machines

MDPI

Article

# Bearing Fault Diagnosis with Variable Speed Based on Fractional Hierarchical Range Entropy and Hunter–Prey Optimization Algorithm–Optimized Random Forest

**Jie Ma * and Fangming Liu**

Mechanical Electrical Engineering School, Beijing Information Science and Technology University, Beijing 100192, China
* Correspondence: mjbeijing@163.com

**Abstract:** It is difficult for rolling bearings to realize high-precision fault diagnosis with variable speed. To obtain the features of variable speed fault signal effectively and complete the classification work of high accuracy, robust local mean decomposition (RLMD), fractional hierarchical range entropy (FrHRE), hunter–prey optimization algorithm (HPO) and random forest (RF) are combined. Then the paper advances a model for fault diagnosis based on RLMD, FrHRE and HPO-RF. Firstly, RLMD is selected to reconstruct the signal to eliminate some noise interference in this paper. Secondly, FrHRE is chosen to extract the useful feature. Next step, HPO is used to optimize the important parameters of RF and enhance RF's classification ability. Finally, these obtained features are imported into the optimized RFmodel to achieve the classification. The experimental data is provided by University of Ottawa. The experiment analysis demonstrates that the proposed method performs very well in classification.

**Keywords:** fractional hierarchical range entropy; variable speed condition; hunter-prey optimization algorithm; random forest



**Citation:** Ma, J.; Liu, F. Bearing Fault Diagnosis with Variable Speed Based on Fractional Hierarchical Range Entropy and Hunter–Prey Optimization Algorithm–Optimized Random Forest. *Machines* **2022**, *10*, 763. https://doi.org/10.3390/machines10090763

Academic Editor: Davide Astolfi

Received: 2 August 2022
Accepted: 31 August 2022
Published: 2 September 2022



## 1. Introduction

As a core component, the health of rolling bearings has a huge influence on performance and safety of rotating machinery. Due to the complexity of the rolling bearings' own structure and the harsh operating environment, they will inevitably have various faults, affecting the safety of rotating machinery. Therefore, the most effective way to avoid bearing faults causing significant economic losses and casualties is to carry out bearing fault diagnosis as soon as possible, and then repair and replace the damaged bearings [1–5]. However, most methods for rolling bearing fault diagnosis are applied to constant speed conditions. In contrast, there has been less research on variable speed conditions, which are common conditions in rotating machinery [6]. When rotating machinery is operating at variable speeds, the signal frequency will change greatly with time, the spectrum will become blurred and the signal will be strongly non-stationary, which makes the processing of vibration signals very challenging. Therefore, variable speed conditions are receiving extensive attention in the fault diagnosis research of rotating machinery.

Because the vibration signals are easily influenced by various noises during the acquisition process, it is an important premise to find an excellent signal preprocessing method to ensure the effective fault feature extraction. In this respect, many researchers have executed extensive research. Among them, Dragomiretskiy et al. [7] proposed variational mode decomposition (VMD). However, the number of modes for VMD must be set in advance, which can make VMD lack of adaptability. In view of this shortcoming, researchers have studied adaptive variational mode decomposition (AVMD) and successive variational mode decomposition (SVMD), respectively [8,9]. Although these methods have reached some achievements, they must cooperate with other methods to complete signal preprocessing in the face of signal processing under variable speed conditions. However, the local mean decomposition (LMD) method proposed by Li [10] can be directly applied to the nonlinear signals and then combined with the envelope spectrum to achieve the identification of bearing fault features. In order to improve the ability of the local mean decomposition (LMD), the ensemble local mean decomposition (ELMD), the complete ensemble local mean decomposition (CELMD) and the complete ensemble local mean decomposition (CELMDAN) have been proposed [11–13]. Although these methods have achieved certain results, they still suffer from poor adaptability, modal mixing, end effects and so on. The robust local mean decomposition (RLMD) has solved the problems of the above methods [14]. Moreover, RLMD combined with excellent time-frequency analysis tools can accurately extract features of variable speed signals [15]. Therefore, this paper chooses RLMD as the tool to reduce noise.

After using RLMD, the useful fault information is extracted from the denoised signal by finding suitable features. Due to the inevitable friction during the bearing operation, the signal will show non-stationary and nonlinear characteristics. At this time, the entropy index is selected to extract different faults' features. For example, Zhang et al. [16] studied the multi-scale entropy (MSE). Li et al. [17] proposed refined composite multiscale fuzzy entropy (RCMFE) to extract fault features, which are hidden in denoised signals. Omidvarnia et al. [18] presented the concept of range entropy. Multiscale range entropy (MRE) was proposed on the basis of range entropy and successfully applied to bearing fault diagnosis [19,20]. Combined with the advantages of hierarchical analysis, the hierarchical range entropy (HRE) index was proposed [21]. However, this index is only applied to constant speed conditions. Due to the complex time-varying modulation and spectral structure, it is very difficult for rolling bearing under variable speed conditions to extract fault features. To address the problem, a new index, called fractional hierarchical range entropy (FrHRE), is presented. Compared with the original HRE, FrHRE reflects the features of time, frequency and time-frequency domain at different scales. It can fully express the signal information for the sake of extracting multi-angle and deep-level signal features with variable speeds.

In addition to excellent fault feature extraction methods, it also needs to be matched with a good fault classifier so as to achieve high-precision classification under variable speed conditions. In recent years, many machine-learning methods have been applied on fault classification problems, for example support vector machine (SVM), extreme learning machine (ELM), kernel extreme learning machine (KELM), least squares support vector machine (LSSVM) and random forest (RF) [22–28]. Due to the advantages of high classification accuracy and high calculation efficiency, RF is often applied to fault diagnosis for fault identification. Han et al. [29] used the RF classifier to achieve accurate classification of rolling bearing faults. Vakharia et al. [30] selected the RF to identify the ball bearing fault and realized high-precision classification. It can be observed that the classification performance of RF is excellent. However, there are many parameters that need to be tuned in RF model, which will affect the classification accuracy of RF to a certain extent. To solve this problem, researchers often use swarm intelligence optimization algorithms, which are very straightforward and easy to comprehend [31]. For example, particle swarm optimization algorithm, artificial fish swarm algorithm and grey wolf optimization have been adopted to choose RF's optimal parameters [32–34]. The classification accuracy of the optimized RF is significantly improved. These swarm intelligence algorithms have been applied to RF's optimization and attained certain results. However, their exploration and development capabilities are not very good. As a result, the hunter–prey optimization algorithm (HPO) has been proposed [35]. The algorithm simulates the behavior of hunters and prey and has sufficient exploration and exploitation capabilities to find the optimal parameters adaptively. Therefore, this paper has applied HPO to the RF and adopted HPO-RF model for fault identification.

In conclusion, the paper presents a method for fault diagnosis with variable speed based on RLMD, FrHRE and HPO-RF. Notably, the paper make a summary on main contributions as follows:

1. A signal preprocessing method (RLMD) that can effectively remove noise from the variable speed signal is adopted.
2. A new feature extraction method applied to variable speed bearing signals, namely fractional hierarchical range entropy (FrHRE), is proposed in this paper.
3. This paper investigates an adaptive optimization-seeking fault identification model on the foundation of RF model. RF model parameters are globally optimized by iterative algorithm.

The rest of the paper is represented as detailed below: Section 2 presents the basic theory of RLMD, FrFT, HRE, RF and HPO. The paper provides the explicit steps of FrHRE, presents the steps of HPO-RF and shows the steps of proposed method in Section 3. Section 4 analyzes the performance of RLMD, FrHRE and HPO-RF, respectively. Section 5 discusses the experimental results. Finally, some conclusions are given in Section 6.

## 2. Basic Theory

### *2.1. Robust Local Mean Decomposition (RLMD)*

Robust Local Mean Decomposition (RLMD) can extract pure FM signal, envelope signal and their product function (PFs) from any complex signal $y(n)$ to be analyzed, so that PFs can fully contain the multi-scale information of the original signal. Next, we will briefly introduce the process of extracting PFs by RLMD algorithm.

Step 1: Find the local extremes of $y(n)$, and calculate the smooth local mean $me(n)$ and smooth local amplitude $am(n)$.

Step 2: The estimated zero-mean signal $hh_{11}(n)$ and the FM signal $ss_{11}(n)$ are calculated by Equation (1). The subscripts represent the $i$th PF and the $j$th sifting process.

$$\begin{aligned} hh_{11}(n) &= y(n) - me_{11}(n) \\ ss_{11}(n) &= \tfrac{hh_{11}(n)}{am_{11}(n)} \end{aligned} \tag{1}$$

Continue to repeat the above steps $R$ times with $ss_{11}(n)$ as a new signal until the sifting conditions proposed by RLMD ($of_{ij} < of_{ij+1} < of_{ij+2}$) are met.

If the above sifting conditions are met, it can be proved that $ss_{1R}(n)$ is a pure FM signal, namely:

$$PF_1(n) = am_1(n) \times ss_1(n) \tag{2}$$

Step 3: Extract the remaining $PF_i$ from $y(n)$. $y(n)$ can be expressed by Equation (3). Where $u_L(n)$ is the residual signal after $L$ repetitions.

$$y(n) = \sum_{i=1}^{L} PF_i(n) + u_L(n) \tag{3}$$

### *2.2. Fractional Fourier Transform (FrFT)*

Performing FrFT on the signal is to rotate the signal counterclockwise on the time axis by $\beta$ angle to the $u$axis and then perform the Fourier transform. Figure 1 shows the process of using FrFT to rotate the $t - f$ plane to the $u - v$ plane.

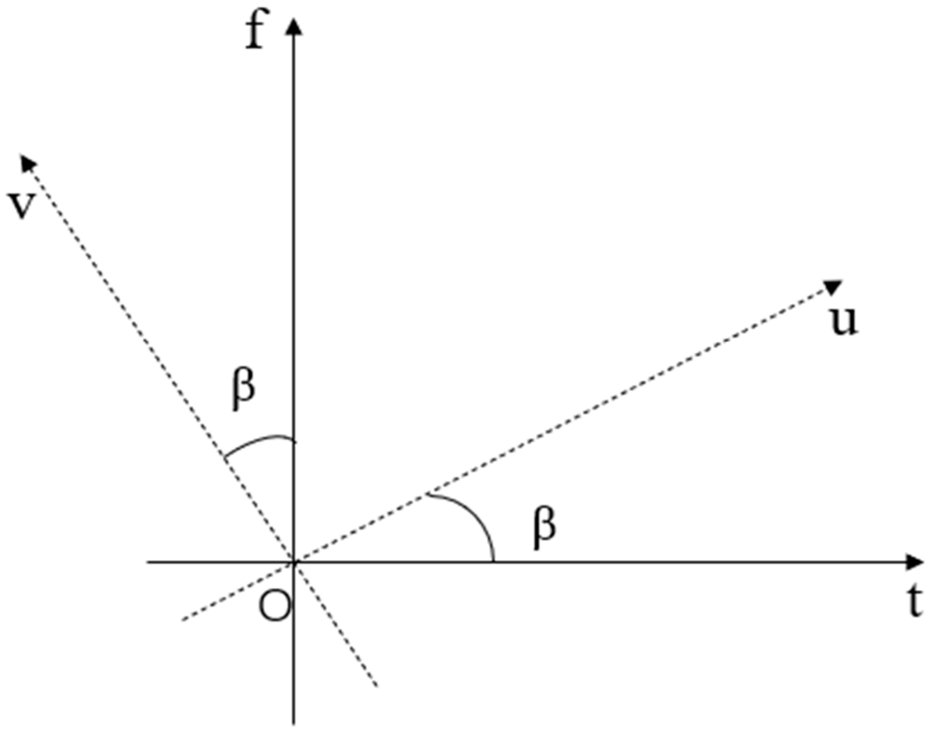



**Figure 1.** The $t-f$ plane is rotated to the u–v plane by FrFT.

The FRFT of signal $y(t)$ is defined as:

$$Y_\beta(u) = F^p[y(t)] = \int_{-\infty}^{\infty} y(t) K_\beta(t,u) dt \tag{4}$$

where $p$ is the order of FRFT, which can change from 0 to 1, meaning that the signal changes from time domain to frequency domain gradually. Moreover, the kernel function of FRFT is:

$$K_\beta(t,u) = \begin{cases} \sqrt{\dfrac{1-j\cot\beta}{2\pi}} \exp(j\dfrac{t^2+u^2}{2}\cot\beta - tu\csc\beta) & \beta \neq n\pi \\ \delta(t-u) & \beta = 2n\pi \\ \delta(t+u) & \beta = (2n\pm 1)\pi \end{cases} \tag{5}$$

Its transformation formula is shown in Equation (6), where the range of $\beta$ is generally $0 \sim \dfrac{\pi}{2}$.

$$\left.\begin{array}{l} u = t\cos\beta + f\sin\beta \\ v = -t\sin\beta + f\cos\beta \end{array}\right\} \tag{6}$$

*2.3. Hierarchical Range Entropy (HRE)*

HRE combines hierarchical decomposition and range entropy index, which can mine signal feature information from multi-scale aspects. Assuming a time series $Y = \{y_1, y_2, \ldots, y_N\}$ is given, the process of calculating its HRE is as follows:

First, we define two operators: the average operator $Q_0$ and the difference operator $Q_1$, the formula is shown in Equation (7).

$$\begin{aligned} Q_0(y) &= \frac{y_i + y_{i+1}}{2} \quad i = 1, 2, \ldots\ldots, N-1 \\ Q_1(y) &= \frac{y_i - y_{i+1}}{2} \quad i = 1, 2, \ldots, N-1 \end{aligned} \tag{7}$$

where $Q_0$ represents the low-frequency component of time series and $Q_1$ represents its high-frequency component [36].

Secondly, given a integer $e$ that is non-negative, there is only one vector $[v_1, v_2, \ldots, v_k]$, $k \in N$, which can represent the non-negative integer $e$, as shown in Equation (8).

$$e = \sum_{m=1}^{k} 2^{k-m} v_m \tag{8}$$

Then, the operator should be used repeatedly in the hierarchical decomposition of time series. The hierarchical decomposition components can be attained by using Equation (9).

$$Y_{n,e} = Q_{v_n} \bullet Q_{v_{n-1}} \bullet \ldots \bullet Q_{v_1}(y) \tag{9}$$

Finally, the range entropy is calculated for the hierarchical components obtained in the above procedures, and the hierarchical range entropy can be obtained.

$$HRE(y, n, e, m, \tau) = RangeEn(V_{n,e}, m, \tau) \tag{10}$$

### 2.4. Random Forest (RF)

RF is a machine learning algorithm, which can be used to classify, cluster and regress. The classification of RF is achieved by training base decision trees, generating models and using the comprehensive results of many decision trees to vote. The content of RF is roughly as follows [37,38]:

Step 1: Select $N$ samples randomly as training sample set $X^* = \{X_1^*, X_2^*, \ldots, X_N^*\}$ by the bootstrap sampling method. A decision tree can be obtained by training using this sample set, as shown in Equation (11).

$$\begin{aligned} &h_i(X^*, \Theta_k) = c(x_1, x_2, \ldots, x_N, root(h_i)) \\ &c(x_1, x_2, \ldots, x_N, h_t) = \begin{cases} label(h_t) & h_t \text{ is the leaf node} \\ c(x_1, x_2, \ldots, x_N, h_t) & h_t \text{ is the inner node} \end{cases} \end{aligned} \tag{11}$$

where $root(h_i)$ is the root node of the decision tree $h_i(X^*, \Theta_k)$, and $c(x_1, x_2, \ldots, x_N, h_t)$ is the splitting criterion of decision tree.

Step 2: When each sample has $M$ features, select $m(m \ll M)$ features randomly, and input the best one at each split node of the decision tree for splitting.

Step 3: Repeat the second step and continue splitting until all training samples of this node belong to the same class. There is no pruning behavior in the whole formation of the decision trees.

Step 4: Build multiple decision trees in accordance with the first three steps, which then can constitute a random forest.

Step 5: After the test samples are input into RF model, the class with the most classification results is selected as the final result. The classification decision model $H(x)$ is exhibited in Equation (12), where it represents the output tag variable and $I(*)$ is the indicative function.

$$H(x) = \arg\max_{Y} \sum_{i=1}^{N} I(h_i(X^*, \Theta) = Y) \tag{12}$$

*2.5. Hunter–Prey Optimizer (HPO)*

HPO is a new algorithm for optimizing proposed in 2021. These sections are described in detail below. First, Equation (13) are used to set the initial population $(\overrightarrow{XX}) = \left\{\overrightarrow{XX_1}, \overrightarrow{XX_2}, \ldots, \overrightarrow{XX_n}\right\}$. Then, the fitness of each scheme is calculated by $O_i = f(\overrightarrow{XX})$.

$$XX_i = rand(1, d).*(uub - llb) + llb \tag{13}$$

where $XX_i$ represents the position of the hunter or prey, $llb$ represents the lower limits of the problem variables and $uub$ represents its upper limits. $d$ is the number of these variables.

*(1) The hunter search mechanism*

In HPO, the hunter search mechanism is represented by Equation (14).

$$XX_{i,j}(t+1) = XX_{i,j}(t) + 0.5\left[\left(2CZP_{pos(j)} - XX_{i,j}(t)\right) + \left(2(1-C)Z\mu_{(j)} - XX_{i,j}(t)\right)\right] \tag{14}$$

where $XX(t)$ is the current position and $XX(t+1)$ is the next position of the hunter. $\mu$ is the mean value of all positions, that is, $\mu = \frac{1}{n}\sum_{i=1}^{n}\overrightarrow{XX_i}$. $C$ is the balance parameter, which is calculated by $C = 1 - it(\frac{0.98}{MAXIt})$. $Z$ is the adaptive parameter, as Equation (15) shows. $P_{pos}$ is the position of the prey, which is shown in Equation (16).

$$Z = R_2 \otimes IDEX + \vec{R_3} \otimes (\sim IDEX) \tag{15}$$

$$\vec{P}_{pos} = \overrightarrow{XX_i} \,|\, i \text{ is sorted } D_{euc}(kbest) \tag{16}$$

*(2) The prey escaped to a safe position*

When the prey is attacked, it will escape to the safe position immediately. At this time, Equation (17) represents the update of the prey's position.

$$XX_{i,j}(t+1) = T_{pos(j)} + CZ\cos(2\pi R_4) \times (T_{pos(j)} - XX_{i,j}(t)) \tag{17}$$

where $XX(t)$ and $XX(t+1)$ are the current position and the next position of the prey, respectively. $T_{pos}$ stands for the optimal global location. $R_4$ is a number randomly selected from $[-1, 1]$.

To solve the problem of how to distinguish the object represented by $XX_i$, HPO proposes an adjustment parameter $\beta$ (its value is 0.1) and a random number $R_5$ (its range is $[0, 1]$). When $R_5 < \beta$, $XX_i$ is regarded as the hunter. Otherwise, $XX_i$ is regarded as the prey.

## 3. Proposed Method

*3.1. Fractional Hierarchical Range Entropy (FrHRE)*

A new feature extraction method, fractional hierarchical range entropy (FrHRE), is proposed by combining FrFT, hierarchical decomposition and range entropy in this paper. From different time-frequency perspectives, FrHRE retains the strength of multi-scale decomposition and adds high frequency components in different scales to make the extracted features contain more information. Figure 2 is the flow chart for calculating FrHRE. Firstly, FrFT is applied to the time series to obtain the time-frequency components of different orders. The parameters are then initialized before calculating the HRE. Next, calculate the HRE for each time-frequency component. Finally, the output is HRE of each time-frequency component. That is, FrHRE is obtained.

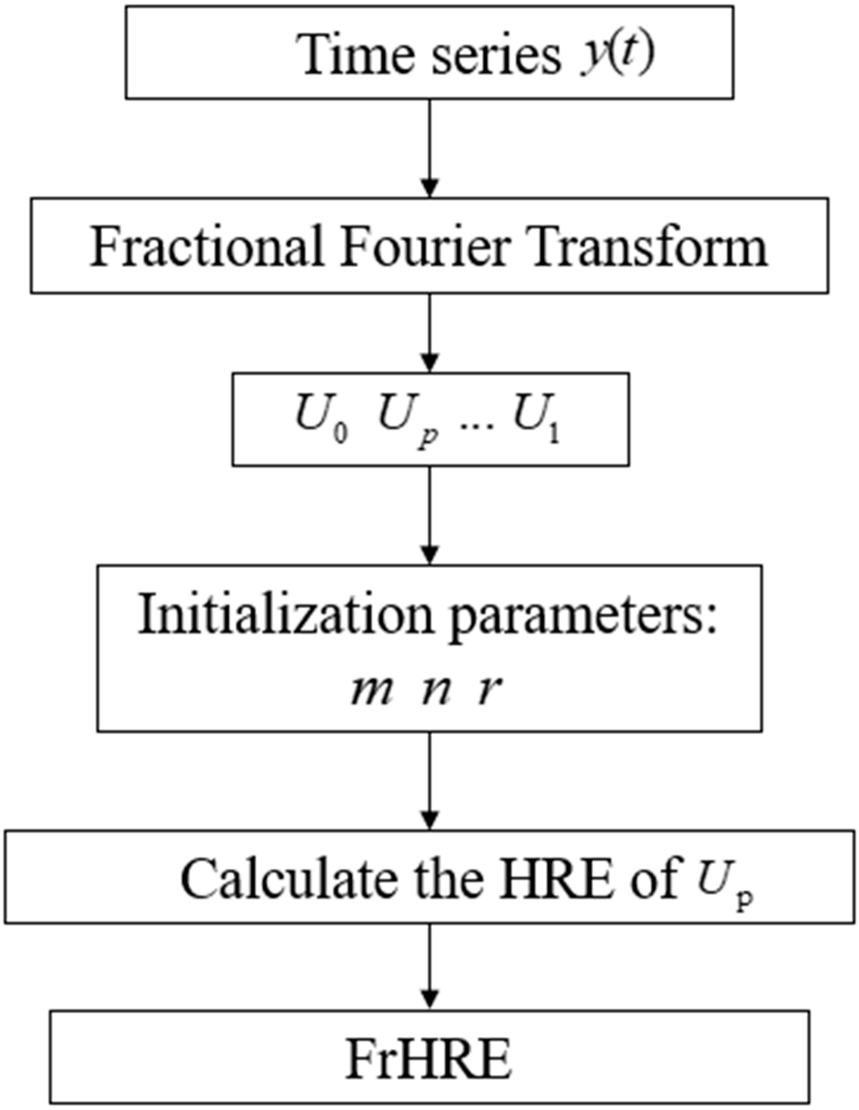



**Figure 2.** The flow chart of calculating FrHRE.

*3.2. Hunter–Prey Optimizer-Random Forest (HPO-RF)*

In most cases, the number of base decision trees (ntrees) and minimum node-divided samples (n_splits) in RF are selected empirically, which may increase the classification error of RF. Therefore, this paper selects these two parameters as tuning objects to decrease the classification error as much as possible and enhance the RF's accuracy. The optimization process of HPO-RF is as follows.

Input: training dataset, test dataset, the upper and lower limits of RF parameters $uub$ and $llb$, the maximum number of iterations $MAXIt$, the number of populations $nPop$, the number of RF parameters $dm$ and the adjustment parameter $\beta$.

Output: the optimal solutions ***Target*** and ***TargetScore***, which correspond to the optimal parameters $ntrees_{best}$ and $n\_splits_{best}$.

(1) According to the input parameters, the initial population is randomly set.
(2) Read the input training data and test data.
(3) Calculate the objective function of all members in the population and store the optimal fitness value found so far. (In this paper, we choose the sum of the classification error rates of training and test set as the fitness value, so the smaller fitness value is considered as the better one.)
(4) Update parameter: $C$, $k_{best}$, $Z$ and set $\vec{R}_1$, $\vec{R}_3$, $R_2$, $R_4$, $R_5$.
(5) Update the position of the current search agent using Equation (18).

$$\begin{aligned} XX_{i,j}(t+1) &= XX_{i,j}(t) + 0.5[(2CZP_{pos(j)} - XX_{i,j}(t)) + (2(1-C)Z\mu_{(j)} - XX_{i,j}(t))] \\ & \qquad R_5 < \beta \\ XX_{i,j}(t+1) &= T_{pos(j)} + CZ\cos(2\pi R_4) \times (T_{pos(j)} - XX_{i,j}(t)) \\ & \qquad R_5 > \beta \end{aligned} \tag{18}$$

### 3.3. The Proposed Method

Figure 3 exhibits the flow chart of the proposed method.

Step 1: The bearing signal is decomposed by RLMD. Moreover, the optimal components are chosen for signal reconstruction by using the principle of the maximum cross-correlation, which can reduce noise and highlight features to a certain extent.

Step 2: FrHRE is calculated separately for the denoised signal of each state, and a feature sample set is constructed. The parameters of FrHRE are set as: $p \in [0, 0.1, \ldots, 1]$, the number of layers is 3, the embedding dimension is 2 and $r$ is 0.15 times of the sample's standard deviation. Eleven components in the fractional Fourier domain (FrFD) can be obtained by performing FrFT of the corresponding order on the denoising signal. Where $p = 0$ is the time domain of signal and $p = 1$ is frequency domain, and the rest orders correspond to the signal's time domain representation at different angles. Moreover, the range entropies at different levels are calculated for the obtained components in FrFD. Then the features in time, frequency and time-frequency domain are obtained.

Step 3: The samples of training and test are obtained by dividing extracted feature samples randomly, and the corresponding labels are given.

Step 4: The RF parameters that need to be optimized and their ranges are determined. Moreover, the input of HPO-RF are the sample set.

Step 5: RF is optimized using HPO. First, the parameters of the HPO are initialized. Secondly, the fitness values of all members in the initial population are computed (the classification error rate of the training and test samples is chosen as the fitness value). Next, the parameter values with the smallest fitness value are chosen as the current optimal parameters. After that, the parameters should continue to be update and agent's current location should be searched for. Finally, when the maximum of iterations number is reached, the outputs are current parameters and the corresponding fitness values. If not, the target parameters should be updated and training continued until the iteration numbers reach the maximum.

Step 6: These obtained optimal parameters are input into RF for training and the classification of the test samples is completed.

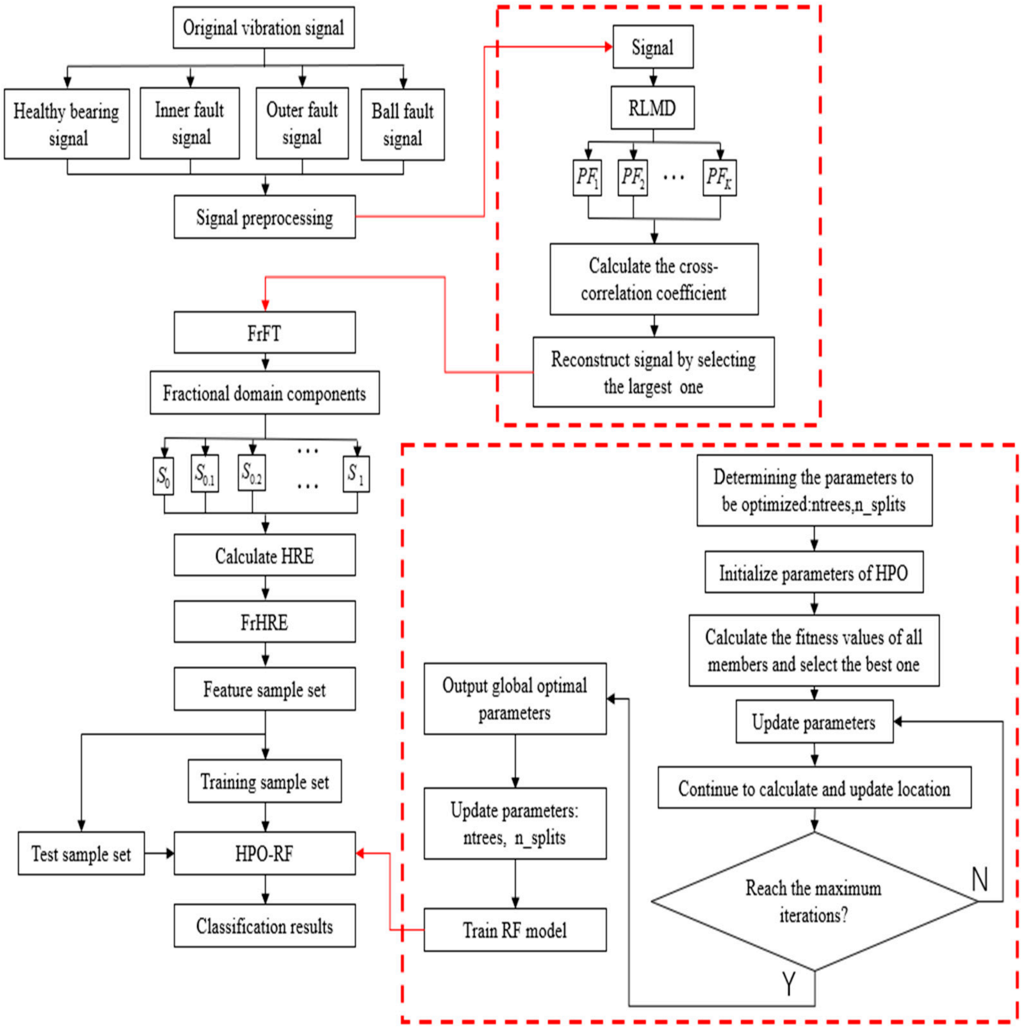



**Figure 3.** The flow chart of RLMD, FrHRE and HPO-RF.

## 4. Performance Analysis

### *4.1. Performance Analysis of RLMD*

To analyze the performance of the RLMD, the simulated fault signal model of rolling bearing under acceleration state is constructed, as shown in Equation (19) [39].

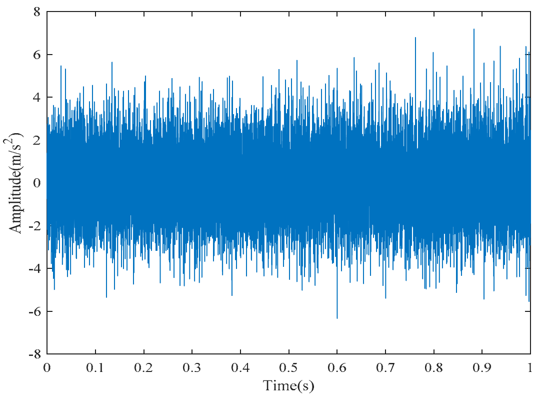
$$\begin{cases} s_1(t) = \sum\limits_{i=1}^{M} C_i e^{(-200(t-t_i))} \sin(2\pi \times 3000(t-t_i))u(t-t_i) \\ s_2(t) = \sum\limits_{i=1}^{M} C_i e^{(-200(t-t_i))} \sin(2\pi \times 3500(t-t_i))u(t-t_i) \\ s_3(t) = \sum\limits_{i=1}^{M} C_i e^{(-200(t-t_i))} \sin(2\pi \times 4000(t-t_i))u(t-t_i) \\ s(t) = s_1(t) + s_2(t) + s_3(t) + n(t) \end{cases} \tag{19}$$

where $s(t)$ is the fault simulation signal of the rolling bearing under variable speed; $n(t)$ is the white Gaussian noise with $SNR = -5dB$; $C_i = c + \lambda * f(t_i)$ represents the amplitude of the $i_{th}$ impact, and both $c$ and $\lambda$ are constants; $f(t_i)$ represents the rotation frequency of the bearing at time $t_i$; and $t_i$ is the time coordinate corresponding to the $i_{th}$ impact, which is calculated in Equation (20).

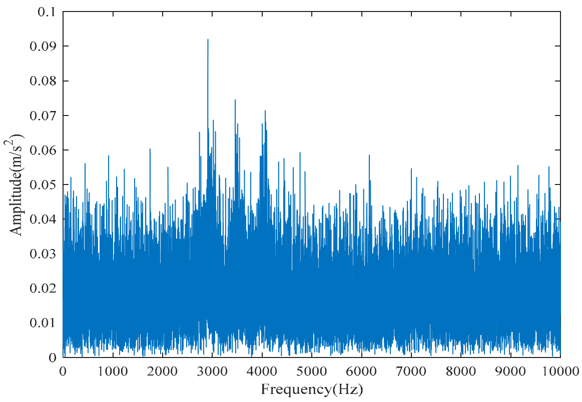
$$\begin{cases} t_1 = (1+\tau)/f(t_0)/FCC \\ t_i = (1+\tau)[1/f(t_0) + 1/f(t_1) + \cdots + 1/f(t_{i-M})]/FCC \;\; i = 2,3,\cdots,M \\ t_0 = 0 \end{cases} \tag{20}$$

In Equation (20), $\tau$ is the sliding error coefficient of the rolling bearing, generally $\tau = 0.01$, and $FCC$ represents the number of fault impact caused by each revolution of the bearing, which is not affected by the rotation speed.

The fault signal model parameters of rolling bearing are as below: the sampling frequency is 204,800 Hz; the sampling points' number is 204,800; $C_i = 1 + 0.05 * f(t_i)$; $f(t) = 10t + 10$; and $FCC_{outer}$ is fault characteristic coefficient of outer race of bearing, that is $FCC_{outer} = 3.7$. At this time, the waveform and spectrum of the simulated outer race fault signal are both exhibited in Figure 4. It can be seen that due to the influence of noise, the frequency range of the components of simulated signal are not clear in the spectrum.



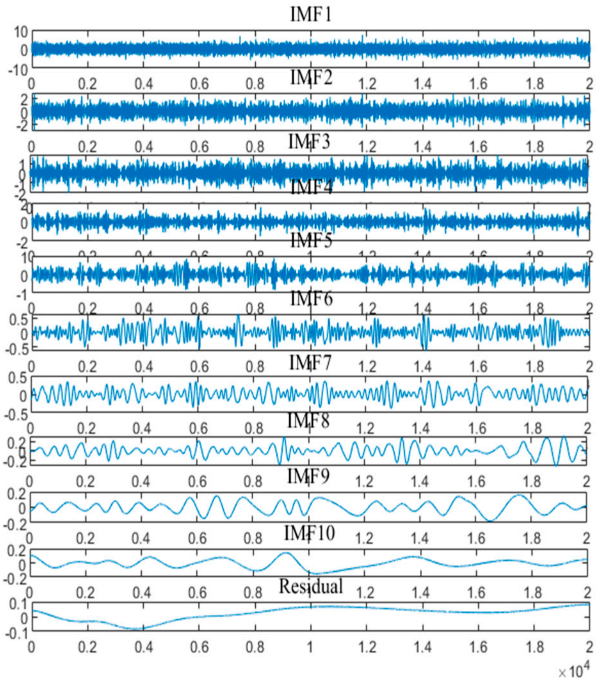

(**a**) The waveform



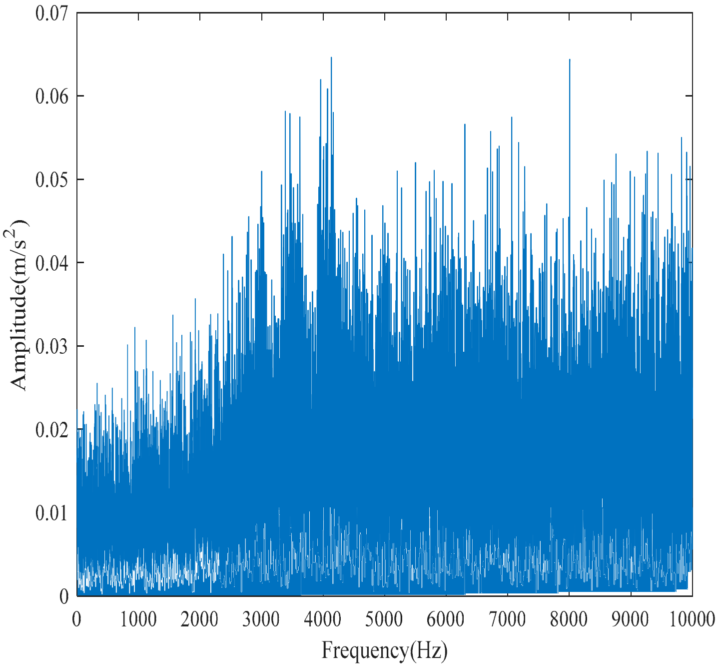

(**b**) The spectrum

**Figure 4.** The waveform and the spectrum diagram of simulated signal.

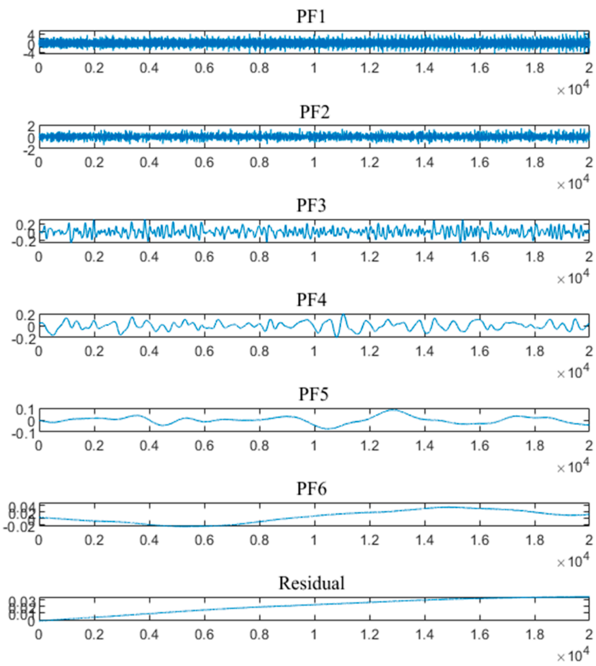
In this regard, CEEMDAN, LMD and RLMD are selected for comparison in terms of noise reduction performance. Firstly, these three decomposition methods are, respectively, applied to the simulated signal, and the component diagrams obtained by decomposition are shown in Figure 5a,c,e. By observing these three subgraphs, we can see that the first PF component obtained by RLMD reveals more impact. However, there is still a large amount of noise in the IMF1 obtained by CEEMDAN and the PF1 obtained by LMD. Then, the cross-correlation coefficients between these components and the simulated signal are calculated, respectively. The component with the largest cross-correlation coefficient under each decomposition method is selected as the denoised signal. The spectra of these three denoised signals obtained are shown in Figure 5b,d,f. We can clearly see that the frequency range cannot be clearly located in Figure 5b. From Figure 5d, the frequency range of the signal components can only be distinguished roughly, and the spectrum still includes a lot of interference. In Figure 5f, the frequency range of three signal components can be distinctly observed, and most of the noise has been eliminated. Therefore, we can conclude that RLMD not only shows an excellent decomposition effect but also exhibits superior performance in noise reduction when performing denoising processing on the bearing vibration signal at variable speeds.



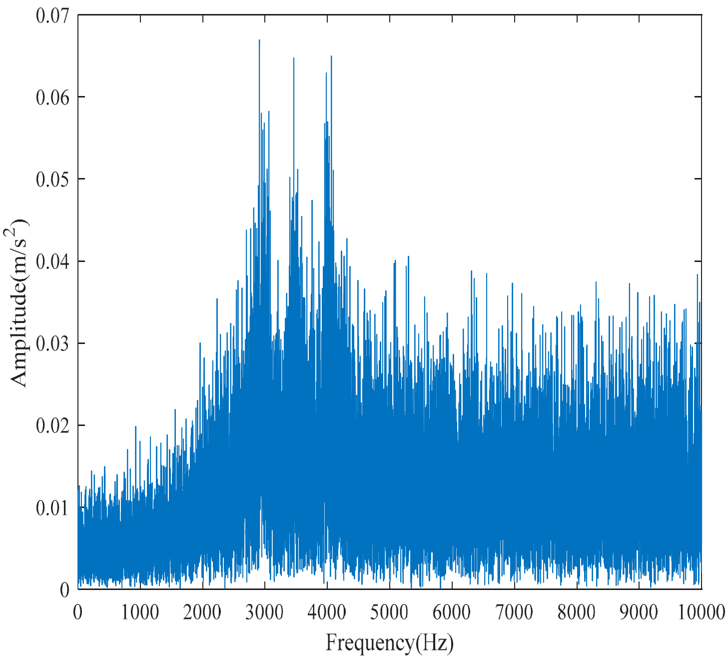

(**a**) The decomposition results of CEEMDAN




(**b**) The processed signal's spectrum obtained by CEEMDAN

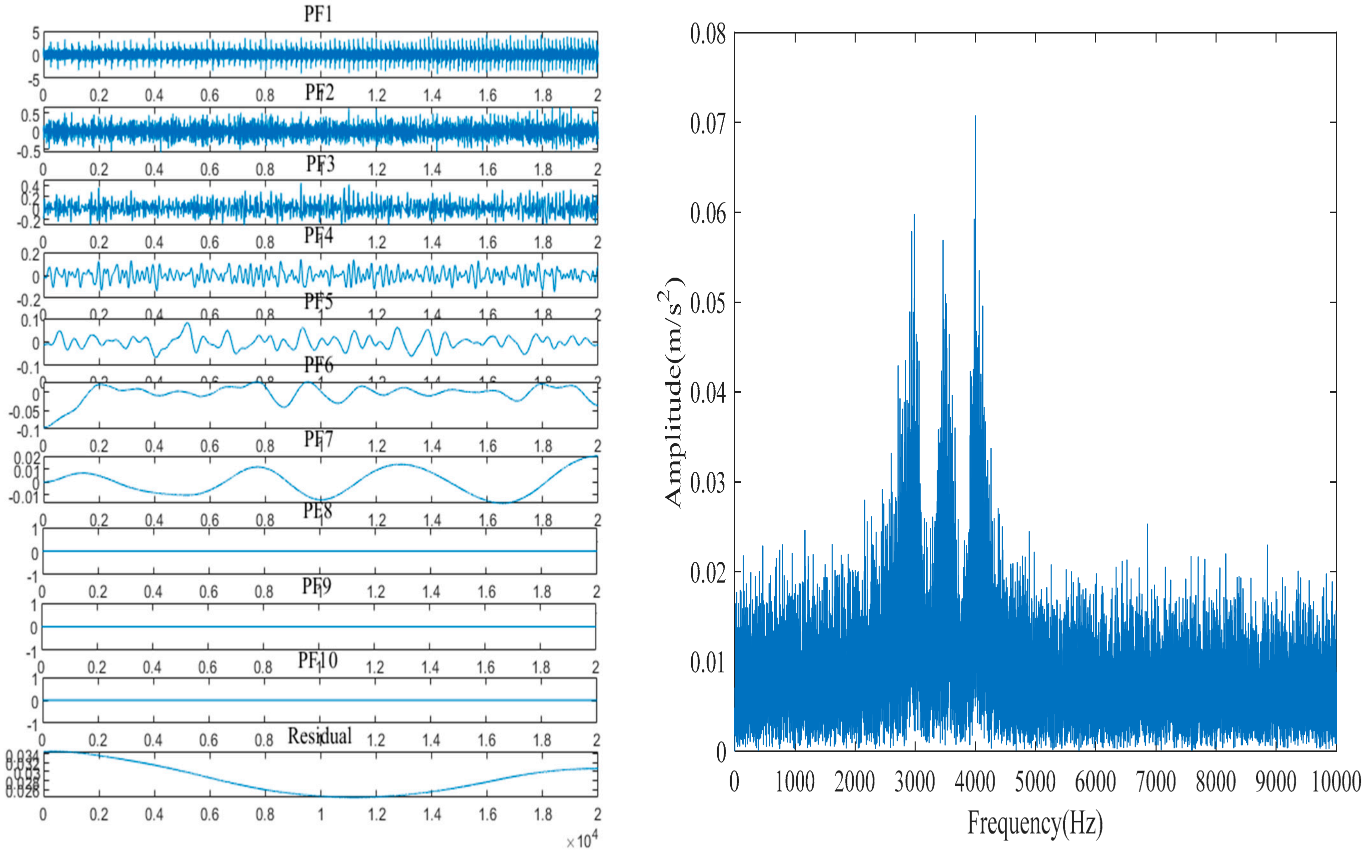



(**c**) The decomposition results of LMD




(**d**) The processed signal's spectrum obtained by LMD

**Figure 5.** *Cont.*

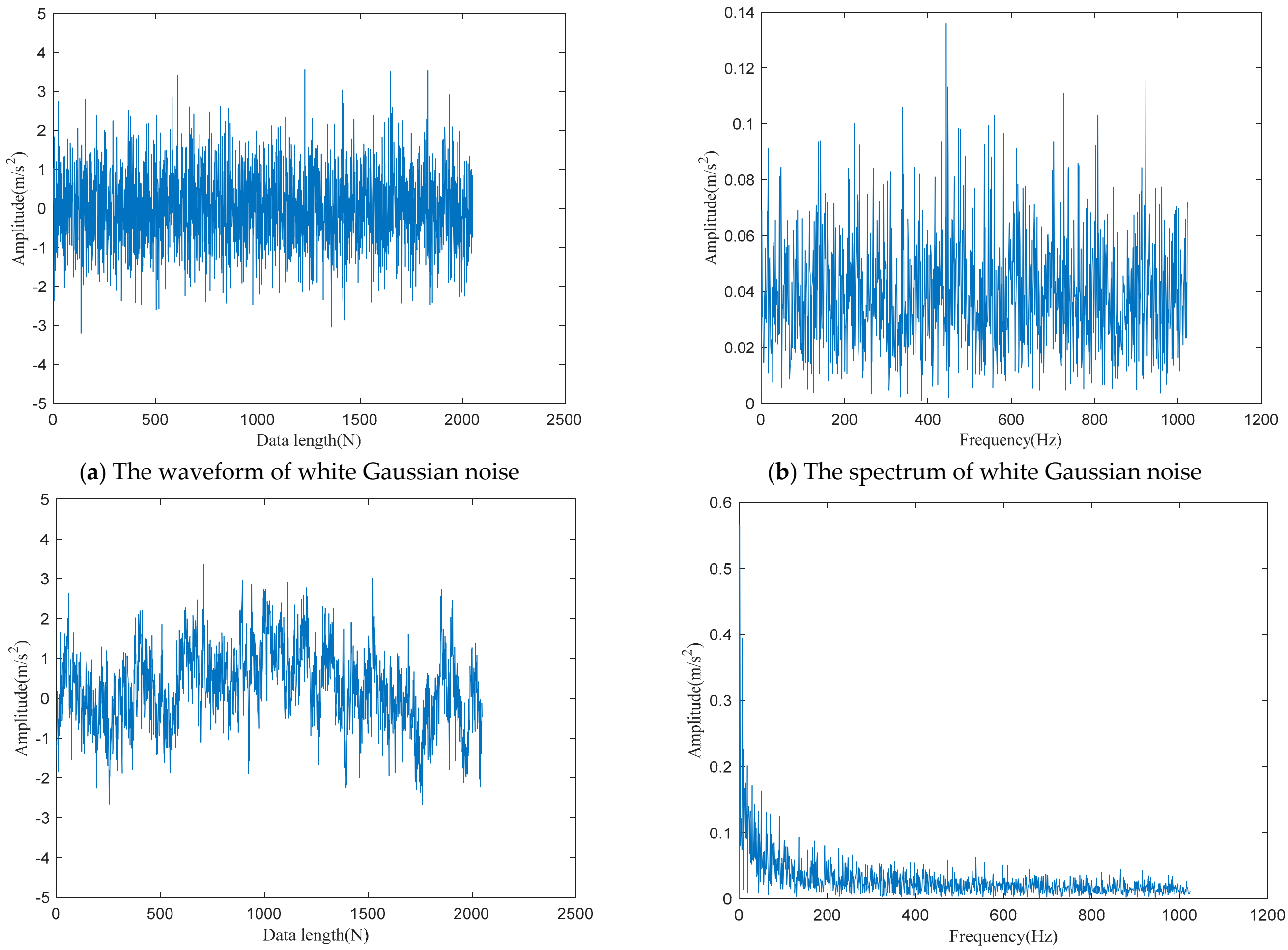



(**e**) The decomposition results of RLMD (**f**) The processed signal's spectrum obtained by RLMD

**Figure 5.** The decomposition results attained by CEEMDAN, LMD and RLMD separately and the spectra of denoised signals obtained through them.

### *4.2. The Analysis of FrHRE*

$1/f$ noise and white Gaussian noise (WGN) are commonly applied to verify the performance of entropy [40]. Therefore, to demonstrate the superiority of FrHRE, these two noise signals are simulated in this paper, and their sampling points are set to 2048. Figure 6 shows their waveform and spectrum. It can be observed from the figure that white Gaussian noise is evenly distributed in the frequency band, while the spectrum of $1/f$ noise is mainly centered in the low frequency area, which means that the periodic components with strong correlation are mainly distributed in the low frequency band of $1/f$ noise, that is, the signal at the low frequency band is more ordered and has a lower entropy value.

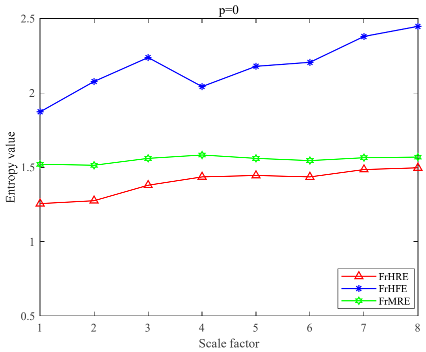



(**a**) The waveform of white Gaussian noise (**b**) The spectrum of white Gaussian noise

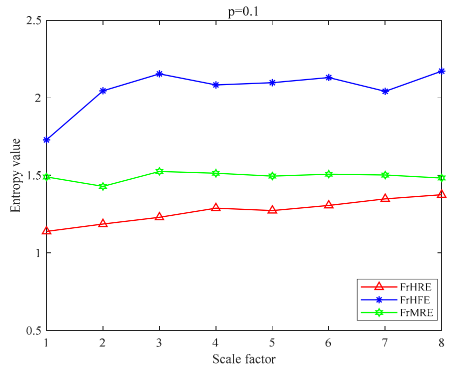
(**c**) The waveform of $1/f$ noise (**d**) The spectrum of $1/f$ noise

**Figure 6.** The time domain waveform and the spectrum of different simulated noise signal.

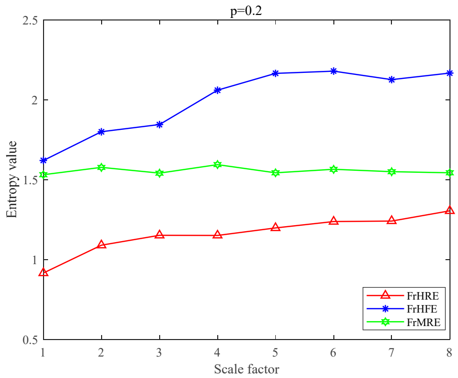
In Figure 7, when these three methods are applied to $1/f$ noise, the entropy curves of FrHRE and FrHFE show an upward trend with the increasing of scale (i.e., the entropy value is small in low scale and large in high scale), and the entropy value curve of FrMRE remains basically unchanged. This illustrates that compared with FrMRE, which only reflects the signal's low-frequency components, FrHRE and FrHFE can reflect the information not only low-frequency but high-frequency, which makes their entropy features contain more comprehensive information. In addition, the paper uses FrHRE, FrHFE and FrMRE to process 20 sets of white Gaussian noise generated randomly so as to verify the stability of FrHRE. Their standard deviations at different orders corresponding to different scales are then calculated by averaging 20 realizations. It can be seen from Figure 8 that the standard deviation of FrHRE in each scale is less than that of FrHFE. In addition, in some orders, the standard deviation of FrMRE is occasionally smaller than that of FrHRE. However, on the whole, the standard deviation curve of FrHRE is always relatively stable and the value of its standard deviation is relatively small no matter its location in the low scale or the high scale. This indicates that FrHRE has excellent stability performance and is superior to the other two in stability. Considered comprehensively, FrHRE is qualified in terms of reflecting the comprehensiveness of signal information and high stability.



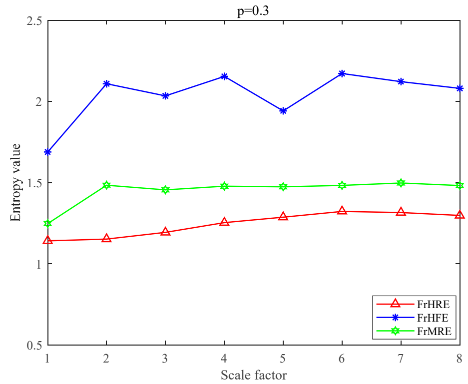

(**a**) Three entropy curves for $1/f$ noise at $p = 0$



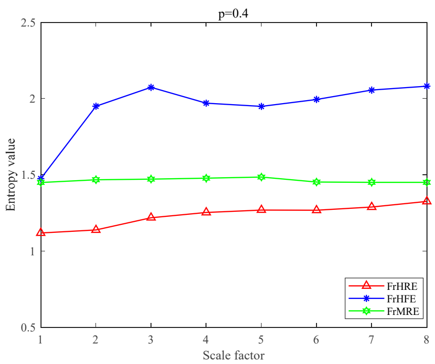

(**b**) Three entropy curves for $1/f$ noise at $p = 0.1$



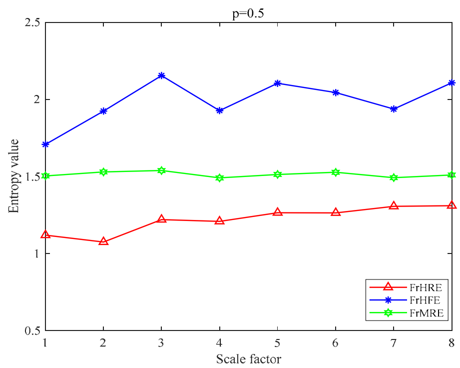

(**c**) Three entropy curves for $1/f$ noise at $p = 0.2$



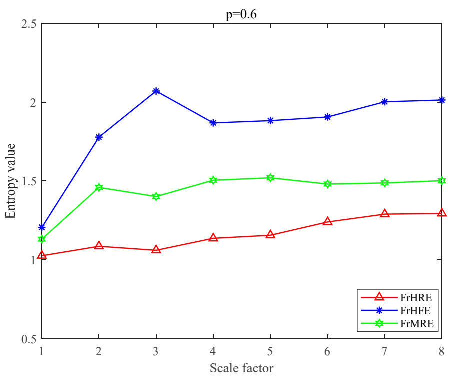

(**d**) Three entropy curves for $1/f$ noise at $p = 0.3$



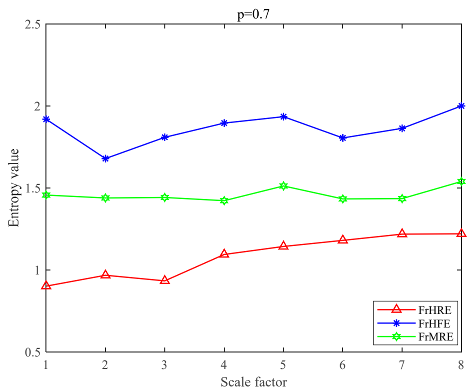

(**e**) Three entropy curves for $1/f$ noise at $p = 0.4$




(**f**) Three entropy curves for $1/f$ noise at $p = 0.5$

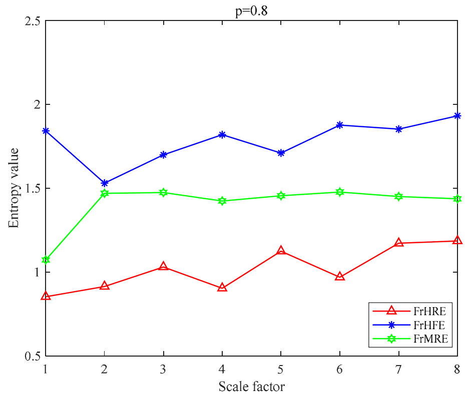



(**g**) Three entropy curves for $1/f$ noise at $p = 0.6$

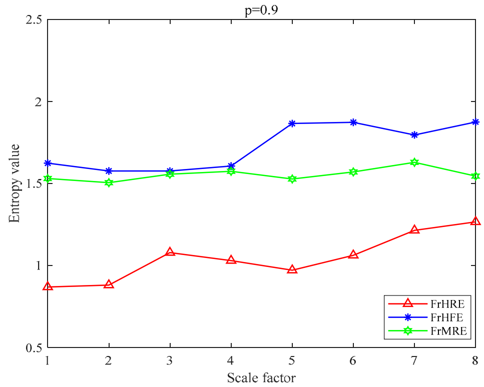



(**h**) Three entropy curves for $1/f$ noise at $p = 0.7$

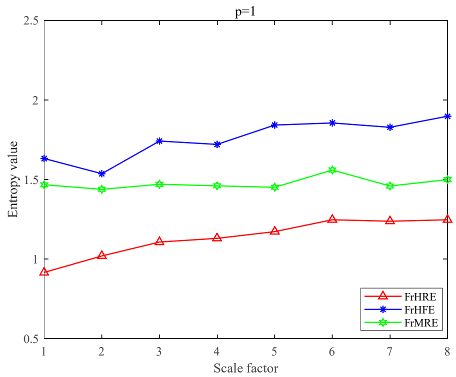
**Figure 7.** *Cont.*




(**i**) Three entropy curves for $1/f$ noise at $p = 0.8$

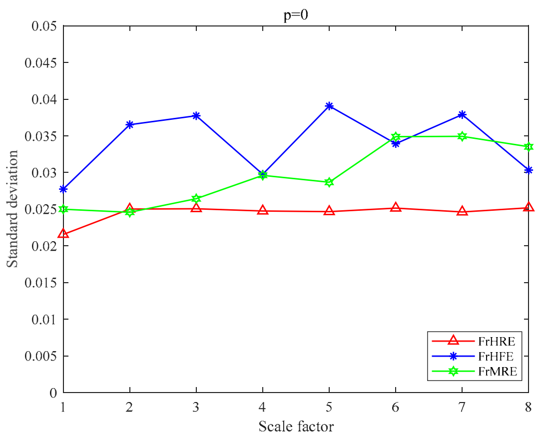



(**j**) Three entropy curves for $1/f$ noise at $p = 0.9$

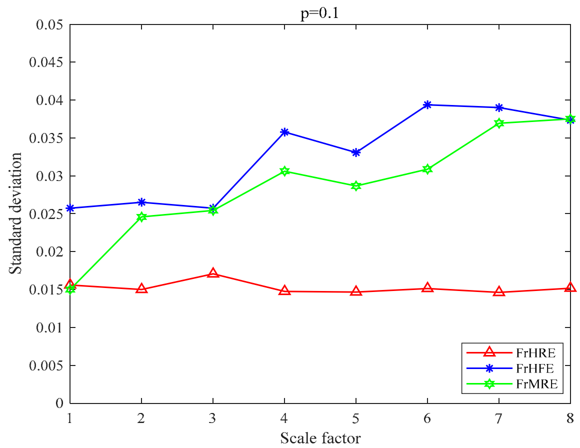



(**k**) Three entropy curves for $1/f$ noise at $p = 1$

**Figure 7.** The entropy curves of FrHRE, FrHFE and FrMRE for $1/f$ noise in different orders.

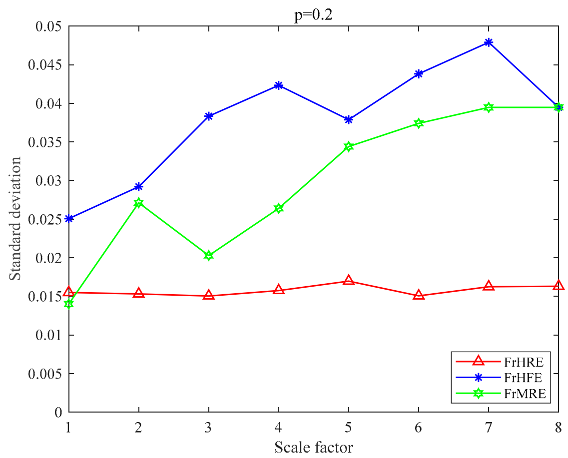



(**a**) The standard deviation curves for white Gaussian noise at $p = 0$

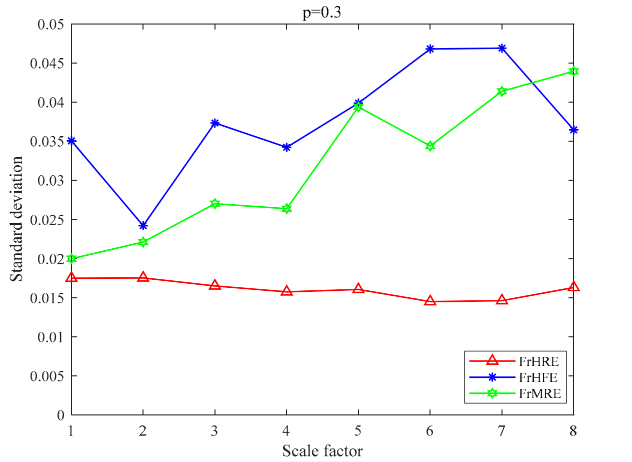



(**b**) The standard deviation curves for white Gaussian noise at $p = 0.1$

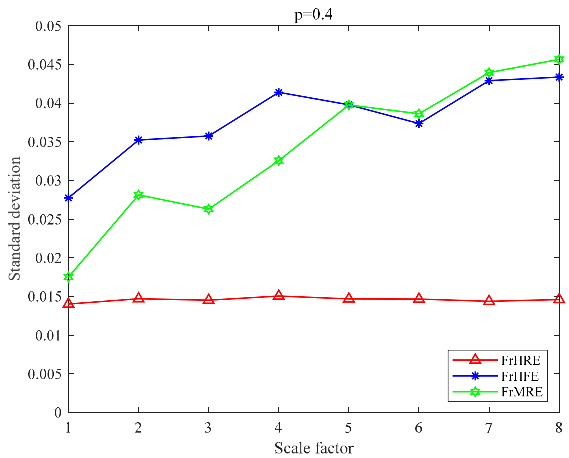
**Figure 8.** *Cont.*



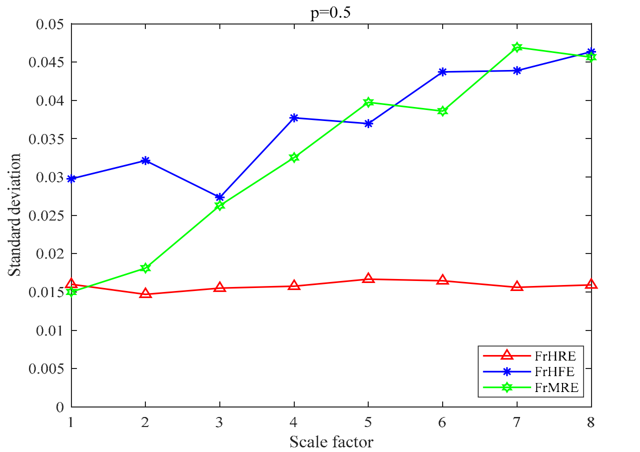

(**c**) The standard deviation curves for white Gaussian noise at $p$ = 0.2



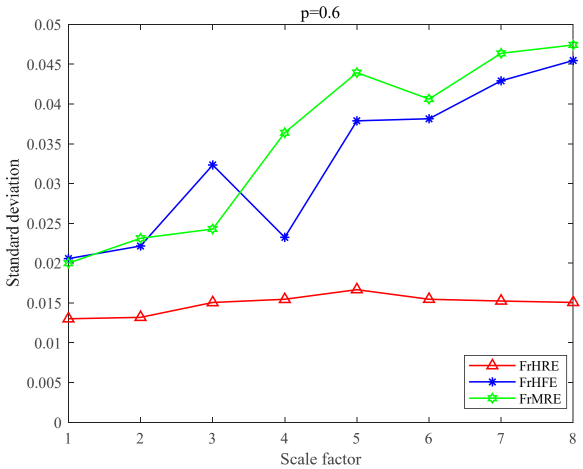

(**d**) The standard deviation curves for white Gaussian noise at $p$ = 0.3



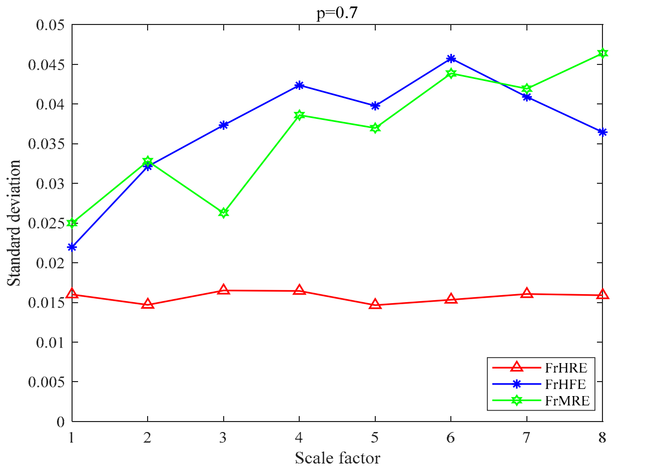

(**e**) The standard deviation curves for white Gaussian noise at $p$ = 0.4




(**f**) The standard deviation curves for white Gaussian noise at $p$ = 0.5

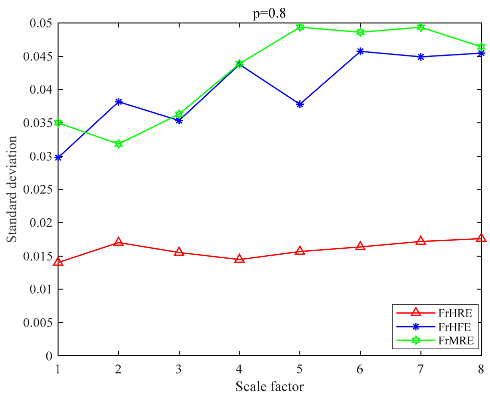



(**g**) The standard deviation curves for white Gaussian noise at $p$ = 0.6

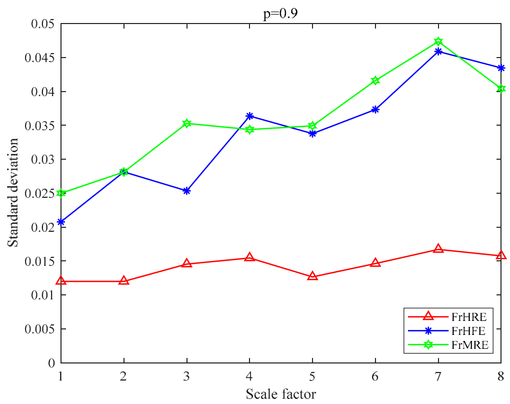



(**h**) The standard deviation curves for white Gaussian noise at $p$ = 0.7

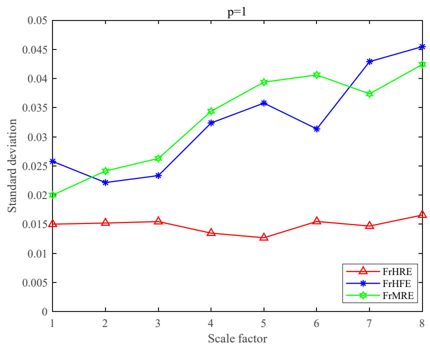
**Figure 8.** *Cont.*




(**i**) The standard deviation curves for white Gaussian noise at $p = 0.8$




(**j**) The standard deviation curves for white Gaussian noise at $p = 0.9$

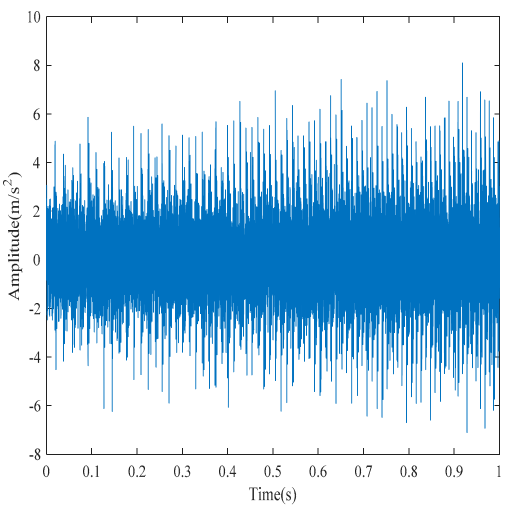



(**k**) The standard deviation curves for white Gaussian noise at $p = 1$

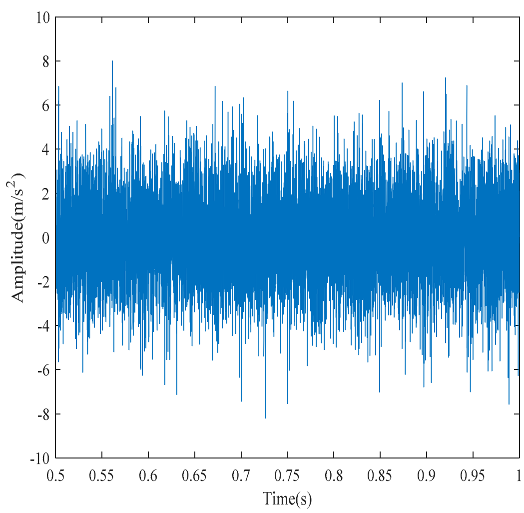
**Figure 8.** The standard deviation curves of FrHRE, FrHFE and FrMRE for white Gaussian noise in different orders.

## *4.3. Classification Performance Analysis of HPO-RF*

Since the signal model in Equation (19) is applicable to the fault of the inner race and outer race and ball, the fault signals of the inner race and ball are, respectively, constructed by setting the parameters of simulated inner race and ball fault signals ($FCC_{inner} = 5.3$, $FCC_{ball} = 2.28$, and the rest of the parameters are the same as the parameters in Section 4.1). Their waveforms are seen in Figure 9.




(**a**) The inner race signal's waveform




(**b**) The ball signal's waveform

**Figure 9.** The waveforms of simulated signals.

Secondly, the four type simulation signals are divided into 100 samples, respectively, according to the division criterion that the length of each sample is 2048. Since the superiority of RLMD and FrHRE in their respective fields have been proved in the previous two sections, RLMD is directly selected in this section to denoise the simulation signals of three fault states, and FrHRE is adopted to extract the entropy features for the three denoised signals in eight scales for different orders, respectively. Therefore, the entropy feature set can be obtained for each fault state signal. Then we randomly select training and test samples at a ratio of 7:3 for the feature sample set of each fault. That is, the training samples' number in each class is 70 and test sample set is 30. The samples are described in Table 1.

**Table 1.** The descriptions of the samples.

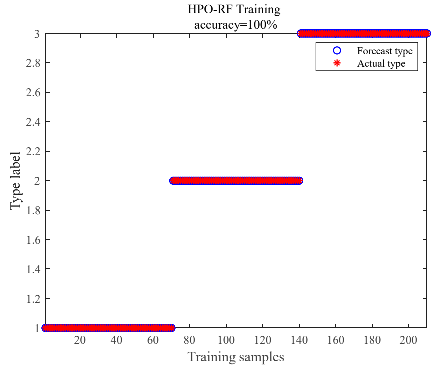
| Fault Type | Training Samples | Test Samples | Label |
|---|---|---|---|
| Inner race | 70 | 30 | 1 |
| Outer race | 70 | 30 | 2 |
| Ball | 70 | 30 | 3 |

Finally, the whale optimization algorithm (WOA), artificial fish swarm optimization algorithm (AFSA), gray wolf optimization algorithm (GWO) and hunter–prey optimization algorithm (HPO) are chosen to optimize the parameters of random forest (RF), respectively, to obtain four improved random forest models. The training and the test sample are then imported into these classification models for learning and testing separately, so as to verify the superiority of the classification performance of HPO-RF. The training results and test results of these models are shown in Figure 10. To make the results more convincing, the paper uses randomly assigned samples for 20 runs. The average value and standard deviation with 20 runs will be used as the evaluation index. The comparative results of the classification performance of the four models are exhibited in Table 2.

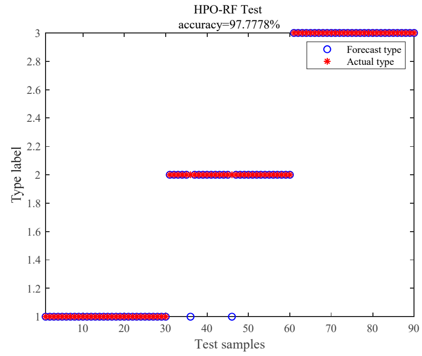
As can be observed, the accuracy of the training model of HPO-RF is 100% and the accuracy of the test model is 97.7778%, both of which ranked first among the four models. Based on the table, some conclusions can be reached: first, the average accuracy of HPO-RF classifier is higher than the other three, which verifies that HPO algorithm has outstanding ability in parameter optimization, making its classification accuracy relatively high; and second, the standard deviation of the HPO-RF model is the least among the four models, which indirectly indicates that the stability of the HPO-RF model is better than the remaining three. Put simply, HPO-RF not only has outstanding recognition ability in fault identification but the stability of the model is also excellent.

**Table 2.** Classification performance of four models.

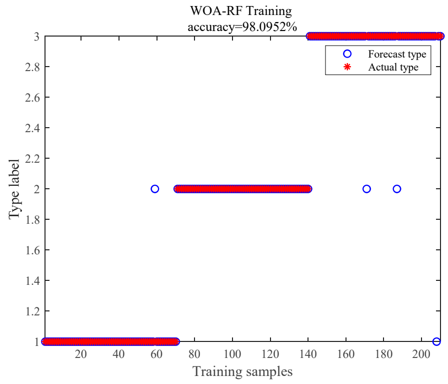
| Models | Training Samples | | Test Samples | |
|---|---|---|---|---|
| | Average Classification Accuracy (%) | Standard Deviation | Average Classification Accuracy (%) | Standard Deviation |
| HPO-RF | 99.67 | 0.0054 | 97.12 | 0.0358 |
| WOA-RF | 96.82 | 0.0408 | 93.08 | 0.0771 |
| GWO-RF | 96.19 | 0.0473 | 90.44 | 0.0964 |
| AFSA-RF | 91.93 | 0.0898 | 82.87 | 0.1831 |



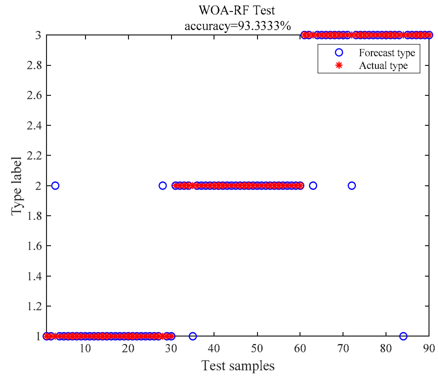

(**a**) The training results of HPO-RF



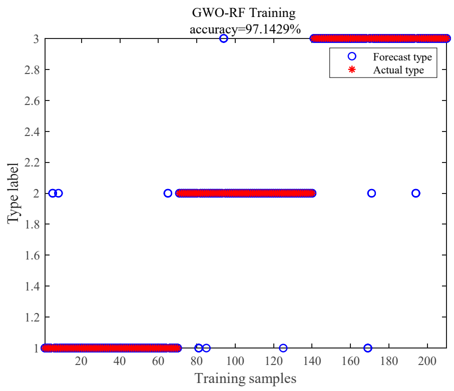

(**b**) The test results of HPO-RF



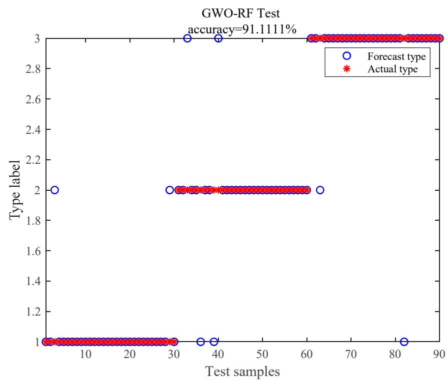

(**c**) The training results of WOA-RF



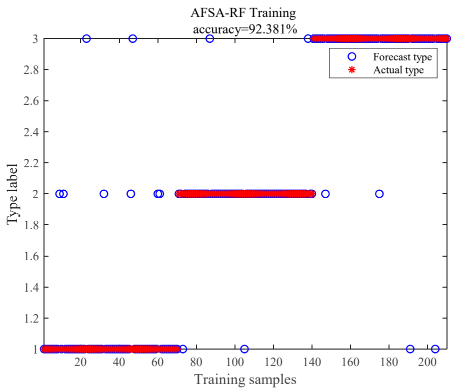

(**d**) The test results of WOA-RF



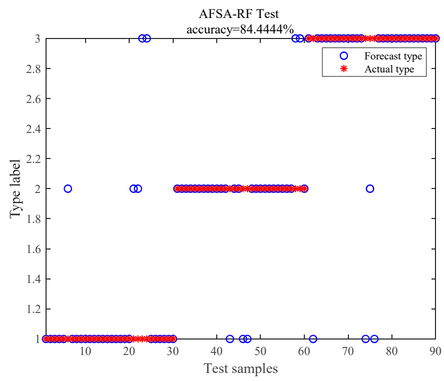

(**e**) The training results of GWO-RF




(**f**) The test results of GWO-RF




(**g**) The training results of AFSA-RF




(**h**) The test results of AFSA-RF

**Figure 10.** Training results and test results of four models.

## 5. Experiments

### *5.1. Experimental Settings*

The proposed method was used on the experimental data provided by Ottawa University to prove the validity of the paper [41]. Experiments were conducted on the SpectraQuest mechanical failure simulator (MFS-PK5M), which is shown in Figure 11.

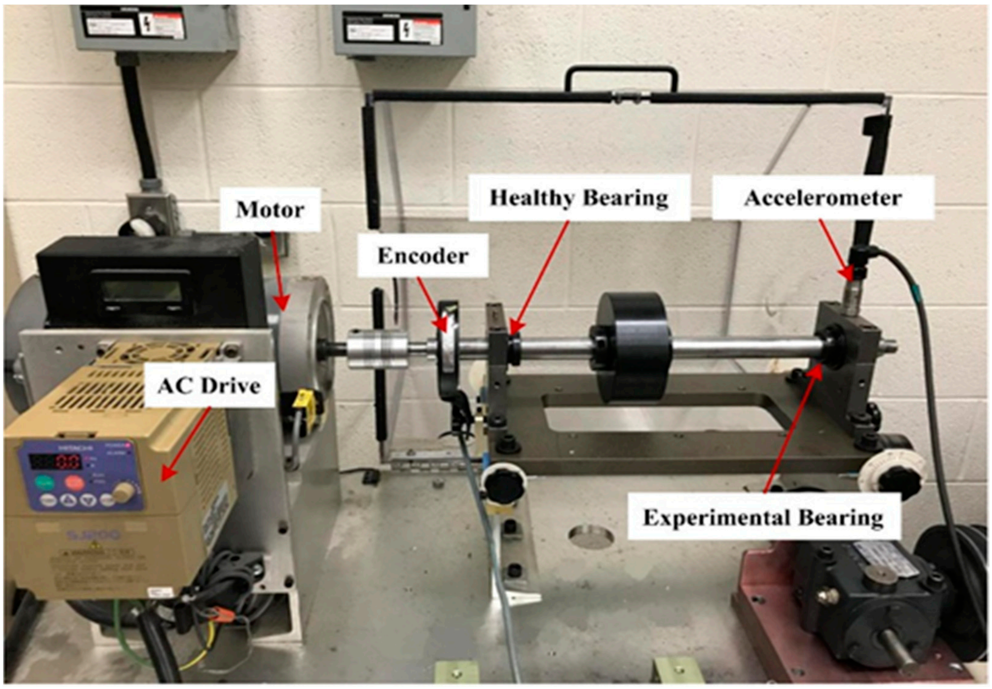



**Figure 11.** The experimental equipment.

It uses a three-phase motor to drive the shaft to rotate and controls the speed through an AC drive. All bearings used in the experiment are ER16K ball bearings. A healthy bearing is installed on the left side of the shaft. The right one is a test bearing that can be substituted by bearings of four health statuses. Moreover, the vibration acceleration signal of the bearing is collected by the ICP accelerometer placed on the experimental bearing shell. In addition, its rotational speed signal is measured by the incremental encoder. The obtained data are collected synchronously on the NI data acquisition boards. The sampling frequency is 200 kHz. The main experimental parameters are exhibited in Table 3.

**Table 3.** The basic parameters in the experiment.

| Name of Parameters | Value of the Parameter |
|---|---|
| Bearing type | ER16K |
| Number of Balls | 9 |
| Pitch diameter | 38.52 mm |
| Ball diameter | 7.94 mm |
| Sampling frequency | 200 kHz |

According to the variable speed condition of the bearing, each health state is divided into four conditions: speed-up, speed-down, speed-up then speed-down and speed-down then speed-up. In this paper, four kinds of signals under the speed-up condition are selected as the research objects. Then the proposed method is applied to those experimental data and analysed in this paper so as to prove its effectiveness.

Under the condition of increasing speed, the experimental signal consists of four health states. Therefore, the experimental analysis can be viewed as a four-class classification problem. In this paper, the first 409600 points of the vibration signal of each class are equally divided into 200 non-overlapping data samples. That is, the sampling length of each sample is 2048, each class has 200 samples and there are 800 samples in total.

### 5.2. Experimental Verification

Figure 12 exhibits four time-domain waveforms of the bearing sampling data under the speed-up condition. It can be viewed that the signal amplitude gradually increases with the speed.

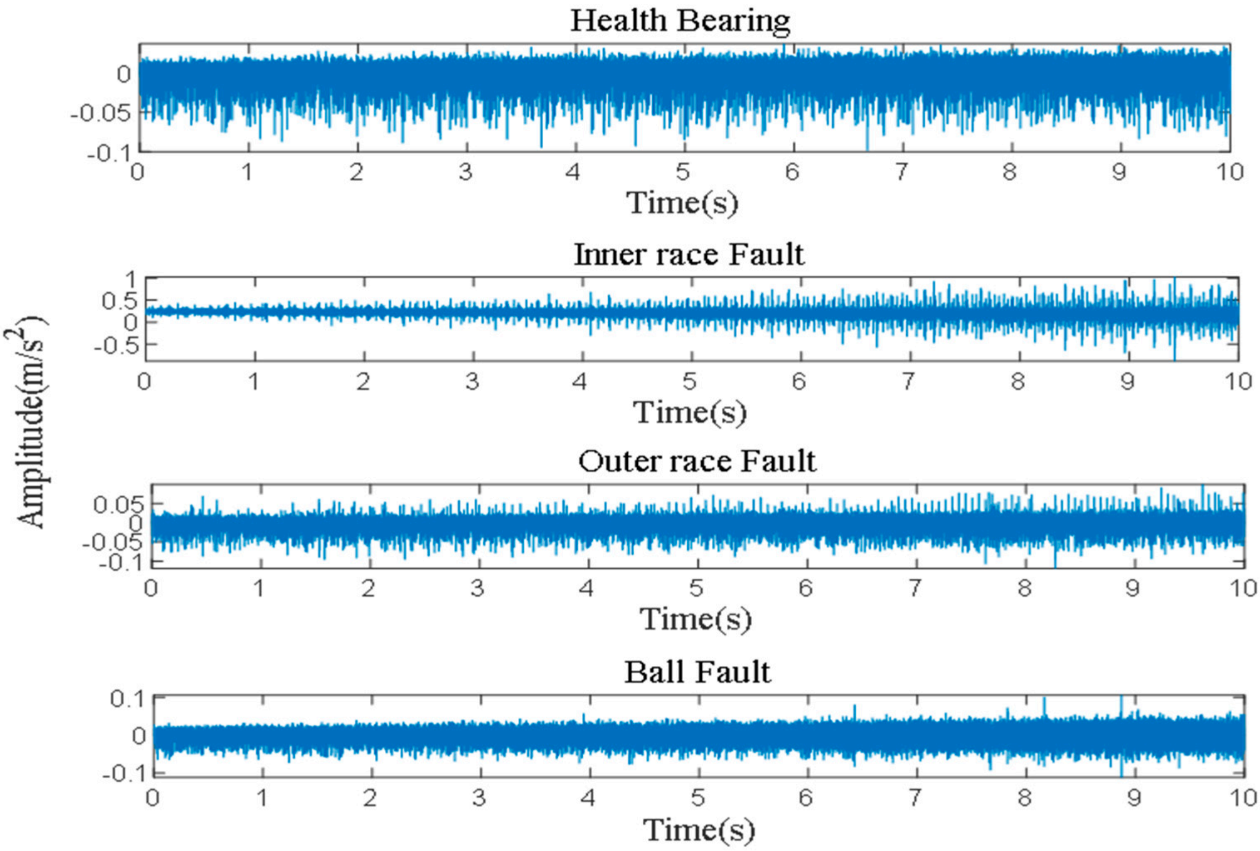



**Figure 12.** The waveforms of four bearing fault signals with speed-up.

Some useful fault feature information is hidden due to the existence of more disturbances. Therefore, RLMD is first selected to decompose the signals of each healthy state, and then multiple PF components and residual terms are obtained in this section. Taking the inner race fault signal as an example, Figure 13 shows the six PF components and a residual term obtained by its decomposition. Moreover, the cross-correlation coefficients between these PF components and original signals of their corresponding health state are calculated, respectively. Signal reconstruction is performed by selecting the component with the largest cross-correlation coefficient. Thus, the effect of removing a certain degree of noise is achieved by reconstructing the signal helps to increase the accuracy of fault classification.

Secondly, the proposed FrHRE was utilized to extract features from denoising signals. Then a $200 \times 88$ time-frequency domain feature set was obtained for each health state. The 11 subgraphs in the following figure are the FrHREs of 11 components with orders from 0 to 1 under 8 scale factors. As can be observed from Figure 14, the trend of the FrHRE curve for the inner and outer race fault signal at each order is almost the same, while the trend of the curve for the healthy state signal and the curve for the ball fault signal are also generally similar. This indicates that the disorder of the inner race and outer race fault signal is similar. (The same can apply to the health state signal and the ball fault signal.) This also implies that misdiagnosis in fault identification can occur when the two signals with similar disorder have very small differences in the value of FrHRE at certain scales. For example, the inner race and outer race fault signal, the health state signal and the ball fault signal in Figure 14f,j, and the health state signal and the ball fault signal in the subgraph(h) have similar entropy values on each scale. However, in most orders, the FrHRE curves of the four health states signals can be clearly distinguished. For example, in Figure 14a–e,g,I,k, the FrHRE values for the four states are more clearly different on most scales. These components can clearly distinguish the FrHRE curves of different states on most scales.

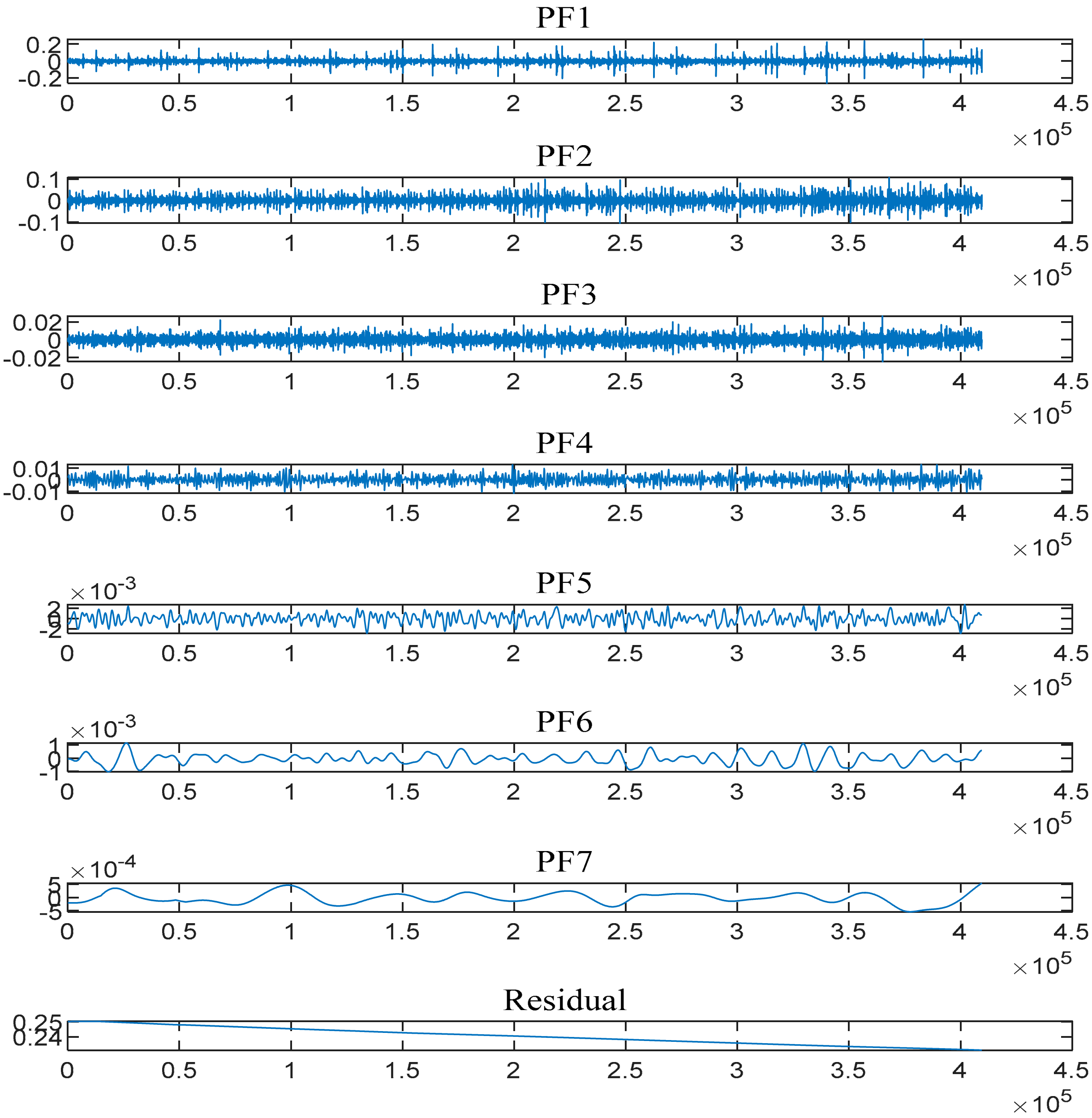
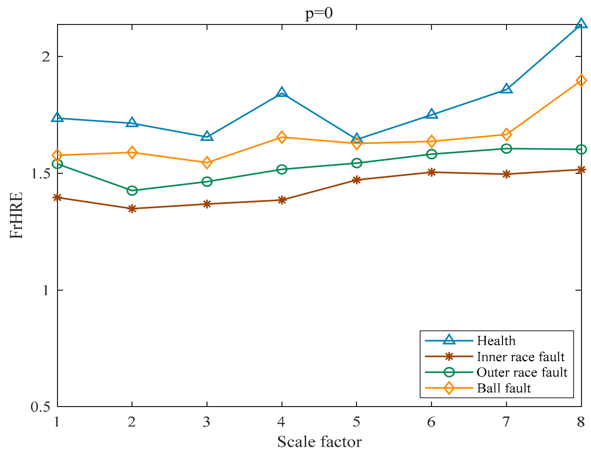
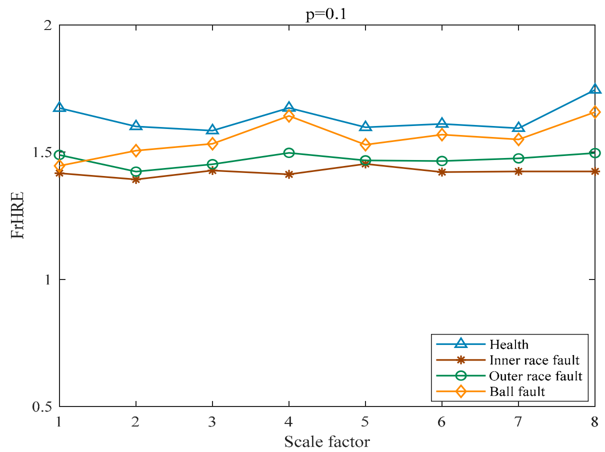



**Figure 13.** The PF components obtained by RLMD.

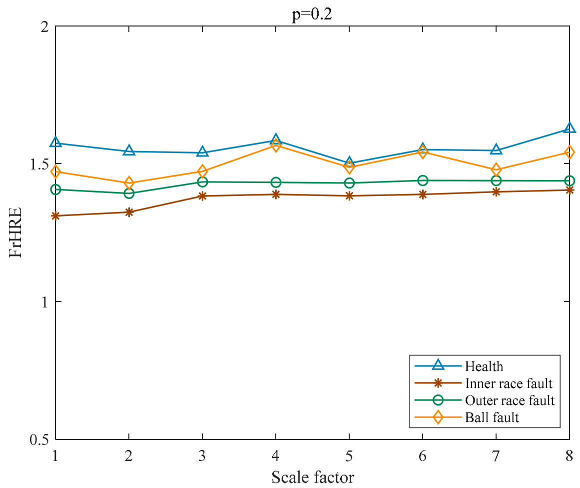
In summary, although the entropy curves of the four signals are overlapped at a small number of orders, the FrHRE curves can still be clearly distinguished in most orders. This shows that FrHRE can play a positive role in feature extraction and make good preparations for subsequent fault identification.



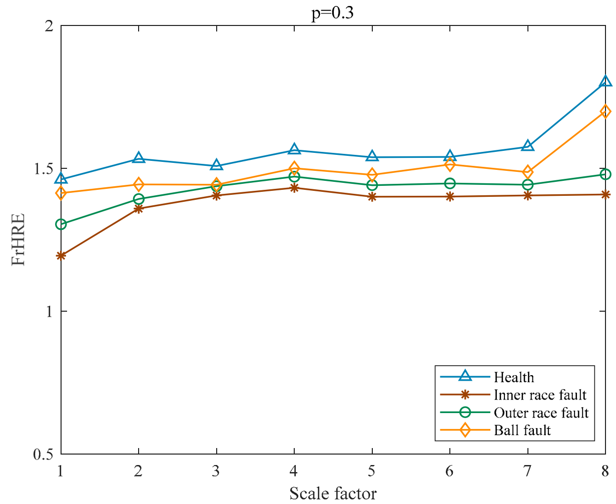

(**a**) The FrHREs of the four fault types on eight scales with $p = 0$.



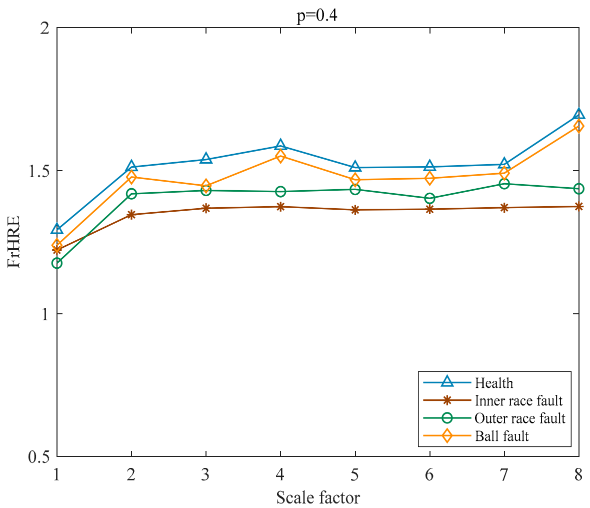

(**b**) The FrHREs of the four fault types on eight scales with $p = 0.1$.



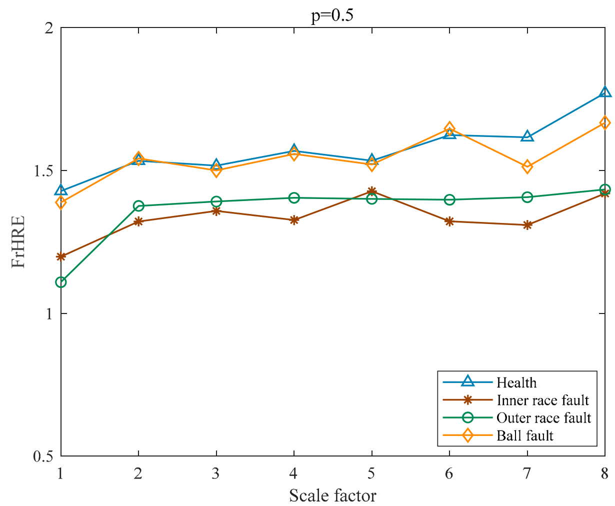

(**c**) The FrHREs of the four fault types on eight scales with $p = 0.2$.




(**d**) The FrHREs of the four fault types on eight scales with $p = 0.3$.

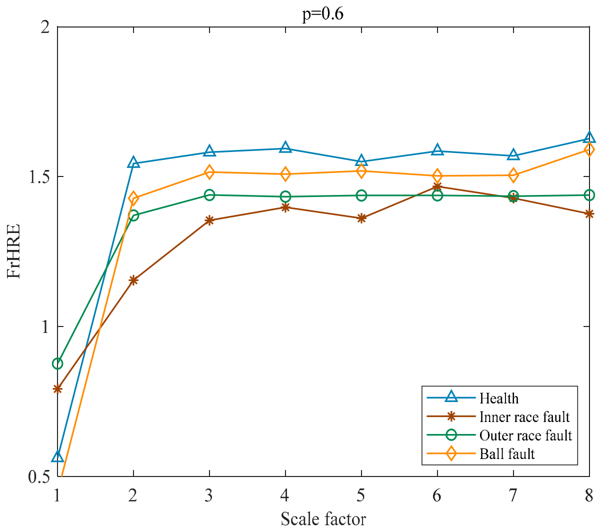



(**e**) The FrHREs of the four fault types on eight scales with $p = 0.4$.

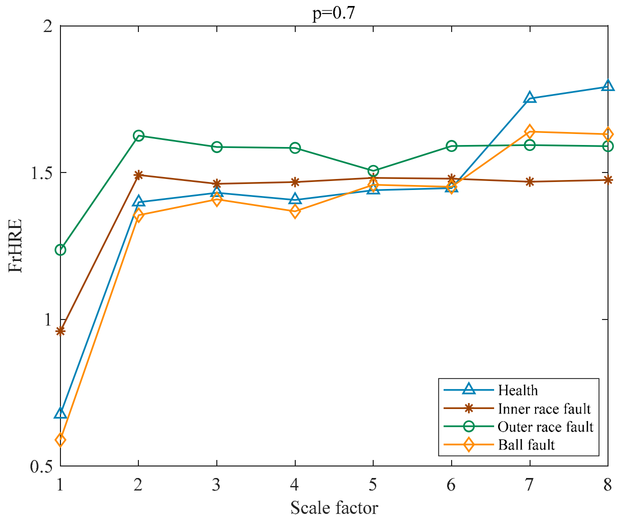



(**f**) The FrHREs of the four fault types on eight scales with $p = 0.5$.

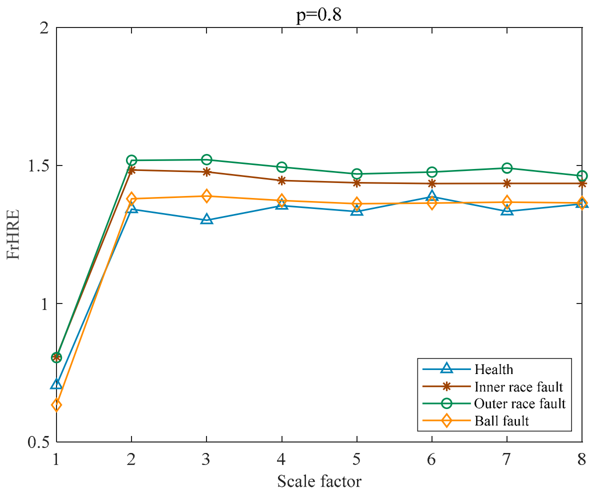
**Figure 14.** *Cont.*



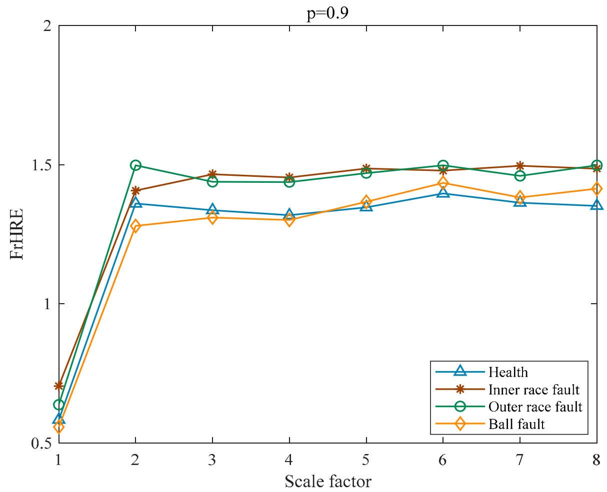

(**g**) The FrHREs of the four fault types on eight scales with $p = 0.6$.



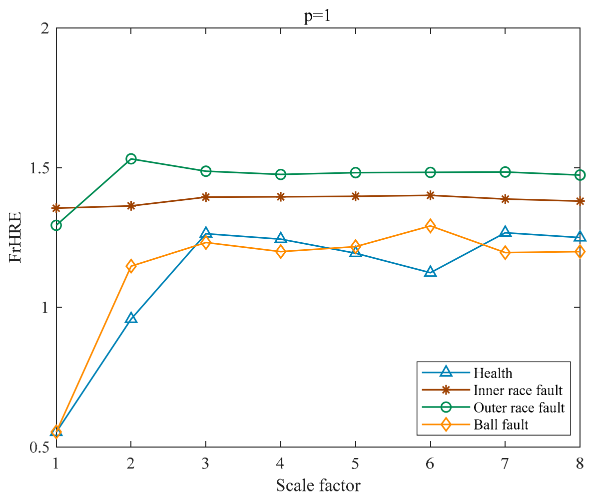

(**h**) The FrHREs of the four fault types on eight scales with $p = 0.7$.




(**i**) The FrHREs of the four fault types on eight scales with $p = 0.8$.




(**j**) The FrHREs of the four fault types on eight scales with $p = 0.9$.




(**k**) The FrHREs of the four fault types on eight scales with $p = 1$.

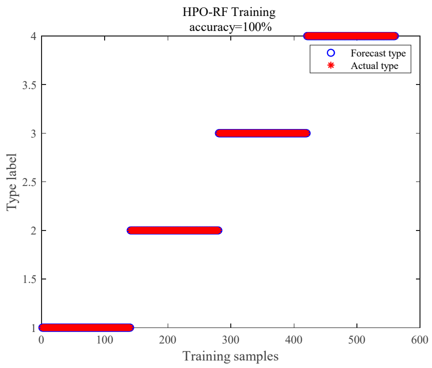
**Figure 14.** The FrHREs of 11 components by using RLMD-FrHRE.

For the sake of testing the modeling accuracy and generalization of HPO-RF, the sample set processed by RLMD-FrHRE is used to test the established HPO-RF model.

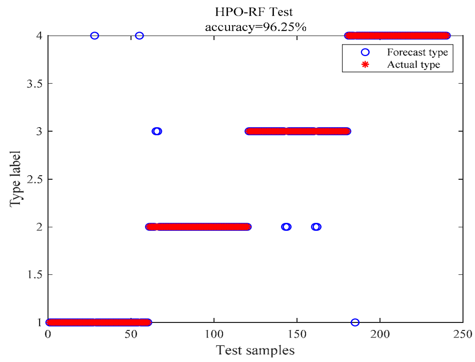
Before testing, the samples are separated into training samples and test samples. By convention, the ratio of the training samples to the test samples should be 7:3 or 8:2, that is, the number of training samples for each class in this paper should be 140 or 160. However, when the number in each class is 140, the training result of the proposed model is 100%, while its test result is 96.25% in Figure 15. When 160 samples are chosen randomly from each class as training samples, the model achieves a training result of 100% and a test result of 97.5%.

Table 4 is presented to provide a clearer visualisation of the results of different training samples numbers of the model's classification accuracy. Different numbers of training samples has no effect on the training results. Moreover, the accuracy of the test results of the model is the highest (97.5%) when the number of training samples randomly extracted from each class is 150 or 160. This indicates that the proposed model has the excellent classification ability with the number of training samples per class being 150 or 160. Although the test accuracy of the model with the training sample size of 150 is the same as that with the training sample size of 160, we believe that the model's classification performance was not improved after adding 10 samples, a factor which was undoubtedly useless.

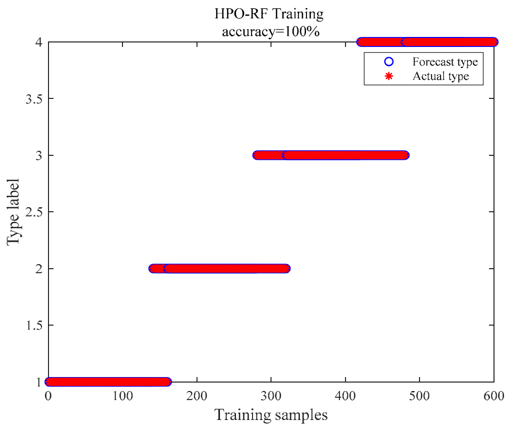
Therefore, this paper chooses 150 samples as training samples of each class, and then the remaining 50 samples are applied for testing. Table 5 gives the detailed introduction of the samples for each health state.



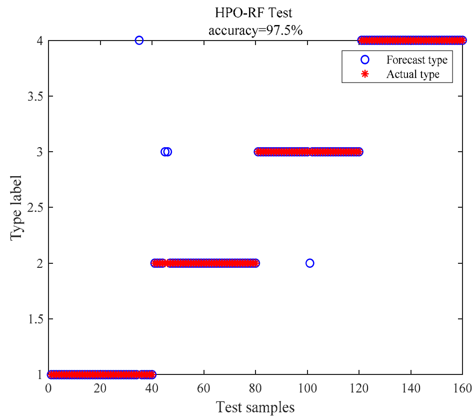

(**a**) The training result when the training samples' number is 140




(**b**) The test result when the training samples' number is 140




(**c**) The training result when the training samples' number is 160




(**d**) The test result when the training samples' number is 160

**Figure 15.** The training and test results when the training samples' number is 140 or 160.

**Table 4.** The different results with different training sample numbers.

| Training Samples Randomly Selected by Each Class | Training Result of HPO-RF | Test Result of HPO-RF |
|---|---|---|
| 120 | 100% | 95.88% |
| 130 | 100% | 96.07% |
| 140 | 100% | 96.25% |
| 150 | 100% | 97.5% |
| 160 | 100% | 97.5% |
| 170 | 100% | 95.83% |

**Table 5.** The detailed description of the samples.

| Fault Type | Training Samples | Test Samples | Label |
|---|---|---|---|
| Healthy bearing | 150 | 50 | 1 |
| Inner race | 150 | 50 | 2 |
| Outer race | 150 | 50 | 3 |
| Ball | 150 | 50 | 4 |

The diagnostic results are seen in Figure 16 and Table 6. As is seen, the classification accuracy of HPO-RF model for training samples has reached 100%, and there is no overfitting. This also verifies that the HPO-RF model has strong learning ability. Subsequently, we input the remaining 200 test samples into the HPO-RF model for classification validation. The results are observed in Table 7 and Figure 17. The test precision of the established HPO-RF model is 97%. It can accurately identify the four signal types. However, among the five misidentified samples, two outer race fault samples were wrongly classified as inner race fault samples. Additionally, two samples of inner race faults were identified as outer race fault samples by mistake. It can be seen from Figure 14 that the FrHRE trends of inner race and outer race, which are of different orders, are similar, that is, the disorder degrees of these two kinds of fault information are similar. This is the reason why the samples that belong to outer race fault or inner race fault account for a large proportion of the incorrectly identified samples.

Therefore, we can see that the HPO-RF model has good modeling accuracy and generalization ability when inputting training samples and test samples into the HPO-RF model, respectively. The paper provides a method for the bearings' fault diagnosis with variable speeds.

**Table 6.** Diagnosis results of training sample.

| HPO-RF | | Real | | | | Precision |
|---|---|---|---|---|---|---|
| | | Healthy Bearing | Inner Race | Outer Race | Ball | |
| Predict | Healthy bearing | 150 | 0 | 0 | 0 | 100% |
| | Inner race | 0 | 150 | 0 | 0 | |
| | Outer race | 0 | 0 | 150 | 0 | |
| | Ball | 0 | 0 | 0 | 150 | |

**Table 7.** Diagnostic results of testing sample.

| HPO-RF | | Real | | | | Precision |
|---|---|---|---|---|---|---|
| | | Healthy Bearing | Inner Race | Outer Race | Ball | |
| Predict | Healthy bearing | 49 | 0 | 0 | 0 | 97.5% |
| | Inner race | 0 | 48 | 2 | 0 | |
| | Outer race | 0 | 2 | 48 | 0 | |
| | Ball | 1 | 0 | 0 | 50 | |

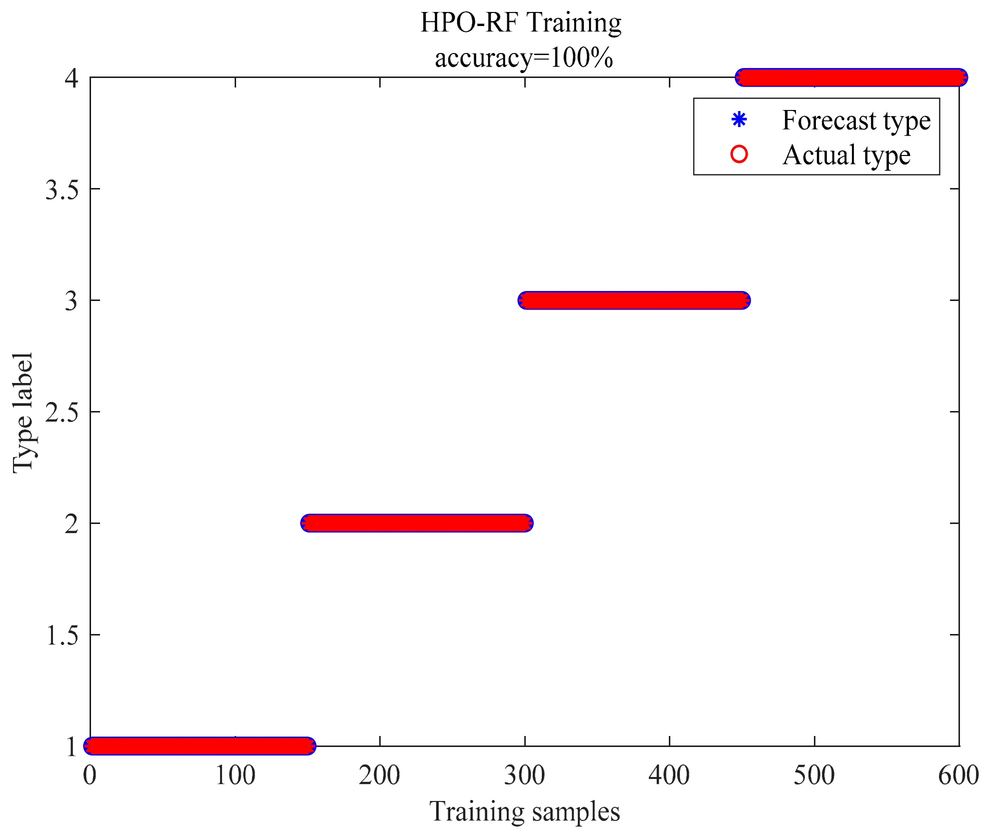



**Figure 16.** The training result.

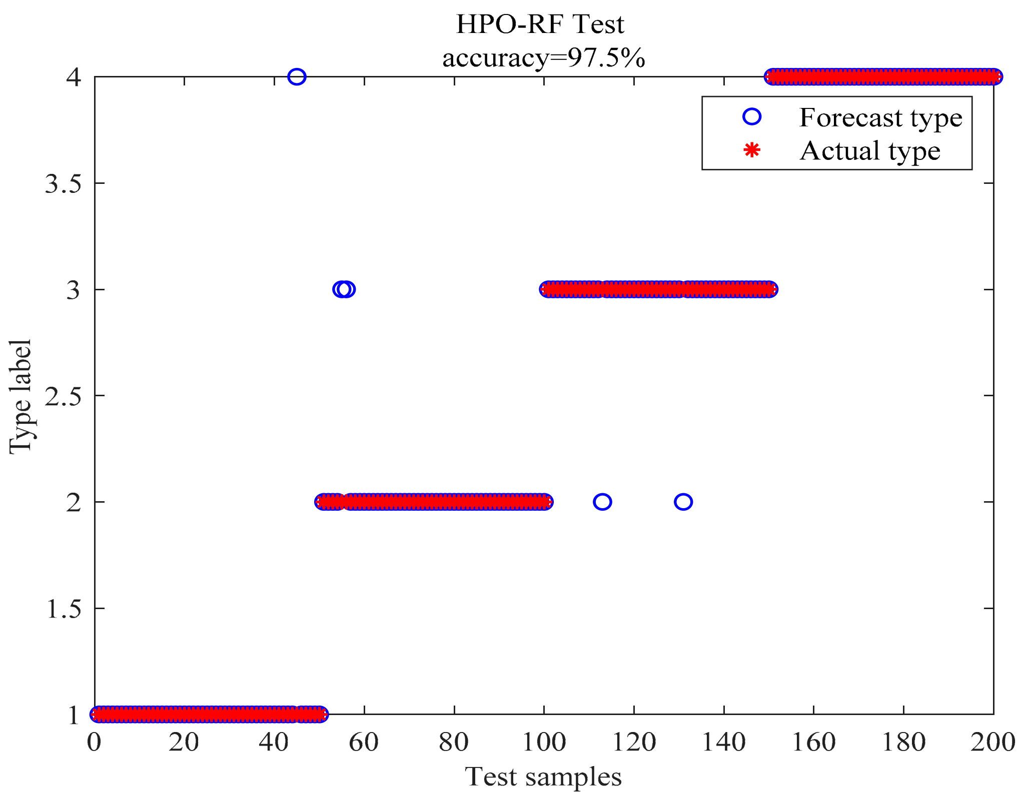



**Figure 17.** The testing result.

## 6. Conclusions

The paper presents a novel model for variable speed bearing signal based on a combination of RLMD, FrHRE and HPO-RF.

In terms of signal preprocessing, the robust local mean decomposition (RLMD) is chosen for noise reduction of variable speed bearing signals. RLMD and other signal decomposition methods are applied to the simulated signal, respectively. By comparing their respective denoising effects, it is proved that RLMD is superior to other methods in eliminating noise, and it also has better performance in noise reduction.

In view of the difficulty of feature extraction for variable speed bearing signals, this paper introduces the fractional Fourier transform to improve it on the basis of hierarchical range entropy, and proposes a new method, namely fractional hierarchical range entropy (FrHRE). The comprehensiveness and stability of FrHRE in extracting entropy features are verified by using simulated noise signal.

Aiming at the problem that the parameters of random forest (RF) cannot be obtained adaptively, this paper improves the RF by using the hunter–prey optimization algorithm (HPO) in order to establish a random forest model with adaptive parameters. Compared with RF model improved by other optimization methods, the stability and accuracy of RF are superior.

The experimental data provided by Ottawa University validates that the proposed method can diagnose the faults effectively under variable speed conditions with good effects.

For the fault diagnosis that rolling bearing is in the working environment of variable speed, we can continue to improve FrHRE to complete higher precision feature extraction in the future. Furthermore, the proposed method should be used on other experimental data to further verify its universality.

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

**Funding:** This research was funded by the National Natural Science Foundation of China under Grant No. 61973041 and the National Key Research and Development Program of China under Grant No. 2019YFB1705403.

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

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are openly available as the University of Ottawa bearing vibration data under time-varying rotational speed conditions.

**Acknowledgments:** We would like to thank the University of Ottawa for providing the bearing dataset under time-varying conditions.

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

## References

1. Li, W.; Chen, J.; Li, J.; Xia, K. Derivative and enhanced discrete analytic wavelet algorithm for rolling bearing fault diagnosis. *Microprocess. Microsyst.* **2021**, *82*, 103872. [CrossRef]
2. Zhu, J.; Hu, T.; Jiang, B.; Yang, X. Intelligent bearing fault diagnosis using PCA-DBN framework. *Neural Comput. Appl.* **2020**, *32*, 10773–10781. [CrossRef]
3. Wang, K.; Dong, S.; Yu, Z.; Shan, S.J. Application of Wavelet Threshold Denoising on Bearing Fault Diagnosis. In Proceedings of the 2019 Chinese Control and Decision Conference (CCDC), Nanchang, China, 3–5 June 2019; pp. 1980–1985.
4. Du, W.H.; Guo, X.M.; Wang, Z.J.; Wang, J.Y. A New Fuzzy Logic Classifier Based on Multiscale Permutation Entropy and Its Application in Bearing Fault Diagnosis. *Entropy* **2019**, *22*, 27. [CrossRef]
5. Wang, J.; Cui, L.; Xu, Y. Quantitative and Localization Fault Diagnosis Method of Rolling Bearing Based on Quantitative Mapping Model. *Entropy* **2018**, *20*, 510. [CrossRef]

6. Gong, T.; Yang, J.H.; Shan, Z. Research on rolling bearing fault diagnosis under strong noise background and variable speed working condition. *Ind. Mine Autom.* **2021**, *47*, 63–71.
7. Dragomiretskiy, K.; Zosso, D. Variational Mode Decomposition. *IEEE Trans. Signal Processing* **2014**, *62*, 531–544. [CrossRef]
8. Lian, J.J.; Liu, Z.; Wang, H.J.; Dong, X.F. Adaptive variational mode decomposition method for signal processing based on mode characteristic. *Mech. Syst. Signal Processing* **2018**, *107*, 53–77. [CrossRef]
9. Nazari, M.; Sakhaei, S.M. Successive Variational Mode Decomposition. *Signal Processing* **2020**, *174*, 107610. [CrossRef]
10. Li, H. Local Mean Decomposition Based Bearing Fault Detection. In Proceedings of the 2nd International Conference on Mechatronics and Intelligent Materials (MIM 2012), Guilin, China, 18–19 May 2012; pp. 360–364.
11. Liao, X.; Wan, Z.; Li, Y.Y.; Cheng, L. Fault Diagnosis Method of Rolling Bearing Based on Ensemble Local Mean Decomposition and Neural Network. In Proceedings of the 2nd International Conference on Mechatronics and Applied Mechanics, Hong Kong, China, 6–7 December 2012; pp. 714–720.
12. Cheng, Y.; Zou, D. Complementary ensemble local means decomposition method and its application to rolling element bearings fault diagnosis. *Proc. Inst. Mech. Eng. Part O-J. Risk Reliab.* **2019**, *233*, 868–880. [CrossRef]
13. Wang, L.; Liu, Z.W.; Miao, Q.; Zhang, X. Complete ensemble local mean decomposition with adaptive noise and its application to fault diagnosis for rolling bearings. *Mech. Syst. Signal Processing* **2018**, *106*, 24–39. [CrossRef]
14. Liu, Z.; Jin, Y.; Zuo, M.J. Time-Frequency Representation Based on Robust Local Mean Decomposition. In *ASME 2016 International Mechanical Engineering Congress and Exposition*; American Society of Mechanical Engineers: New York, NY, USA, 2016.
15. Ke, W.; Jin, Z.P.; Lv, X.C.; Liu, S.H. Fault Diagnosis Method of Variable Speed Rolling Bearing Based on RLMD and WSET. *J. Mech. Electr. Eng.* **2022**, *39*, 300–308.
16. Zhang, L.; Song, C.; Wang, C. Bearing Performance Degradation Assessment based on A Combination of Multi-Scale Entropy and K-medoids Clustering. In Proceedings of the IEEE Prognostics and System Health Management Conference (PHM-Qingdao), Qingdao, China, 25–27 October 2019.
17. Li, Y.B.; Wei, Y.; Feng, K.; Wang, X.Z. Fault Diagnosis of Rolling Bearing Under Speed Fluctuation Condition Based on Vold-Kalman Filter and RCMFE. *IEEE Access* **2018**, *6*, 37349–37360. [CrossRef]
18. Omidvarnia, A.; Mesbah, M.; Pedersen, M.; Jackson, G. Range Entropy: A Bridge between Signal Complexity and Self-Similarity. *Entropy* **2018**, *20*, 962. [CrossRef]
19. Zhang, S.F.; Chen, X.Q. Fault Diagnosis of Bearing Based on Multi-scale Range Entropy and Expert Forest. *J. Mech. Electr. Eng.* **2022**, *39*, 47–52. [CrossRef]
20. Li, F.G.; Wang, J.Y.; Wu, Z.R.; Lin, B.Q.; Lv, P.D.; Fan, R.T. Bearing Fault Diagnosis Method Based on MRE and EigenClass. *Manuf. Technol. Mach. Tool.* **2022**, *6*, 50–54.
21. Zhou, J. *Research On Intelligent Fault Diagnosis Method of Rolling Bearing Based on CYCBD and HRE*; North University Of China: Taiyuan, China, 2020.
22. Zhu, H.; Li, X.Y.; Liu, H.M. Fault Diagnosis of Rolling Bearing Based on WT-VMD and Random Forest. In Proceedings of the 2020 Chinese Control And Decision Conference (CCDC), Shenyang, China, 9–11 June 2020; pp. 2130–2135.
23. Chen, P.; Zhao, X.; Zhu, Q. A novel classification method based on ICGOA-KELM for fault diagnosis of rolling bearing. *Appl. Intell.* **2020**, *50*, 2833–2847. [CrossRef]
24. Zhou, S.H.; Qian, S.L.; Chang, W.B.; Xiao, Y.Y.; Cheng, Y. A Novel Bearing Multi-Fault Diagnosis Approach Based on Weighted Permutation Entropy and an Improved SVM Ensemble Classifier. *Sensors* **2018**, *18*, 1934. [CrossRef]
25. Yang, S.Y.; Yue, J.H. Fault Diagnosis of EMU Rolling Bearing Based on EEMD and SVM. In Proceedings of the International Conference on Computer-Aided Design, Marrakesh, Morocco, 19–21 March 2018.
26. The, C.; Aziz, A.; Elamvazuthi, I.; Man, Z. Classification of Bearing Faults using Extreme Learning Machine Algorithm. In Proceedings of the IEEE 3rd International Symposium in Robotics and Manufacturing Automation (ROMA), Kuala Lumpur, Malaysia, 19–21 September 2017.
27. Zhu, K.H.; Song, X.G.; Xue, D.X. Roller bearing fault diagnosis based on IMF kurtosis and SVM. *Adv. Mater. Res.* **2013**, *694–697*, 1160–1166. [CrossRef]
28. Sui, W.; Zhang, D.; Wang, W. Roller Bearings Fault Diagnosis Based on LS-SVM. In Proceedings of the 2009 IEEE International Conference on Automation and Logistics, Shenyang, China, 5–7 August 2009.
29. Han, T.; Jiang, D. Rolling Bearing Fault Diagnostic Method Based on VMD-AR Model and Random Forest Classifier. *Shock Vib.* **2016**, *2016*, 1–11. [CrossRef]
30. Vakharia, V.; Gupta, V.K.; Kankar, P.K. Efficient fault diagnosis of ball bearing using Relief-F and Random Forest classifier. *J. Braz. Soc. Mech. Sci. Eng.* **2017**, *39*, 2969–2982. [CrossRef]
31. Yazdani-Asrami, M.; Taghipour-Gorjikolaie, M.; Razavi, S.M.; Gholamian, S.A. A novel intelligent protection system for power transformers considering possible eletrical faults, inrush current, CT saturation and over-excitation. *Electr. Power Energy Syst.* **2015**, *64*, 1129–1140. [CrossRef]
32. Tang, G.J.; Pang, B.; Tian, T.; Zhou, C. Fault Diagnosis of Rolling Bearing Based on Improved Fast Spectral Correlation and Optimized Random Forest. *Appl. Sci.* **2018**, *8*, 1859. [CrossRef]
33. Wang, D.; Qu, Y.; Liu, Y.H.; Zhu, X.J. Classification diagnosis of breast cancer based on optimized random forest algorithm. *Comput. Eng. Des.* **2022**, *8*, 1859.
34. Fan, H.D. Study on Stochastic Forest Model Based on Grey Wolf optimization. *Electron. Test* **2022**, *36*, 45–47.

35. Naruei, I.; Keynia, F.; Molahosseini, A.S. Hunter-prey optimization: Algorithm and applications. *Soft Comput.* **2022**, *26*, 1279–1314. [CrossRef]
36. Jiang, Y.; Peng, C.K.; Xu, Y. Hierarchical entropy analysis for biological signals. *J. Comput. Appl. Math.* **2011**, *236*, 728–742. [CrossRef]
37. Tian, J.; Liu, L.L.; Zhang, F.L.; Ai, Y.T.; Fei, C.W. Multi-Domain Entropy-Random Forest Method for the Fusion Diagnosis of Inter-Shaft Bearing Faults with Acoustic Emission Signals. *Entropy* **2020**, *22*, 57. [CrossRef]
38. Zhang, T.; Chen, W.; Li, M. AR based quadratic feature extraction in the VMD domain for the automated seizure detection of EEG using random forest classifier. *Biomed. Signal Processing Control.* **2017**, *31*, 550–559. [CrossRef]
39. Li, X.; Ma, Z.Q.; Kang, D.; Li, X. Fault diagnosis for rolling bearing based on VMD-FRFT. *Measurement* **2020**, *155*, 107554. [CrossRef]
40. Zheng, J.D.; Pan, H.Y.; Cheng, J.S. Rolling bearing fault detection and diagnosis based on composite multiscale fuzzy entropy and ensemble support vector machines. *Mech. Syst. Signal Processing* **2017**, *85*, 746–759. [CrossRef]
41. Huang, H.; Baddour, N. Bearing vibration data collected under time-varying rotational speed conditions. *Data Brief.* **2018**, *21*, 1745–1749. [CrossRef] [PubMed]