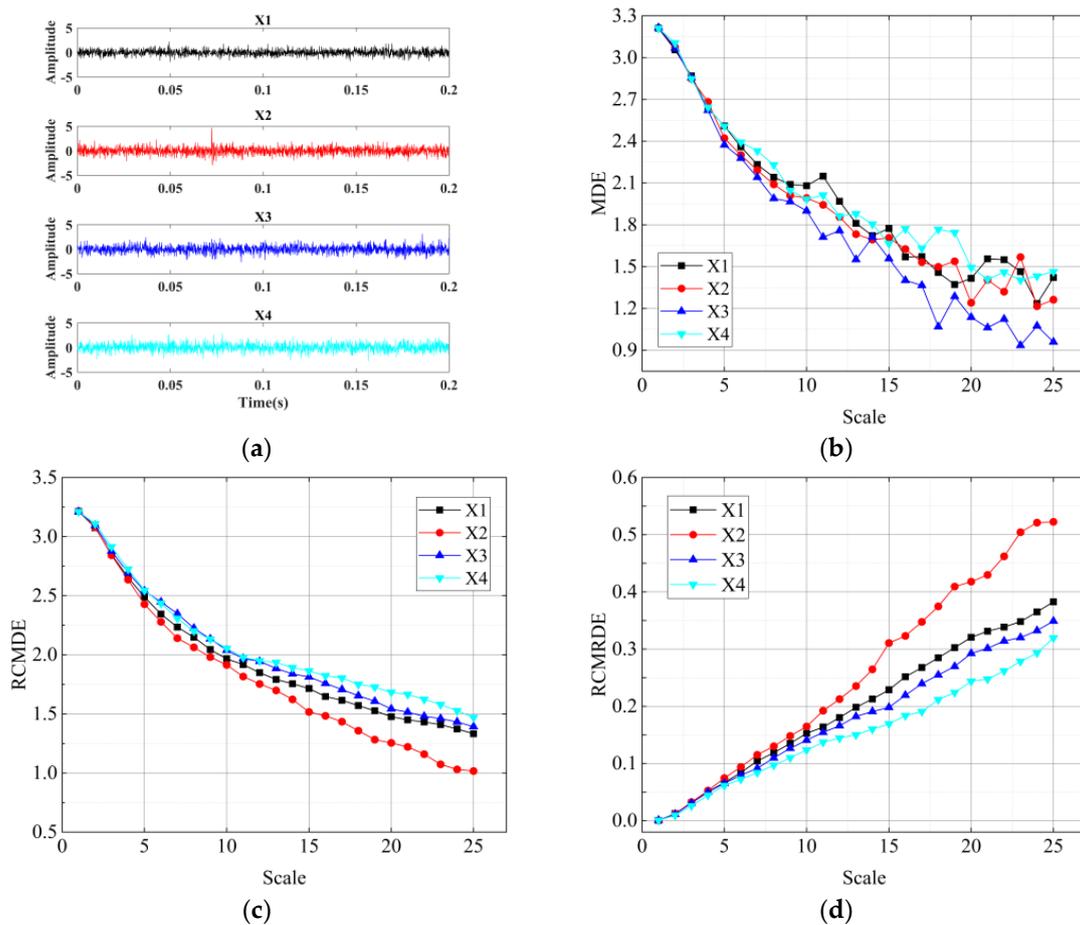
$$p(t) = \sum_{k=1}^{M_3} C_k \sin(2\pi f_k t + \theta_k) \quad (12)$$

where  $M_3$  is the number of periodic harmonics.  $C_k$ ,  $f_k$  and  $\theta_k$  are the amplitude, frequency, and phase of components, respectively.  $n(t)$  in Equation (9) is the added noise. The model parameters are given in Table 2.

**Table 2.** Parameter values of bearing outer race fault simulation model.

Parameter	Value	Parameter	Value	Parameter	Value
$M_1$	73	$M_2$	3	$M_3$	2
$C_2(m/s^2)$	0.025	$B_s$	N(2, 5, 1)	$C_2(m/s^2)$	0.025
$A(t)(m/s^2)$	0.5, 0.6, 0.7, 0.8	$T_s(s)$	U(1, 8192)/ $f_s$	$f_1(Hz)$	10
$f_\alpha(Hz)$	2600	$f_b(Hz)$	1700	$f_2(Hz)$	20
$\varphi_\alpha(rad)$	0	$\varphi_b(rad)$	0	$\theta_1(rad)$	$\pi/6$
$\zeta_\alpha$	1000	$\zeta_b$	800	$\theta_2(rad)$	$-\pi/3$

In the simulation experiment, the sampling frequency was 12,800 Hz, and the signal length was 8192 points. N(2.5, 1) and U(1, 8192) denoted a normal distribution and a uniform distribution, respectively. Adjust the size of  $A(t)$  to 0.5, 0.6, 0.7, and 0.8, respectively. Therefore, four simulation signals that represented by  $X_1$ ,  $X_2$ ,  $X_3$ ,  $X_4$  were generated, corresponding to the damage degree of the outer ring from slight to serious. The waveforms of four simulation signals are given in Figure 5a. However, the fault types and degrees cannot be distinguished by the time domain waveform. Next, we calculated the MDE, RCMDE, and RCMRDE of simulation signals with four different fault degrees, as displayed in Figure 5.



**Figure 5.** Simulation signals and corresponding three entropy values: (a) four bearing outer fault signals with different damage degrees; (b) MDE value curves; (c) RCMDE value curves; (d) RCMRDE value curves.

It can be seen from Figure 5 that the four curves of MDE intersect with each other, while RCMDE and RCMRDE show good independence, which indicates that RCMDE and RCMRDE have outstanding fault discrimination ability. Furthermore, RCMRDE shows better distinguishing performance than RCMDE in determining the degree of bearing fault. Thus, the RCMRDE can more accurately characterize the bearing health status.

### 3. The Proposed Fault Diagnosis Method

#### 3.1. Variational Mode Decomposition

VMD is an adaptive non-stationary signal decomposition technology, which can decompose the given signal into an ensemble of intrinsic mode functions (IMFs), and each IMF has limited bandwidth in its spectrum [43]. VMD determines the correlation frequency band and decomposes the original signal into  $k$  modal components. The constrained variational problems are given by Equation (13):

$$\begin{aligned} \min_{\{u_k\}, \{\omega_k\}} & \left\{ \sum_k \|\partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \|_2^2 \right\} \\ \text{s.t.} & \sum_k u_k = f(t) \end{aligned} \quad (13)$$

where  $u_k$  is the  $k$ -th IMF of the signal.  $\omega_k$  is the center frequency.  $f(t)$  is the input signal and  $\delta(t)$  is the Dirac function.  $\partial_t$  is the partial derivative of a function to  $t$ .  $*$  is the convolution

symbol. Considering a quadratic penalty term and Lagrange multipliers  $\lambda$ , Equation (13) can be written as follows:

$$L(\{u_k\}, \{\omega_k\}, \lambda) = \alpha \sum_k \|\partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t}\|_2^2 + \|f(t) - \sum_k u_k(t)\|_2^2 + \left\langle \lambda(t) f(t) - \sum_k u_k(t) \right\rangle \quad (14)$$

where  $\alpha$  is the balance parameter of the data fidelity constraint.  $\lambda$  is a common method of strictly enforcing constraints.  $\|\bullet\|_2^2$  represent the squared  $L^2$  norm. VMD employs the alternating direction method of multipliers to solve Equation (14). The solution of the second-order optimization problem can be obtained. Each estimated IMF can be expressed as:

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i^n(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2} \quad (15)$$

where  $\hat{f}(\omega)$ ,  $\hat{u}_i(\omega)$ ,  $\hat{\lambda}(\omega)$  stand for the Fourier transforms of  $f(\omega)$ ,  $u_i(\omega)$  and  $\lambda(\omega)$ , respectively.  $n$  represents the iterations number. As above, optimization is performed in the Fourier transform domain to find the optimal central frequency  $\omega_k$ , which can be obtained by Equation (16). The acquired new center frequency is given as:

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega^2 |\hat{u}_k(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k(\omega)|^2 d\omega} \quad (16)$$

### 3.2. Feature Selection Based on JMIM

Feature selection is widely used in many fields, such as data mining and machine learning. Feature selection based on information theory is a popular method because it has a tremendous advantage in computational efficiency. However, the disadvantage of this method is lacking information about the interaction between features and classifiers, as well as the selection of redundancy features. Therefore, it is crucial to select a reasonable feature selection method to reduce dimension and improve classification accuracy.

The joint mutual information maximization (JMIM) is an effective feature selection algorithm, which can extract features and create a feature subset efficiently based on joint mutual information [44]. Compared with many other feature selection methods such as joint mutual information (JMI), maximum relevancy minimum redundancy (mRMR), etc. JMIM feature selection method can achieve the best trade-off in accuracy and stability [45]. Therefore, this paper employs JMIM as the sensitive feature selection criterion to reduce features redundancy and improve recognition accuracy.

### 3.3. Random Forest

Random Forest (RF) is a machine learning algorithm that consists of many independent decision trees [46]. RF uses multiple CART (Classification and Regression Tree) as meta-classifier and applies bagging algorithms to produce different training sample sets. Meanwhile, it randomly selects features to split the internal nodes when constructing a single tree. Therefore, RF can tolerate noise better and have better classification performance. As a multi-functional machine learning algorithm in practical application, RF is not only used for regression and classification but also for processing missing values, outliers, and other data exploration. The RF algorithm includes the following steps [47]:

Step 1: set a sample set of  $N$  for the samples number,  $M$  for the variables number;

Step 2: each node will randomly select  $m$  ( $m < M$ ), a specific variable, which is then used to determine the optimal splitting point. The value of  $M$  remains constant during the generation of the decision tree;

Step 3: sample  $N$  times from the sample set ( $N$  samples) to form a set of training sets;

Step 4: for each node,  $m$  variables based on this point are randomly selected, and their optimal splitting points are calculated;

Step 5: each decision tree will grow as much as possible without pruning and will forecast new data by adding all the decision trees together.

### 3.4. Proposed Fault Diagnostic Framework

In this paper, an intelligent diagnosis method integrated VMD, RCMRDE, and RF model is shown in Figure 6, and specific implementation steps include:

- (1) VMD is applied to decompose the original signal into several modal components. The modal number is based on the decomposition criterion that the frequency center frequency of each component is not overlapping.
- (2) Based on the original signal and VMD decomposed components, RCMRDE at 25 scales is calculated as the initial feature set. The high-dimensional features contain redundant information. Subsequently, JMIM is employed to select sensitive features, thereby removing redundant information and reducing the dimension of feature set data.
- (3) Input sensitive features selected in steps (2) into the RF model to identify bearing health status. The presented method performance is checked by the rolling bearing vibration signals under different conditions.

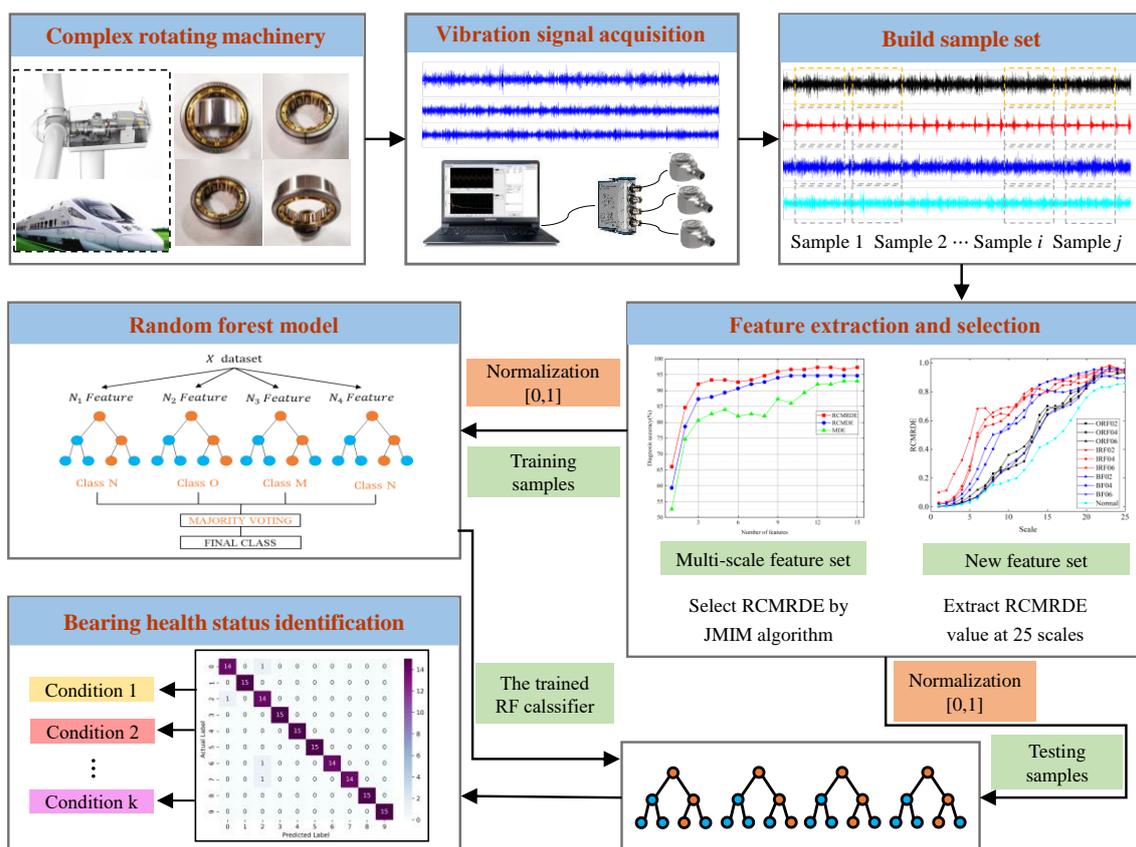


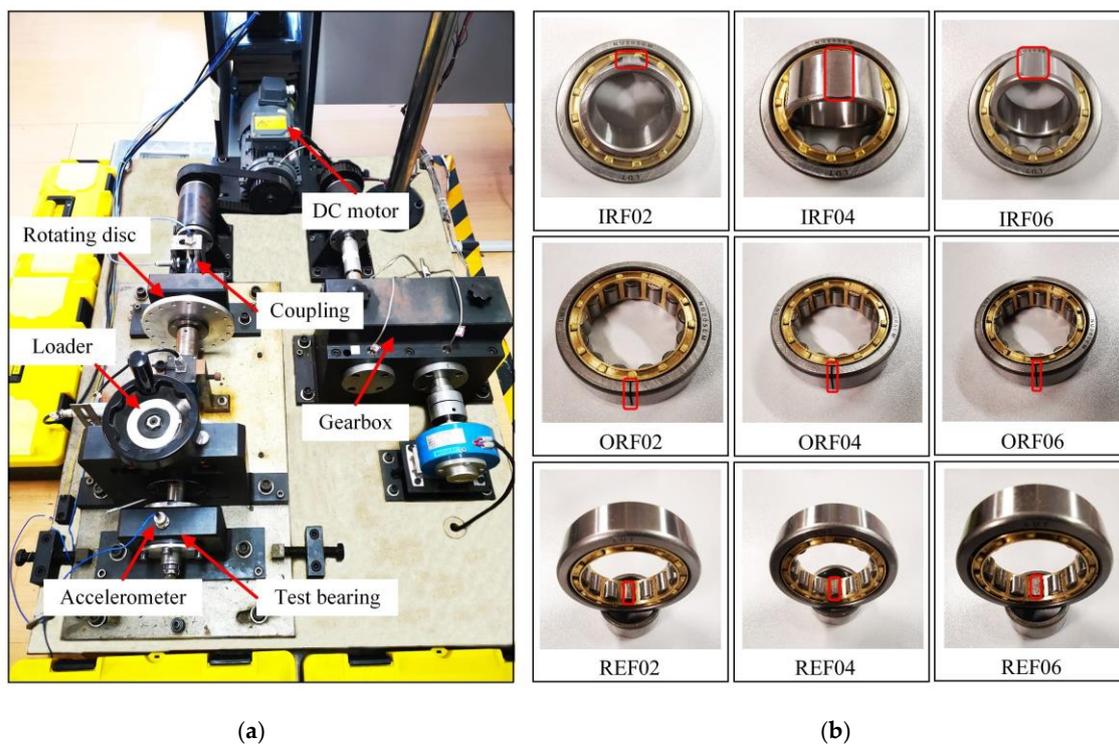
Figure 6. The flowchart of the proposed method.

## 4. Experiments and Data Analysis

### 4.1. Experimental Setup and Data Acquisition

To verify the effectiveness and feasibility of the presented method, actual data collection was performed using the bearing fault diagnosis test bench QPZZ-II, which is displayed in Figure 7a. The test bench can simulate multiple bearing failure types. In order to simulate the bearing fault damage in the real situation, 2 different types of bearings were processed by wire cutting. The fault types were inner race 0.2 mm, 0.4 mm, and

0.6 mm wear, outer race 0.2 mm, 0.4 mm, and 0.6 mm crack, and rolling element 0.2 mm, 0.4 mm, and 0.6 mm wear. The bearings of the 3 fault types are shown in Figure 7b. The sensor used 3035B accelerometer produced by the DYTRAN company in the United States, which was installed in the radial direction of the rolling bearing. The sampling frequency was 12,800 Hz, and the rotating speed was 1200 r/min. The bearing models used were N205EM/PS with a detachable outer ring and NU205EM/PS with a detachable inner ring. A total of 50 non-overlapping samples were selected from the vibration signals of each fault type, and the length of each sample was 2048 points, of which 35 groups were used for training and 15 groups were used for testing. The dimensional parameters of the bearing are shown in Table 3. Table 4 gives the description of experimental data. Bearing vibration signals under 10 different health conditions are illustrated in Figure 8. One cannot judge the fault types and damage degrees only by observing the time domain waveforms. Therefore, it was necessary to select reasonable feature extraction and recognition methods to realize the automatic and accurate discrimination of bearing health states.



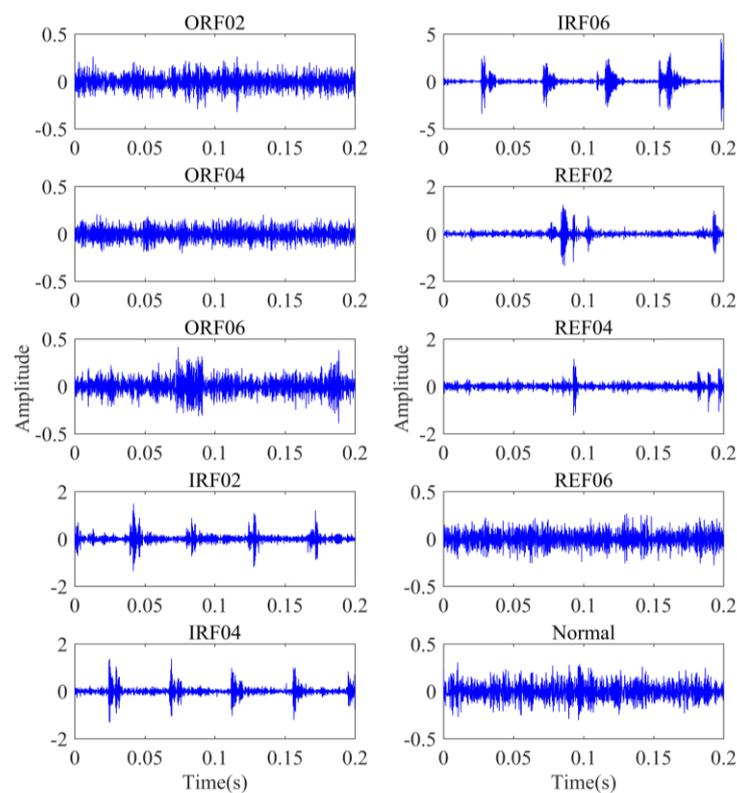
**Figure 7.** Rolling bearing fault diagnosis test bench and faulty parts: (a) test bench; (b) different types of faulty parts.

**Table 3.** The information of bearing dimension.

Type	Inner Diameter (mm)	Outer Diameter (mm)	Rolling Element Diameter (mm)	Pitch Circle Diameter (mm)	Number of Rolling Element	Characteristic
N205EM/PS	25	52	7.5	39	13	Detachable outer ring
NU205EM/PS	25	52	7.5	39	13	Detachable inner ring

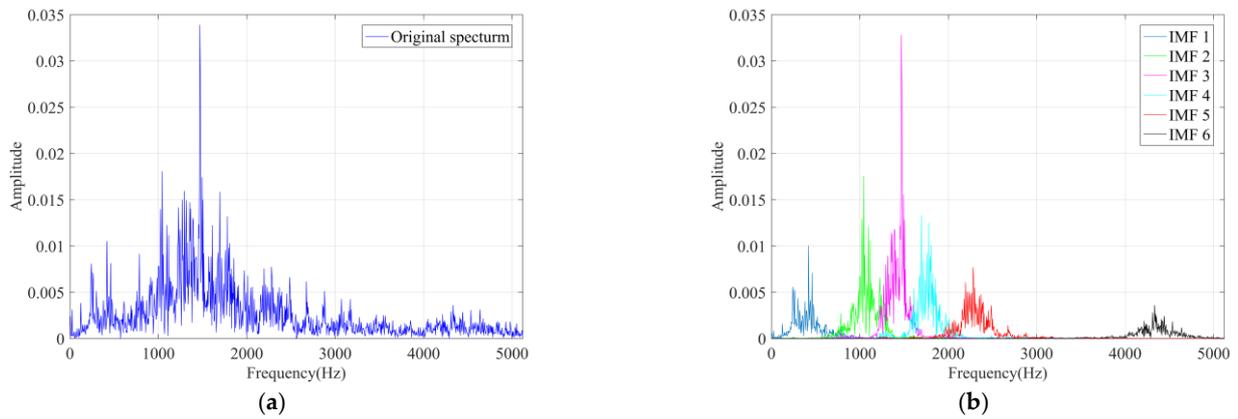
**Table 4.** The description of experimental data.

Bearing State	Fault Size (mm)	Abbreviation	Label	Training Data	Test Data
Outer race fault	0.2	ORF02	0	35	15
Outer race fault	0.4	ORF04	1	35	15
Outer race fault	0.6	ORF06	2	35	15
Inner race fault	0.2	IRF02	3	35	15
Inner race fault	0.4	IRF04	4	35	15
Inner race fault	0.6	IRF06	5	35	15
Rolling element fault	0.2	REF02	6	35	15
Rolling element fault	0.4	REF04	7	35	15
Rolling element fault	0.6	REF06	8	35	15
Normal	-	Normal	9	35	15

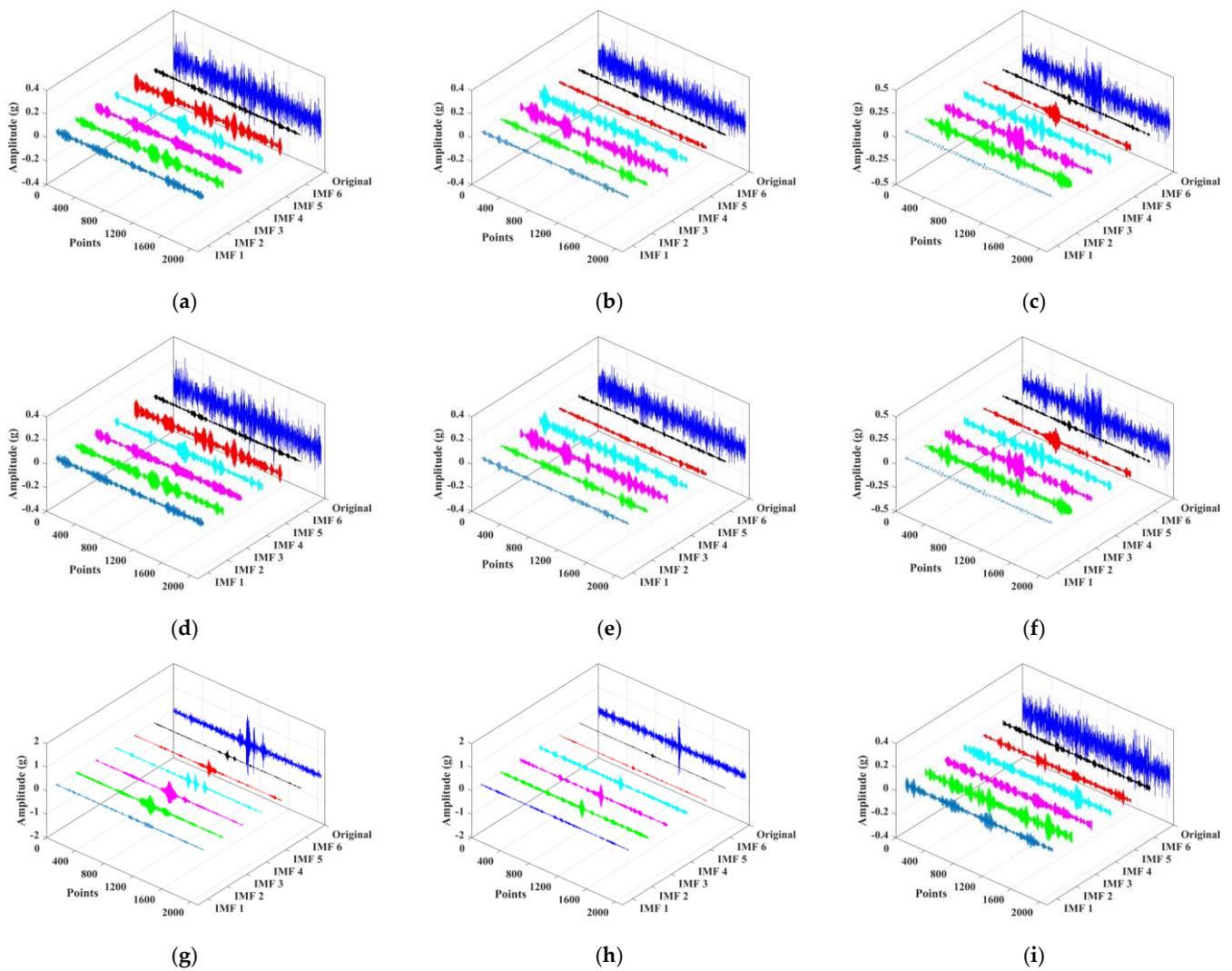
**Figure 8.** Bearing vibration signals under 10 different health conditions.

#### 4.2. Feature Extraction by VMD-Based RCMRDE

In this section, the VMD was applied to decompose the original signal into multiple IMF components. The principle of setting the number of VMD decomposition was that the center frequency of each component spectrum will not overlap after decomposition. When the decomposition level was set to 6, all 10 types of signals were decomposed sufficiently without overlapping of the center frequency of each component. Therefore, the decomposition level was set to 6. Here, taking the normal bearing signal as an example, its frequency spectrum and corresponding decomposition components are shown in Figure 9. It can be found that the original signal was decomposed into 6 IMF components, which achieved complete decomposition and no frequency aliasing. Subsequently, the parameters of RCMRDE were set as  $m = 2$ ,  $d = 1$ ,  $c = 5$ ,  $\tau_{\max} = 25$ . The VMD decomposition results of each fault type signal are given in Figure 10.



**Figure 9.** The frequency spectrum of normal bearing signal: (a) spectrum of original signal; (b) spectrum of each component.



**Figure 10.** The VMD decomposition results of each fault type signal: (a) outer race 0.2 mm crack; (b) outer race 0.4 mm crack; (c) outer race 0.4 mm crack; (d) inner race 0.2 mm wear; (e) inner race 0.4 mm wear; (f) inner race 0.6 mm wear; (g) rolling element 0.2 mm wear; (h) rolling element 0.4 mm wear; (i) rolling element 0.6 mm wear.

According to Steps 2 in Section 3.4, there were 25 RCMRDE values obtained from the original signal, and 150 RCMRDE values were obtained from the VMD decomposed components. Therefore, a total of 175 eigenvalues were obtained for each type of fault signal. The RCMRDE values of the original signal and the first IMF component are displayed in Figure 11. We can observe that the RCMRDE values of different fault types were well separable, which indicates the proposed method can mine the distinguishing characteristics between different types and degrees of bearing fault.

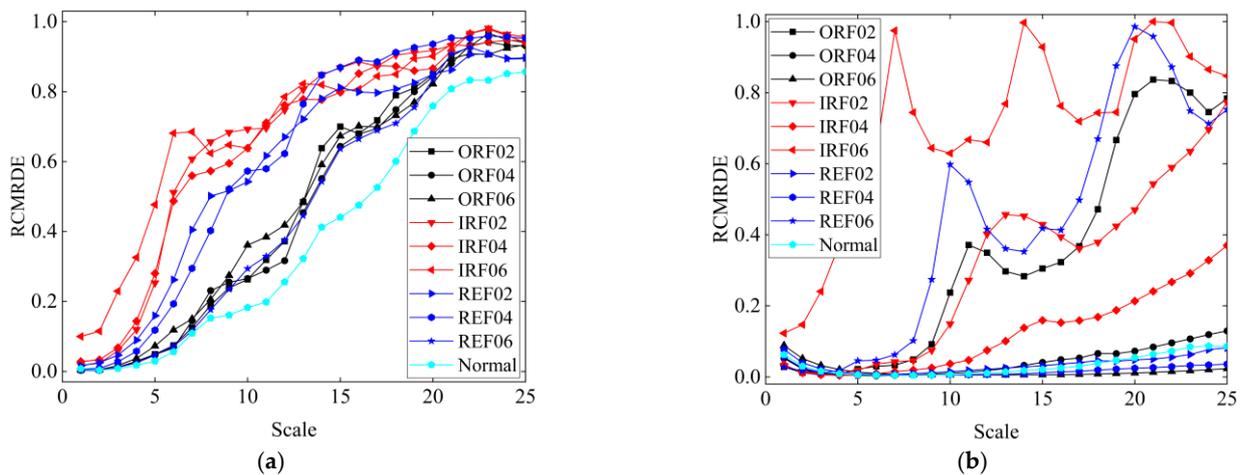


Figure 11. The RCMRDE values obtained under scale 25: (a) RCMRDE values of original signal; (b) RCMRDE values of the first IMF components.

### 4.3. Diagnosis Results and Analysis

Considering that high-dimensional features data contain redundant information, the JMIM method is utilized to select the sensitive RCMRDE values with large contributions and excellent classification performance. Finally, the top 15 features selected are used for establishing the new feature set and then feed them into the RF model. Figure 12 depicts the multi-class confusion matrix of the presented method. It can be observed that the classification accuracy of ORF02, ORF06, REF02, and REF06 was 93.33%, and the other fault types were 100%, and the average classification accuracy was 97.3%, which indicates that the proposed method can accurately and effectively identify bearing health status and obtain satisfactory diagnosis results.

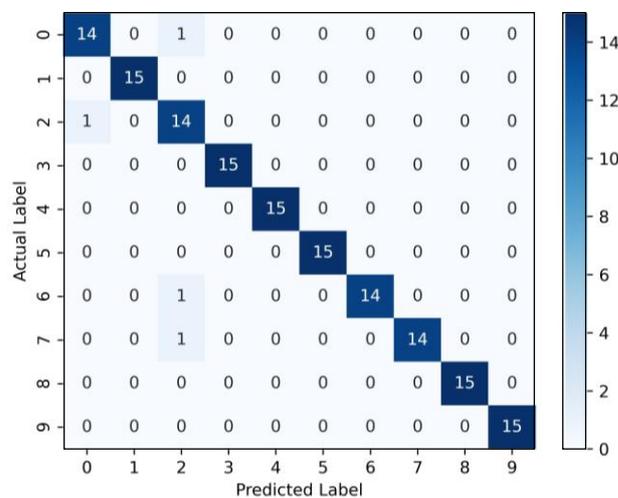


Figure 12. Multi-class confusion matrix of the proposed method.

Furthermore, in order to illustrate the superiority of the presented feature extraction algorithm, the RCMRDE features were compared with RCMDE and MDE, and the entropy values of the top 15 features were selected, respectively, and then input into the RF model. The diagnostic accuracy is given in Figure 13. It can be known that the diagnostic accuracy of RCMRDE was the best, up to 97.33%, followed by RCMDE, up to 94.67%. Meanwhile, with the increase of features number, the diagnostic accuracy of RCMRDE shows the fastest growth rate. The above results indicate that RCMRDE can more thoroughly detect the dynamic mutation of the bearing fault signal.

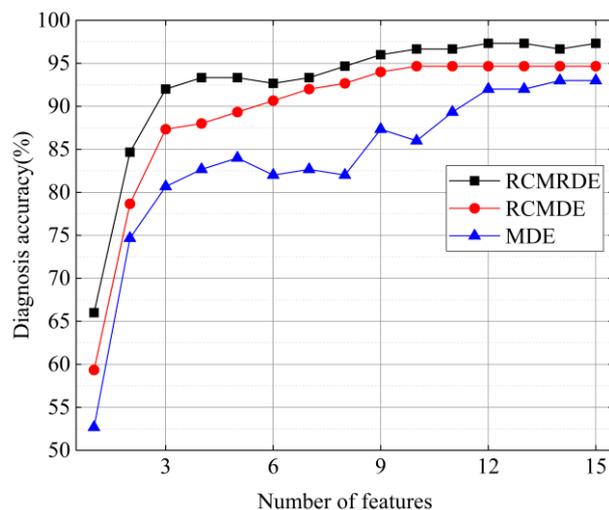


Figure 13. The diagnosis accuracy of three methods.

Subsequently, the JMIM is compared with Fisher [48] and LS [49]. Figure 14 illustrates the diagnostic accuracy of the 3 methods. We can see from Figure 14 that the presented method achieves the highest diagnostic accuracy of 97.33% (12 sensitive features are selected), the Fisher method obtained the diagnostic accuracy of 94.67% (14 sensitive features are selected), and the LS method gets the diagnostic accuracy of 90.67% (14 sensitive features are selected). Some conclusions can be drawn through comparison: (1) the sensitive feature selection criteria can reduce the dimension of feature data and improve computational efficiency; (2) JMIM is better than Fisher and LS in terms of sensitive feature selection; (3) the effective sensitive features can significantly improve diagnostic accuracy.

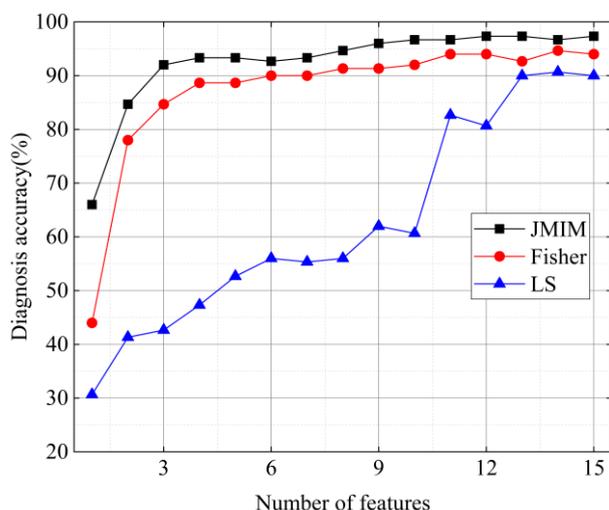


Figure 14. The diagnosis accuracy of JMIM, Fisher and LS feature selection methods.

#### 4.4. Comparison with Other Methods

To further evaluate the presented method, some machine learning algorithms, such as KNN, BPNN, and SVM, were selected for comparison. In total, the proposed method was compared with the other 23 methods, and the results are shown in Figure 15. In the figure, the MDE\_Original stands for extracting MDE features from the original signal, where  $\tau_{\max}$  is set to 25 and the top 15 sensitive features are selected as the final feature set. It should be noted the feature set construction process in RCMDE\_Original and RCMRDE\_Original are the same as those in MDE\_Original. In addition, the MDE\_VMD stands for extracting MDEs from the original signal and the VMD decomposition components, where the  $\tau_{\max}$  of the original signal and each IMF were set to 25, the decomposition number of VMD was set to 6. Then, the top 15 features were selected from 175 features data to build the sensitive feature set. Similarly, the feature set construction process in RCMDE\_VMD and RCMRDE\_VMD were the same as those in MDE\_VMD.

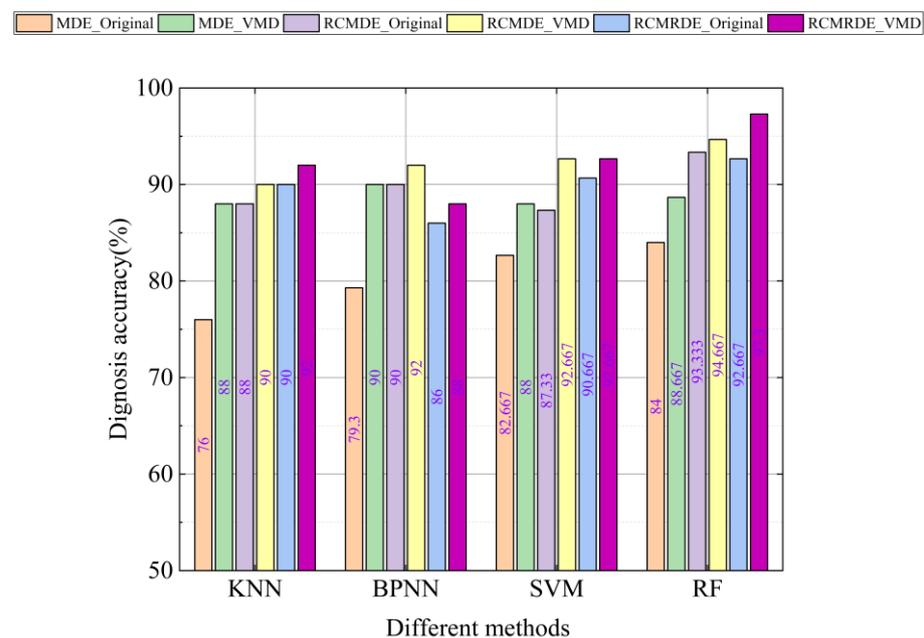


Figure 15. Comparison of diagnostic accuracy of different methods.

It can be found from Figure 15 that when using the same machine learning model (except BPNN), the accuracy of RCMRDE\_VMD was higher than or equal to MDE\_VMD and RCMDE\_VMD. This proves that RCMRDE can more clearly characterize the bearings health status. In addition, for the 4 machine learning models, the diagnostic accuracy of MDE\_VMD was better than MDE\_Original, RCMDE\_VMD was better than RCMDE\_Original, and RCMRDE\_VMD was better than RCMRDE\_Original. This indicates that the sensitive feature information can be more fully excavated after the original signal is decomposed by VMD. It is worth noting that the diagnostic accuracy of the presented method was highest, with a recognition rate of 97.33%. As a result, the presented method has a superior performance in the intelligent diagnosis of bearing fault type and damage degree.

## 5. Conclusions

In this research, a novel intelligent diagnosis of rolling bearings combining RCMRDE and RF model is proposed. First, aiming at the disadvantage of traditional multiscale entropy cannot accurately extract the useful feature from non-stationary fault signal, RCMRDE is created to characterize different fault types and damage degrees of rolling bearings. The simulated experiment shows that RCMRDE performs better in mutation signal detection, noise robustness test, stability, and signal discrimination. Secondly, the JMIM method

is applied to select sensitive features and build feature datasets, thus reducing redundant information and improving computational efficiency. Then, the sensitive feature set is fed into the RF model to realize the automatic and accurate discrimination of bearing health status. The effectiveness of the proposed method is fully demonstrated by the simulation signals and actual bearing diagnosis experiment. Furthermore, the proposed method shows superior performance compared with other machine learning methods and feature extraction methods. The comparisons show that the proposed method achieves outstanding diagnostic results, and its recognition accuracy is 97.33%. However, the decomposition number of VMD still needs further optimization. Future work will consider choosing a good optimization algorithm to realize the best mode decomposition of the rolling bearing fault signal.

**Author Contributions:** Conceptualization, A.L. and Z.Y.; methodology, A.L. and H.L.; software, A.L. and X.L.; validation, A.L. and C.W.; formal analysis, A.L.; investigation, H.L. and Z.Y.; resources, A.L., X.L. and C.W.; data curation, A.L.; writing—original draft preparation, A.L. and C.W.; writing—review and editing, A.L. and C.W.; visualization, A.L. and X.L.; supervision, H.L. and Z.Y.; project administration, H.L.; funding acquisition, H.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research work is supported by the National Key Research and Development Program of China (Grant No.2019YFB2004600).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Yan, X.; Jia, M. Intelligent fault diagnosis of rotating machinery using improved multiscale dispersion entropy and mRMR feature selection. *Knowl.-Based Syst.* **2019**, *163*, 450–471. [[CrossRef](#)]
2. Chen, Z.; Mauricio, A.; Li, W.; Gryllias, K. A deep learning method for bearing fault diagnosis based on Cyclic Spectral Coherence and Convolutional Neural Networks. *Mech. Syst. Signal Process.* **2020**, *140*, 106683. [[CrossRef](#)]
3. Kaplan, K.; Lmaz, K.Y.; Melih, K.; Recep, M.N.M.; Metin, E.H. An improved feature extraction method using texture analysis with LBP for bearing fault diagnosis. *Appl. Soft Comput.* **2020**, *87*, 106019. [[CrossRef](#)]
4. Wang, Z.; Yao, L.; Cai, Y. Rolling bearing fault diagnosis using generalized refined composite multiscale sample entropy and optimized support vector machine. *Measurement* **2020**, *156*, 107574. [[CrossRef](#)]
5. Prasannamoorthy, V.; Devarajan, N. Fault Detection and Classification in Power Electronic Circuits Using Wavelet Transform and Neural Network. *J. Comput. Sci.* **2011**, *7*, 95–100. [[CrossRef](#)]
6. Wang, C.; Li, H.; Ou, J.; Hu, R.; Hu, S.; Liu, A. Identification of planetary gearbox weak compound fault based on parallel dual-parameter optimized resonance sparse decomposition and improved MOMEDA. *Measurement* **2020**, *165*, 108079. [[CrossRef](#)]
7. Tuncer, E.; Bolat, E.D. Classification of epileptic seizures from electroencephalogram (EEG) data using bidirectional short-term memory (Bi-LSTM) network architecture. *Biomed. Signal Process. Control* **2021**, *73*, 103462. [[CrossRef](#)]
8. Li, R.; Ran, C.; Zhang, B.; Han, L.; Feng, S. Rolling Bearings Fault Diagnosis Based on Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise, Nonlinear Entropy, and Ensemble SVM. *Appl. Sci.* **2020**, *16*, 5542. [[CrossRef](#)]
9. Li, Z.; Chen, J.; Zi, Y.; Pan, J. Independence-oriented VMD to identify fault feature for wheel set bearing fault diagnosis of high speed locomotive. *Mech. Syst. Signal Process.* **2017**, *85*, 512–529. [[CrossRef](#)]
10. Shaomin, Z.; Hong, X.; Binsen, P.; Enrico, Z.; Zhichao, W.; Yingying, J. Feature extraction for early fault detection in rotating machinery of nuclear power plants based on adaptive VMD and Teager energy operator. *Ann. Nucl. Energy* **2021**, *160*, 108392.
11. Huang, Y.; Yan, C.J.; Xu, Q. On the difference between empirical mode decomposition and Hilbert vibration decomposition for earthquake motion records. In Proceedings of the 15th World Conference on Earthquake Engineering, Lisbon, Portugal, 24–28 September 2012.
12. Dragomiretskiy, K.; Zosso, D. Variational Mode Decomposition. *IEEE Trans. Signal Process.* **2014**, *62*, 531–544. [[CrossRef](#)]
13. Civera, M.; Surace, C.A. Comparative Analysis of Signal Decomposition Techniques for Structural Health Monitoring on an Experimental Benchmark. *Sensors* **2021**, *21*, 1825. [[CrossRef](#)] [[PubMed](#)]

14. Yang, W.; Peng, Z.; Wei, K.; Shi, P.; Tian, W. Superiorities of variational mode decomposition over empirical mode decomposition particularly in time–frequency feature extraction and wind turbine condition monitoring. *IET Renew. Power Gener.* **2017**, *11*, 443–452. [[CrossRef](#)]
15. Chen, X.; Yang, Y.; Cui, Z.; Shen, J. Vibration fault diagnosis of wind turbines based on variational mode decomposition and energy entropy. *Energy* **2019**, *174*, 1100–1109. [[CrossRef](#)]
16. Kumar, A.; Zhou, Y.; Xiang, J. Optimization of VMD using kernel-based mutual information for the extraction of weak features to detect bearing defects. *Measurement* **2021**, *168*, 108402. [[CrossRef](#)]
17. Chang, L.; Gang, C.; Xihui, C.; Yusong, P. Planetary Gears Feature Extraction and Fault Diagnosis Method Based on VMD and CNN. *Sensors* **2018**, *5*, 1523.
18. Liang, Z.; Cao, J.; Ji, X.; Wei, P. Fault severity assessment of rolling bearings method based on improved VMD and LSTM. *J. Vibroeng.* **2020**, *6*, 1338–1356.
19. Gai, J.; Shen, J.; Hu, Y.; Wang, H. An integrated method based on hybrid grey wolf optimizer improved variational mode decomposition and deep neural network for fault diagnosis of rolling bearing. *Measurement* **2020**, *162*, 107901. [[CrossRef](#)]
20. Xu, F.; Tse, P.W. A method combining refined composite multiscale fuzzy entropy with PSO-SVM for roller bearing fault diagnosis. *J. Cent. South Univ.* **2019**, *9*, 2404–2417. [[CrossRef](#)]
21. Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. *American journal of physiology. Am. J. Physiol.-Heart Circ. Physiol.* **2000**, *6*, 2039–2049. [[CrossRef](#)]
22. Tylová, L.; Kukul, J.; Hubata-Vacek, V.; Vyšata, O. Unbiased estimation of permutation entropy in EEG analysis for Alzheimer’s disease classification. *Biomed. Signal Process. Control* **2018**, *39*, 424–430. [[CrossRef](#)]
23. Zair, M.; Rahmoune, C.; Benazzouz, D. Multi-fault diagnosis of rolling bearing using fuzzy entropy of empirical mode decomposition, principal component analysis, and SOM neural network. *Proc. Inst. Mech. Eng. Part C* **2019**, *233*, 3317–3328. [[CrossRef](#)]
24. Yin, Y.; Sun, K.; He, S. Multiscale permutation Rényi entropy and its application for EEG signals. *PLoS ONE* **2018**, *13*, 0202558. [[CrossRef](#)]
25. Ceravolo, R.; Civera, M.; Lenticchia, E.; Miraglia, G.; Surace, C. Detection and Localization of Multiple Damages through Entropy in Information Theory. *Appl. Sci.* **2021**, *13*, 5773. [[CrossRef](#)]
26. Civera, M.; Surace, C. An Application of Instantaneous Spectral Entropy for the Condition Monitoring of Wind Turbines. *Appl. Sci.* **2022**, *3*, 1059. [[CrossRef](#)]
27. Rostaghi, M.; Ashory, M.R.; Azami, H. Application of dispersion entropy to status characterization of rotary machines. *J. Sound Vib.* **2019**, *438*, 291–308. [[CrossRef](#)]
28. Jinde, Z.; Junsheng, C.; Yu, Y. A rolling bearing fault diagnosis approach based on LCD and fuzzy entropy. *Mech. Mach. Theory* **2013**, *70*, 441–453.
29. Zhenya, W.; Ligang, Y.; Gang, C.; Jiaxin, D. Modified multiscale weighted permutation entropy and optimized support vector machine method for rolling bearing fault diagnosis with complex signals. *ISA Trans.* **2021**, *114*, 470–480.
30. Jinde, Z.; Haiyang, P.; Jinyu, T.; Qingyun, L. Generalized refined composite multiscale fuzzy entropy and multi-cluster feature selection based intelligent fault diagnosis of rolling bearing. *ISA Trans.* **2021**. [[CrossRef](#)]
31. Yongjian, L.; Hao, S.; Bingrong, M.; Weihua, Z.; Qing, X. Improved multiscale weighted-dispersion entropy and its application in fault diagnosis of train bearing. *Meas. Sci Technol.* **2021**, *32*, 075002.
32. Costa, M.; Goldberger, A.L.; Peng, C.-K. Multiscale entropy analysis of complex physiologic time series. *Phys. Rev. Lett.* **2002**, *6*, 137–140. [[CrossRef](#)] [[PubMed](#)]
33. Wu, S.; Wu, C.; Lin, S.; Wang, C.; Lee, K. Time Series Analysis Using Composite Multiscale Entropy. *Entropy* **2013**, *3*, 1369–1374. [[CrossRef](#)]
34. Gituku, E.W.; Kimotho, J.K.; Njiri, J.G. Cross-domain bearing fault diagnosis with refined composite multiscale fuzzy entropy and the self organizing fuzzy classifier. *Eng. Rep.* **2021**, *3*, 12307. [[CrossRef](#)]
35. Songrong, L.; Wenxian, Y.; Youxin, L. Fault Diagnosis of a Rolling Bearing Based on Adaptive Sparsest Narrow-Band Decomposition and Refined Composite Multiscale Dispersion Entropy. *Entropy* **2020**, *22*, 375.
36. Li, Y.; Gao, X.; Wang, L. Reverse Dispersion Entropy: A New Complexity Measure for Sensor Signal. *Sensors* **2019**, *19*, 5023. [[CrossRef](#)]
37. Ding, S.; Zhao, H.; Zhang, Y.; Xu, X.; Nie, R. Extreme learning machine: Algorithm, theory and applications. *Artif. Intell. Rev.* **2015**, *44*, 103–115. [[CrossRef](#)]
38. GuangBin, H.; Hongming, Z.; Xiaojian, D.; Rui, Z. Extreme learning machine for regression and multiclass classification. *IEEE Trans. Syst. Man Cybern.* **2012**, *42*, 513–529. [[CrossRef](#)]
39. Maglogiannis, I.; Karpouzis, K.; Wallace, M.; Soldatos, J.; Kotsiantis, S.B. Supervised Machine Learning: A Review of Classification Techniques. *Emerg. Artif. Intell. Appl. Comput. Eng.* **2007**, *160*, 249–268.
40. Svetnik, V.; Liaw, A.; Tong, C.; Culberson, J.C.; Sheridan, R.P.; Feuston, B.P. Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1947–1958. [[CrossRef](#)]
41. Wei, Y.; Yang, Y.; Xu, M.; Huang, W. Intelligent fault diagnosis of planetary gearbox based on refined composite hierarchical fuzzy entropy and random forest. *ISA Trans.* **2021**, *109*, 340–351. [[CrossRef](#)]

42. Fezai, R.; Dhibi, K.; Mansouri, M.; Trabelsi, M.; Hajji, M.; Bouzrara, K.; Nounou, H.; Nounou, M. Effective Random Forest-Based Fault Detection and Diagnosis for Wind Energy Conversion Systems. *IEEE Sens. J.* **2021**, *21*, 6914–6921. [[CrossRef](#)]
43. Liu, W.; Cao, S.; Wang, Z. Application of variational mode decomposition to seismic random noise reduction. *J. Geophys. Eng.* **2017**, *14*, 888–898.
44. Xi, C.; Yang, G.; Liu, L.; Jiang, H.; Chen, X. A Refined Composite Multivariate Multiscale Fluctuation Dispersion Entropy and Its Application to Multivariate Signal of Rotating Machinery. *Entropy* **2021**, *23*, 128. [[CrossRef](#)] [[PubMed](#)]
45. Bennasar, M.; Hicks, Y.; Setchi, R. Feature selection using Joint Mutual Information Maximisation. *Expert Syst. Appl.* **2015**, *42*, 8520–8532. [[CrossRef](#)]
46. Cabrera, D.; Sancho, F.; Sánchez, R.; Zurita, G.; Cerrada, M.; Li, C.; Vázquez, R.E. Fault diagnosis of spur gearbox based on random forest and wavelet packet decomposition. *Front. Mech. Eng.* **2015**, *10*, 277–286. [[CrossRef](#)]
47. Qin, X.; Li, Q.; Dong, X.; Lv, S. The Fault Diagnosis of Rolling Bearing Based on Ensemble Empirical Mode Decomposition and Random Forest. *Shock. Vib.* **2017**, *2017*, 2623081. [[CrossRef](#)]
48. LI, Y.; Zhang, J.; Lin, Y. Combining Fisher Criterion and Deep Learning for Patterned Fabric Defect Inspection. *IEICE Trans. Inf. Syst.* **2016**, *11*, 2840–2842. [[CrossRef](#)]
49. Tsapparellas, G.; Jin, N.; Dai, X.; Fehringer, G. Laplacian Scores-Based Feature Reduction in IoT Systems for Agricultural Monitoring and Decision-Making Support. *Sensors* **2020**, *20*, 5107. [[CrossRef](#)]