machines

MDPI

Article

An Improved Fault Diagnosis Method for Rolling Bearing Based on Relief-F and Optimized Random Forests Algorithm

Yueyi Yang [1], Jiabo Zhai [2], Haiquan Wang [1,*], Xiaobin Xu [3], Yabo Hu [2] and Jinxia Wen [2]

[1] School of Intelligent Perception and Instrumentation, Zhongyuan University of Technology, Zhengzhou 450007, China; yangyy@zut.edu.cn
[2] School of Automation and Electrical Engineering, Zhongyuan University of Technology, Zhengzhou 450007, China; 2022106230@zut.edu.cn (J.Z.); 2024106042@zut.edu.cn (Y.H.); 2023006218@zut.edu.cn (J.W.)
[3] School of Automation, Hangzhou Dianzi University, Hangzhou 310018, China; xuxiaobin1980@hdu.edu.cn
* Correspondence: wanghq@zut.edu.cn

Abstract: Rolling Bearings are important supporting components of rotating machines in industrial processes; the faults of rolling bearings will cause the deterioration of the operation conditions of rotating machines. How to effectively extract the fault features and identify the fault modes of rolling bearings quickly and accurately has become a key issue for the safe operation of rotating machines. In this paper, a novel hybrid fault diagnosis method of an optimized random forests classifier for rolling bearings is proposed. Firstly, the original vibration signals are decomposed by recursive variational mode decomposition (RVMD), and the typical time–frequency domain features are extracted from decomposed signals at different scales. The Relief-F ranking method is utilized to assess the quality of time–frequency domain features, and the top-ranked features with high weight gain are selected for identifying the fault modes. Then, an improved bee colony algorithm is proposed based on the simulated binary crossover criterion, which is used to optimize the key parameters of random forests (RF). Finally, several experiments are conducted on the Case Western Reserve University bearing dataset and the dataset collected from our rolling bearing fault testbed. The experimental results show that the proposed method can accurately identify bearing faults and outperform other state-of-the-art methods.

Keywords: bearing fault diagnosis; recursive variational mode decomposition; parameter optimization; improved bee colony algorithm; random forest

check for updates

Academic Editors: Jie Liu and Wentao Mao

Received: 15 January 2025
Revised: 6 February 2025
Accepted: 14 February 2025
Published: 25 February 2025

Citation: Yang, Y.; Zhai, J.; Wang, H.; Xu, X.; Hu, Y.; Wen, J. An Improved Fault Diagnosis Method for Rolling Bearing Based on Relief-F and Optimized Random Forests Algorithm. *Machines* **2025**, *13*, 183. https://doi.org/10.3390/machines13030183

Copyright: © 2025 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

Rolling bearings are an important supporting component in most rotating equipment, which is widely applied in aerospace, energy power, transportation, and other industrial fields of rotating machinery. The health condition of rolling bearings often directly affects the safe operation of rotating machinery, and bearing failures may cause serious damage and casualties [1]. Therefore, it is of great significance to research fault diagnosis methods to quickly and accurately identify bearing fault modes.

With the rapid development of machine learning algorithms and deep learning algorithms, data-driven fault diagnosis methods have been paid more and more attention. Owing to the inherent interpretability and lower hardware dependencies, many machine learning algorithms have been continuously studied and applied for the fault diagnosis of bearings, which usually include feature extraction and fault mode classification. Since vibration signals are related to the structural dynamics of the machine, to extract richer

features, many signal processing methods are used for the vibration analysis of defective signals, such as Fourier transform, Wavelet transform, and Hilbert–Huang transform [2]. However, these methods have limitations when dealing with nonlinear and non-stationary vibration signals. To mitigate the limitations of conventional vibration signal analysis methods, Huang et al. proposed an empirical mode decomposition (EMD) algorithm [3], which is suitable for processing non-stationary signals, but it has some problems, such as mode aliasing, the under-envelope, over-envelope, and boundary effect. Dragomiretskiy et al. [4] proposed a signal processing method, called variational mode decomposition (VMD), which can overcome the issues of mode mixing and compensate for the shortcomings of the EMD [5,6]. However, the vibration signal of the rolling bearing contains a lot of noise, which may affect the decomposition results. Moreover, the computational complexity of VMD is high, which needs to be solved iteratively and regularized, so the decomposition of large-scale data may face the challenge of computational efficiency and storage space. These deficiencies can be improved by using recursive variational mode decomposition (RVMD) [7], which decomposes the signal into multiple intrinsic mode functions (IMFs). Each IMF is an oscillating function, with adaptability and low redundancy, which can better adapt to different signals. RVMD can be used to find, quantify, and analyze transient and non-stationary dynamics in bearing fault diagnosis with low computational costs [8].

After the vibration signal decomposition of rotating bearings, it is necessary to extract fault features from vibration signals to identify different fault modes. There are several traditional feature extraction methods, such as time domain analysis [9], frequency domain analysis [10], and time–frequency analysis [11]. The time domain signals can be directly analyzed, and time domain features are calculated from signal waveforms, which include the mean value, kurtosis, root mean square, and so on [12]. The vibration signal of rolling bearings can be of a stationary or non-stationary nature; the frequency domain features are very suitable for analyzing stationary signals, which are mainly the position of the main frequency band, the energy of the reaction vibration, and the energy of the reaction spectrum. Han et al. [13] established a semi-coupled model to reduce the computational complexity in the time–frequency domain analysis of offshore wind turbines. Rauber et al. [14] integrated the time–frequency domain indexes as features and considered the relationship between the values of the time–frequency domain features and the fault degree to realize the effective fault diagnosis of rolling bearings. After feature extraction, there may be redundant feature information, and too high a feature dimension will make model training difficult, so the feature selection method is employed by many researchers in various applications. To eliminate the irrelevant features, a feature selection criterion is applied which can measure the significance of each feature with respect to fault modes, such as principal component analysis [15], global–local marginal Fisher analysis [16], the hierarchical clustering model [17], mutual information [18], and Relief-F [19]. Compared with other feature-selection method, the Relief-F algorithm [20] is the only individual evaluation filter algorithm capable of detecting feature dependencies and the selected features are not induction-algorithm-dependent, which has gained appeal by striking an effective balance between the computationally efficient and sensitive to complex patterns of association.

Based on the constructed fault features, various intelligent fault classification methods have been applied for the fault diagnosis of mechanical equipment, such as the artificial neural network (ANN) [21], kernel extreme-learning machine (KELM) [22], and support vector machine (SVM) [23], but it is difficult to deal well with complex classification problems by these single classifiers. Ensemble learning is utilized to combine multiple weakly supervised models to obtain a better and more comprehensive strongly supervised model [24]. The underlying idea of ensemble learning is that if one weak classifier produces a wrong prediction, the other weak classifiers can correct it. Random Forest is the most

popular ensemble learning algorithm in recent years, which can provide a good generalization performance in the field of fault diagnosis [17,25]. Random forest is constructed by decision trees based on the bagging ensemble and further introduces the selection of random attributes in the training process of decision trees. The random forest has high randomness, and the feature selection is different in each sampling and each splitting. It can be effectively applied to large-scale samples with high accuracy, and the number of decision trees and the number of features are key parameters, which affect the classification performance. Xu et al. [26] employed the continuous wavelet transform to convert the vibration signal into a two-dimensional grayscale image and extracted the features from the grayscale image by CNN, and the random forest was applied to identify the rolling bearing fault modes. Tahi et al. [27] proposed a bearing fault diagnosis method based on decision trees, where genetic packaging is used for feature selection and kurtosis and the crest factor are the main features. Tian et al. [28] applied multi-domain entropy to obtain the fault features of the acoustic emission signal and the fault diagnosis model was established based on the random forest algorithm. To improve the classification performance, the heuristic algorithms are used to optimize the super-parameters of the random forest [29]. Compared with other heuristic algorithms, the artificial bee colony algorithm [30] is an optimization algorithm that simulates the behavior of a natural bee colony, which has fewer parameters and good search characteristics for the global optimization of continuous functions.

Consequently, an effective fault diagnosis method for rolling bearings is proposed with the combination of RVMD, Relief-F, and optimized RF. The real-time vibration signal of the rolling bearing is processed by RVMD to obtain the decomposition signal of multiple dimensions, and the time–frequency domain features are extracted from the decomposed signals. Then, the Relief-F is used to construct a feature set by selecting several top-ranked features. In addition, an improved artificial bee colony algorithm is proposed based on the simulated crossover binary algorithm, which can adaptively optimize the key parameters and improve the classification accuracy of RF. The main contributions of this study mainly include three aspects:

(1) The feature extraction and selection method for bearing fault diagnosis based on the integration of RVMD and Relief-F are proposed. The real-time vibration signals are first automatically decomposed by the RVMD and multi-scale time–frequency domain features are extracted from the decomposed signals. The Relief-F method can effectively evaluate the distinguishing ability of each feature and select several sensitive features to represent the fault signals.
(2) An improved bee colony algorithm is proposed based on the simulated binary crossover criterion, and the key parameters of random forests are adaptively optimized which can improve the classification accuracy of RF.
(3) Extensive comparison experiments with other state-of-the-art methods are conducted on the Case Western Reserve University (CWRU) bearing dataset and the rolling bearing fault dataset collected from our real testbed. The experimental results showed that the proposed method outperforms other methods in term of the diagnosis accuracy.

The remaining sections of this paper are arranged as follows. Section 2 briefly introduces the RVMD, Relief-F, artificial bee colony algorithm, and RF. Section 3 describes the proposed diagnostic framework and provides the details of the optimized RF model. In Section 4, the CWRU dataset and the fault bearing dataset collected from our real testbed are analyzed to evaluate the proposed method. Finally, some conclusions are drawn in Section 5.

2. Related Theories

2.1. Recursive Variational Mode Decomposition

VMD has been widely applied in the fault diagnosis of rolling bearings, which can decompose multi-component signals into a series of sub-modes by some adaptive Wiener filter banks. However, the number of the modes K and the bandwidth balance parameter α are determined on the basis of experience; too small K will lead to an insufficient resolution, and too large K usually causes the modes containing fault information to be over-decomposed and seriously affects the efficiency, which cannot meet the requirement of automatic signal decomposition in fault diagnosis. In addition, an unreasonable value of α will affect the proportion of noise in the fault modes and interfere with the identification of transient pulse components.

RVMD is a data-driven mode decomposition method, which is used to decompose nonlinear and non-stationary signals. By introducing variational principles and regularization techniques, it can improve the accuracy and reliability of the decomposition results. The main steps of RVMD can be summarized as follows:

(1) Initialize $F(t)$ and $U_i(t)$. $F(t)$ is the original signal $x(t)$ and the initial low-frequency mode $U_l(t)$ is a zero sequence.
(2) Set the parameters of VMD and remove the $U_i(t)$ from $F(t)$. VMD has a better performance in extracting lower-frequency components when α is set to a larger value [31], so α is set to 2000 in this paper. To decompose the low-frequency mode $U_l(t)$ and high-frequency mode $U_h(t)$ in each iteration, the K is set to 2.
(3) Run the VMD method to obtain the low-frequency mode $U_l(t)$ and high-frequency mode $U_h(t)$.
(4) A low-frequency mode is decomposed from the original signal until the $U_l(t)$ and $U_h(t)$ separated by VMD are close enough in the center frequencies. The stopping criterion is dependent on the nearness of two center frequencies which can be expressed as:

$$f_c = (f_h - f_l)/f_l \tag{1}$$

where f_l and f_h correspond to the center frequency of the low-frequency mode $U_l(t)$ and high-frequency mode $U_h(t)$, respectively. If f_c is lower than a refined threshold and the max iteration $i_{\max}$ is not reached, then the flow return to step (2) is $U_i(t) = U_l(t)$. If the stopping criterion is reached, $U_i(t) = U_l(t) + U_h(t)$. Since the vibration signal of rolling bearings exhibits high levels of second-order cyclostationarity [32], the modes in this paper are selected according to the unbiased autocorrelation (AC) of the envelope of each mode, which can be obtained:

$$R_m(\tau) = \frac{1}{N-q}\sum_{i=1}^{N-q} T(t_i)T(t_i+\tau) \tag{2}$$

where T is the envelope of each mode, $\tau = q/fs$ represents the delay factor, and $q = 0,1,\ldots,N-1$. RVMD is applied to obtain multiple sub-signals in different frequency bands, and then the time–frequency domain features are extracted from each sub-signal.

2.2. Relief-F

The relief-F algorithm is a feature-selection method that assigns different weights to each feature according to the correlation between each feature and the class, and the weight of a feature below a certain threshold will be removed [33]. The relevance of features and categories in the Relief-F algorithm is based on the ability of features to discriminate close samples. In the Relief-F algorithm, the distance between two objects R_1 and R_2 on features f can be calculated.

$$L(f, R_1, R_2) = \frac{R_1(f) - R_2(f)}{\max(f) - \min(f)} \tag{3}$$

where $R_1(f)$ represents the value of R_1 on feature f, $R_2(f)$ represents the value of R_2 on feature f, max(f) represents the maximum value on feature f, and min(f) represents the minimum value on feature f. Then, the weight of each feature is updated,

$$W(f) = W(f) - \sum_{j=1}^{k} \frac{L(f, R, H_j)}{mk} + \sum_{l \notin L_{R_i}} \frac{P(l)}{1 - P(L_{R_i})} \sum_{i=1}^{m} \frac{diff(f, R, M_j)}{mk} \quad (4)$$

where H_j and M_j represent the j neighbor of object R and the heterogeneous object, respectively, $L(f,R,M_j)$ and $L(f,R,M_j)$ denote the distances between objects R and H_j, R and M_j, respectively, m is the number of iterations of the algorithm, k is the number of adjacent objects, L_{Ri} is the label of object R_i, $P(L_{Ri})$ is the probability of object R_i belonging to the label, and $P(l)$ denotes the probability of the label l.

2.3. Random Forest

Random forest is a classifier consisting of a collection of decision tree base classifiers, which employs the bagging method to select samples from the original sample set in a random and repeatable manner to train the classifier. Random forest takes the largest number of votes as the final classification result of the sample, which solves the low accuracy of a single classifier and overfitting. The main procedure implementing RF for classification is briefly summarized as follows.

(1) Some new training sets are generated by bootstrap sampling from the training set, which are used as the training sample of each decision tree, respectively.
(2) The feature sets n_splits are sampled without reusing and m features are extracted. The samples corresponding to the n_splits features are retained as training samples.
(3) Each decision tree is trained using the training samples generated in step (2) and (3).
(4) The pruning threshold is used to prune the trained decision trees.
(5) If the number of trained decision trees is less than the number of decision tree n_trees, it will return to step (2) to continue execution. Otherwise, the random forest is generated, the classification is obtained by aggregating the results of all generated trees. During this aggregation process, the majority voting strategy is adopted.

Due to the bootstrap training set sampling and random selection of features n_splits at each node of the tree, random forest has performed well in many applications related to classification problems. In addition, the number of decision tree n_trees is also a key parameter of random forest; fewer decision trees may lead to the insufficient diversity of base classifiers, and a large number of decision trees may obtain complex models which cause overfitting problems. In this paper, the artificial bee colony was applied to random forest for parameter optimization to achieve a more accurate classification result.

3. The Proposed Fault Diagnosis Method

In this paper, we propose an effective fault diagnosis method for rolling bearings and the flowchart of the proposed method is shown in Figure 1. The proposed method includes three part: signal decomposition by RVMD, construction of the mutual dimensionless features, and establishment of optimized random forest model for fault diagnosis. The implementation process of each part is described as follows.

(1) Part 1: Signal decomposition by RVMD. In this part, each original vibration signal is decomposed by RVMD into different modes, which include different frequency bands ranging from low to high. In this paper, first n modes containing fault information are selected and decomposed signals of each mode are used to extract the fault features.
(2) Part 2: Construction of the mutual dimensionless features. This part includes feature extraction and feature evaluation. Multi-scale fault features in the time domain

and frequency domain are extracted from decomposed signal of different frequency bands by RVMD, which simultaneously contain the fault information in the time domain and frequency domain and can measure the complexity of nonlinear vibration signals under different scales. Moreover, considering that different fault features have varying amounts of sensitivity for fault modes and high-dimensionality feature set will generate redundant and irrelevant information, Relief-F method is employed to evaluate the fault features. Through Relief-F method, the mutual dimensionless fault features are constructed, which can improve the identification accuracy and efficiency.

(3) Part 3: Establishment of optimized random forest model for fault diagnosis. The datasets are divided into training set and testing set. The training set is used to train random forest model, and improved bee colony algorithm is designed to optimize the key parameters of random forest. The testing samples are classified by each decision tree in the RF model, and the fault modes are determined by taking the majority of the outputs from each tree.

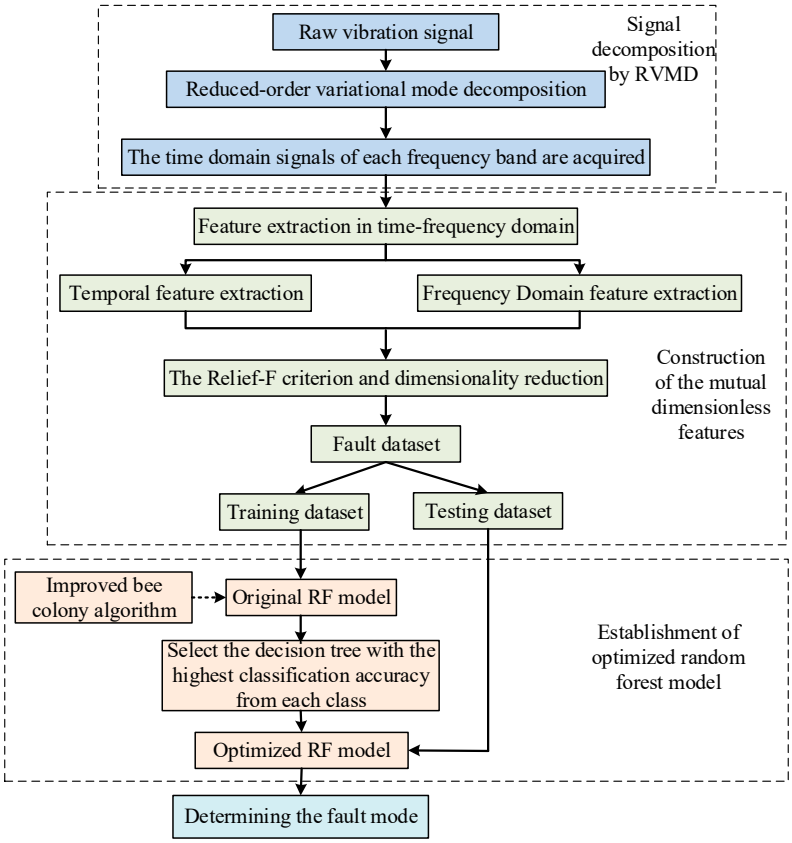


Figure 1. Overall flow chart of the proposed fault diagnosis method.

3.1. Signal Decomposition and Feature Selection

RVMD is utilized to decompose each vibration signal into a discrete number of modes, and these modes contain different frequency bands ranging from low to high. It is supposed that the first n modes including fault information are selected and the decomposed signals for each mode are used for feature extraction.

The features can reflect the information contained in the time domain and frequency domain signal to indicate the occurrence of faults. The time domain vibration signal is beneficial to maintain the basic characteristics of the signal, which are sensitive to failure rather than working conditions. However, the traditional dimensionless indicators have different degrees of aliasing in the feature space, which leads to the uncertainty of diagnosis. As shown in Table 1, 29 features in the time domain and frequency domain are selected

based on existing literature [34] and following the convention in the gearbox fault diagnosis of rotating machinery, where $x(t)$ is a time domain signal (t = 1,2,3,...), T represents a serial number of sampling points, $f(n)$ represents the signal spectrum (n = 1,2,3,...), N is the number of spectral lines, and v_n is the frequency of the nth spectral line. The features in Table 1 have specific physical meanings: parameters C1–C6 reflect the mean, size, and oscillation range of the time domain signal amplitude; parameters C7–C8 reflect the energy of the time domain signal; parameters C9–C16 and C29 reflect the time distribution of the time domain signal; parameter C17 reflects the energy of the frequency domain signal; parameters C18–C21 reflect the position of the main frequency band; and parameters C22–C28 reflect the degree of dispersion of the frequency domain signal. With the precise physical meanings, these features are beneficial to improve the interpretability of fault diagnosis results.

Table 1. Typical time–frequency domain features.

Code	Calculation	Code	Calculation	Code	Calculation
C_1	$\frac{1}{T}\sum_{t=1}^{T} x(t)$	C_2	$\frac{1}{T}\sum_{t=1}^{T} \lvert x(t)\rvert$	C_3	$\min(x(t))$
C_4	$\min(x(t))$	C_5	$C_4 - C_3$	C_6	$\max(\lvert x(t)\rvert)$
C_7	$\left(\frac{\sum_{t=1}^{T}\sqrt{\lvert x(t)\rvert}}{T}\right)^2$	C_8	$\sqrt{\frac{\sum_{t=1}^{T}(x(t))^2}{T}}$	C_9	$\sqrt{\frac{\sum_{t=1}^{T}(x(t)-C_1)^2}{T-1}}$
C_{10}	$\frac{\sum_{t=1}^{T}(x(t)-C_1)^3}{(T-1)C_9^3}$	C_{11}	$\frac{\sum_{t=1}^{T}(x(t)-C_1)^4}{(T-1)C_9^4}$	C_{12}	C_6/C_8
C_{13}	C_6/C_7	C_{14}	$\frac{TC_8}{\sum_{t=1}^{T}\lvert x(t)\rvert}$	C_{15}	$\frac{TC_6}{\sum_{t=1}^{T}\lvert x(t)\rvert}$
C_{16}	$\frac{TC_6}{\sum_{t=1}^{T}(x(t))^2}$	C_{17}	$\sum_{n=1}^{N}\frac{f(n)}{N}$	C_{18}	$\sqrt{\frac{\sum_{n=1}^{N}v_n^2 f(n)}{\sum_{n=1}^{N}f(n)}}$
C_{19}	$\frac{\sum_{n=1}^{N}v_n^2 f(n)}{\sqrt{\sum_{n=1}^{N}v_n^4 f(n)\sum_{n=1}^{N}f(n)}}$	C_{20}	$\sqrt{\frac{\sum_{n=1}^{N}v_n^4 f(n)}{\sum_{n=1}^{N}v_n^2 f(n)}}$	C_{21}	$\frac{\sum_{n=1}^{N}v_n f(n)}{\sum_{n=1}^{N}f(n)}$
C_{22}	$\frac{\sum_{n=1}^{N}(f(n)-C_{17})^2}{N-1}$	C_{23}	$\frac{\sum_{n=1}^{N}(f(n)-C_{17})^3}{N(\sqrt{C_{22}})^3}$	C_{24}	$\frac{\sum_{n=1}^{N}(f(n)-C_{17})^4}{NC_{22}^4}$
C_{25}	$\sqrt{\frac{\sum_{n=1}^{N}(v_n-C_{21})^2 f(n)}{N}}$	C_{26}	C_{25}/C_{21}	C_{27}	$\frac{\sum_{n=1}^{N}(v_n-C_{21})^3 f(n)}{NC_{25}^3}$
C_{28}	$\frac{\sum_{n=1}^{N}(v_n-C_{21})^4 f(n)}{NC_{25}^4}$	C_{29}	$\frac{C_6}{((1/T)\sum_{t=1}^{T}\sqrt{\lvert x(t)\rvert})^2}$		

Although the high-dimensional feature set will increase the useful information for fault diagnosis, it will inevitably produce redundant and irrelevant information, which can reduce the accuracy and efficiency of identification. Therefore, the Relief-F algorithm is used to evaluate all the features of each sub-signal, and the features with high degree of sensitivity are selected.

3.2. Optimized Random Forest

The artificial bee colony is a swarm intelligence algorithm that simulates the honey-gathering process of bees [35]. Different from the other swarm intelligence algorithms, role conversion is a unique mechanism of the ABC algorithm. The bee colony can work together to find high-quality honey sources through the conversion of three different roles:

leading bees, following bees, and scouting bees. Compared with the classical optimization methods, the ABC algorithm has almost no requirements for the objective function and constraints, it only uses the fitness function as the basis for evolution. The artificial bee colony algorithm has the characteristics of simple operation, few control parameters, high search accuracy, and strong robustness.

When the ABC algorithm is used to solve an optimization problem, the location of the honey source is abstracted into a point in the solution space, which represents the potential solution of the problem. The quality of the honey source X_i (i = 1,2,. . .,N) corresponds to the fitness value of the solution f_i, where n is the number of honey sources. The dimension of solving the problem is set as D (d = 1,2,. . .,D), and the position of honey source i in t iterations can be expressed as $X_i^t = [x_{i1}^t, x_{i2}^t, \ldots x_{iD}^t]$. The initial location of honey source x_{id} is randomly generated in the search space and can be expressed as:

$$x_{id} = L_d + rand(0,1)(U_d - L_d) \tag{5}$$

where $x_{id} \in (L_d, U_d)$, and L_d and U_d represent the lower and upper limits of the search space, respectively. At the beginning of the search, a new honey source v_{id} is generated by the hired bees which search around the honey source x_{id}:

$$v_{id} = x_{id} + \varphi(x_{id} - x_{jd}) \tag{6}$$

where φ is a random number in uniform distribution, $j \in$ {1,2,. . .,N}, and $j \neq i$. When the fitness of the new honey source $V_i = [v_{i1}, v_{i2}, \ldots, v_{id}]$ is better than X_i, the greedy algorithm is adopted to replace X_i with V_i. After completing the calculation of the new honey source, all the bees flew back to the information exchange area to share the honey source information. The probability P_i of following bees to choose honey sources depends on the fitness value of the amount of nectar, and the elite bees selected by the following bees are determined based on the roulette wheel strategy:

$$P_i = fit_i \Big/ \sum_{i=1}^{N} fit_i \tag{7}$$

During the search process, if honey source X_i reaches threshold t_{limit} after t_i iterations and no better honey source is found, this honey source will be abandoned, and its corresponding role as an elite bee will be changed to a scout bee. The honey source is updated as:

$$X_i^{t+1} = \begin{cases} L_d + rand(0,1)(U_d - L_d), t_i \geq t_{\text{lim}it} \\ X_i^t, t_i < t_{\text{lim}it} \end{cases} \tag{8}$$

Considering the optimization problem of the classification performance of RF, the judgment criterion of the fitness of the solution is calculated:

$$fit_i = \begin{cases} \frac{1}{(1+f_i)}, f_i \geq 0 \\ 1 + abs(f_i), otherwise \end{cases} \tag{9}$$

where f_i is the diagnosis accuracy of RF model and the lower fitness represents higher diagnosis accuracy.

An improved artificial bee colony algorithm is proposed in this paper based on simulated binary crossover (SBX-ABC). The simulated binary crossover is used to enhance the ability of the information exchange between each generation of honey sources, which can avoid trapping in the local optimum. Therefore, a new honey source is generated based on the simulated binary crossover criterion as follows:

$$v_{id} = \frac{1}{2}\Big[(1 - \beta_i) \cdot x_{id} + (1 + \beta_i) \cdot x_{jd}\Big] \tag{10}$$

where β is a random number which is generated:

$$\beta_i = \begin{cases} [(\mu \times 2)]^{1/(1+\eta)} & \text{if } \mu \leq 0.5 \\ [1/2(1-\mu)]^{1/(1+\eta)} & \text{if } \mu = otherwise \end{cases} \quad (11)$$

where μ is a random number that belongs to (0,1), η is the crossover index, and values of β close to 1 indicate the local perturbation.

SBX-ABC algorithm is used to improve the classification ability of RF in this paper. The purpose of the algorithm is to explore the relationship between more accurate classification results and the key parameters of RF to obtain the difference between the best parameter combination and traditional value selection. The principle of optimized RF can be shown in Figure 2, and the specific steps of SBX-ABC optimization for RF are as follows:

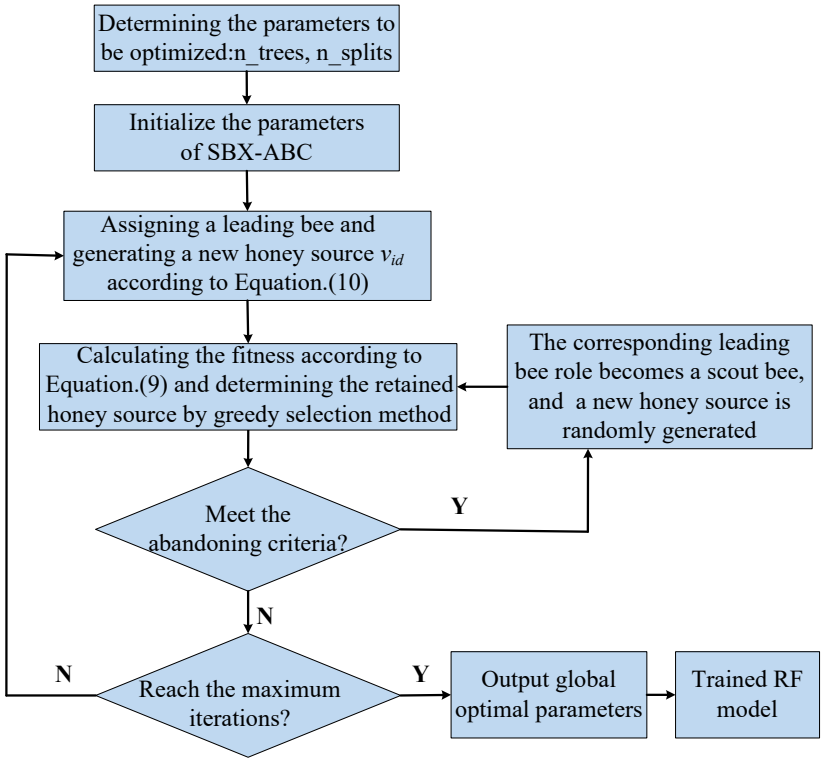


Figure 2. Flow chart of SBX-ABC-RF.

(1) The number of decision tree n_trees and features n_splits is selected as optimized key parameters, their optimization ranges are determined, and the input of SBX-ABC-RF is the training set.

(2) The parameters of the SBX-ABC are initialized, such as the number of honey sources N, initial honey sources X_i, the maximum number of iterations, and threshold value t_{limit}.

(3) The initial location of honey source x_{id} is randomly generated, a leading bee is assigned to the honey source x_{id}, and a new honey source v_{id} is generated by searching around honey source x_{id} according to Equation (10).

(4) The fitness with key parameters of RF is calculated according to Equation (9), and the honey source is determined as whether to be retained according to the greedy selection method. The probability is calculated that the honey source found by the leading bee is followed according to Equation (7), and the scout bees search in the same way as the leading bee.

(5) If the honey source X_i meets the abandoning criteria, the corresponding leading bee role becomes a scout bee, and a new honey source is randomly generated. Otherwise, go to step (6).

(6) Iterations $t = t + 1$, when the maximum of iterations number is reached, the target parameters of RF are obtained. If not, the key parameters of RF will be updated, and training will continue until the iteration numbers reach the maximum.

4. Experimental Verification

The proposed fault diagnosis method is fully verified in this section based on two bearing fault datasets. By comparing with other state-of-the-art methods, the effectiveness and superiority of the proposed method is demonstrated. The proposed model is written in the Matlab R2021b environment and runs on Windows 10 with Intel Core i7-8700.

4.1. Descriptions of Datasets

A well-known adopted experimental platform and our test rig are introduced for data acquisition, the details of different datasets used in two cases are described as follows.

Case 1: The Case Western Reserve University (CWRU) bearing failure dataset is used to evaluate the performance of the proposed method, which is a famous open dataset [36]. The construction of the experimental platforms is shown in Figure 3, which consists of a 2 HP motor, a torque sensor/encoder, a dynamometer, and control electronics. The bearing in the system is a 6205-2RSJEMSKF deep-groove ball bearing. Vibration signals are collected for four types including Normal, Ball fault, inner ring fault, and outer ring fault by the acceleration sensors. In each fault type, three fault diameters including 0.007 in., 0.014 in., and 0.021 in. are selected to simulate different fault severity. The speed of the drive motor can be set to 1772 rpm and the sampling frequency is 48 kHz. In addition, the overlapping sampling method is employed to enhance the original vibration data, sampling length is set to 1500, and the detailed experimental data in Case 1 are listed in Table 2. All samples are randomly divided into the training set and test set according to the ratio of 7:3.

Case 2: The fault bearing dataset from our rolling bearing fault testbed is provided and our developed real test bed is shown in Figure 4. The real test bed mainly consists of a 1.5 KW three-phase asynchronous drive motor, a rotor supported by two bearings, a planetary gearbox with a gear ratio of 10:1, and a magnetic powder brake. Two US CTC accelerometer sensors are placed in the vertical and horizontal directions of the first bearing block, respectively. The collected vibration signals are analyzed and saved in the computer through data acquisition and a processing station. The rolling bearing used in Case 2 are UC203 with a pitch diameter of 33.53 mm and contains eight balls. The fault modes in the Case 2 dataset are simulated with reference to several well-known open available datasets, such as the CWRU dataset, Paderborn university (PU) dataset, XJTU-SY dataset, and SEU datasets [37]. The fault components are used to generate five fault modes which include a bearing cage fault, rolling ball fault (0.5 mm hole), inner ring fault (1 mm in diameter), outer ring fault (1 mm in diameter), and both inner ring; the fault components are shown in Figure 5. The normal condition and these five fault modes are denoted as S1, S2, S3, S4, S5, and S6. The vibration signals of Case 2 were collected with the sampling frequency of 48 kHz which is consistent with the acquisition frequency of the CWRU dataset. The data length of the sample is set to 2000 and 800 samples are obtained in each fault mode by the fixed-length sliding window. The detailed experimental data in Case 2 are listed in Table 3. All samples are randomly divided into the training set and test set according to the ratio of 7:3, and the original time domain waveform of each fault mode is shown in Figure 6.

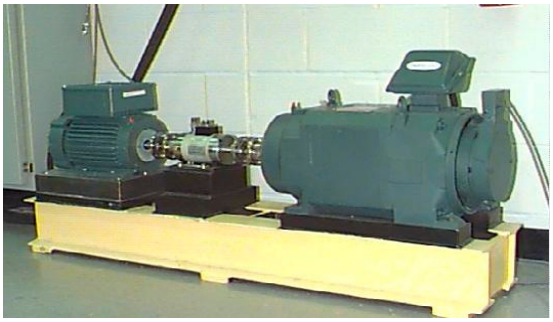


Figure 3. Bearing test rig of Western reserve university.

Table 2. Detailed experimental data description in Case 1.

Label	Fault Mode	Fault Diameter (mm)	Sampling Length	Number of Samples
F0	Normal	-	1500	100
F1	Inner ring	0.007	1500	100
F2	Inner ring	0.014	1500	100
F3	Inner ring	0.021	1500	100
F4	Ball	0.007	1500	100
F5	Ball	0.014	1500	100
F6	Ball	0.021	1500	100
F7	Outer ring	0.007	1500	100
F8	Outer ring	-	1500	100
F9	Outer ring	-	1500	100

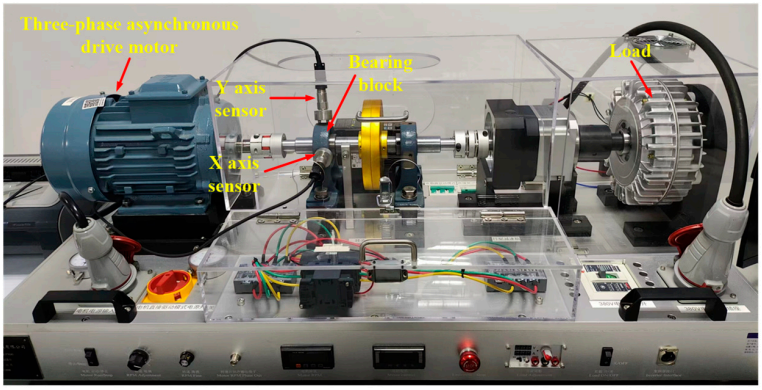


Figure 4. Our developed real testbed.

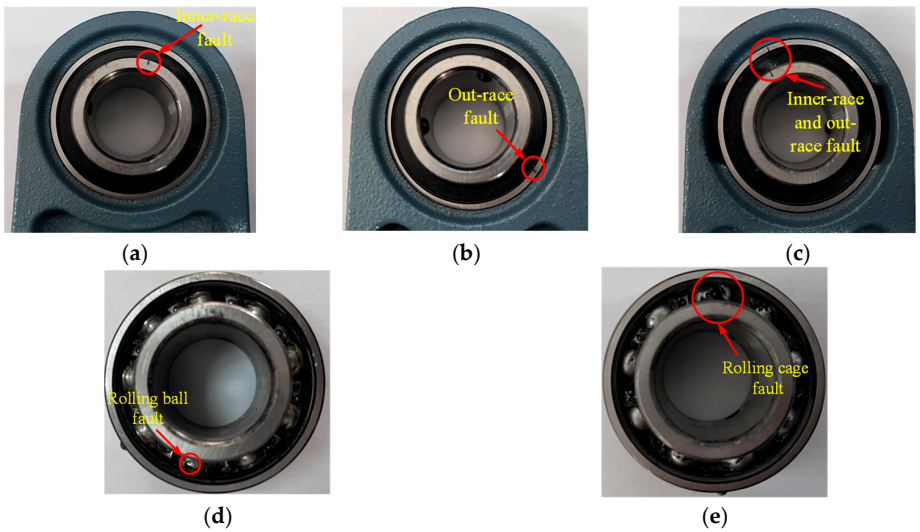


Figure 5. Fault components in our testbed: (**a**) Inner-ring fault, (**b**) Outer-ring fault, (**c**) Inner-ring and outer-ring fault, (**d**) Rolling ball fault, (**e**) Rolling cage fault.

Table 3. Detailed experimental data description in Case 2.

Label	Fault Mode	Fault Diameter (mm)	Sampling Length	Number of Samples
S1	Normal	-	2000	800
S2	Inner ring	1	2000	800
S3	Outer ring	1	2000	800
S4	Inner ring and Outer ring	1	2000	800
S5	Rolling ball	0.5	2000	800
S6	Rolling cage	-	2000	800

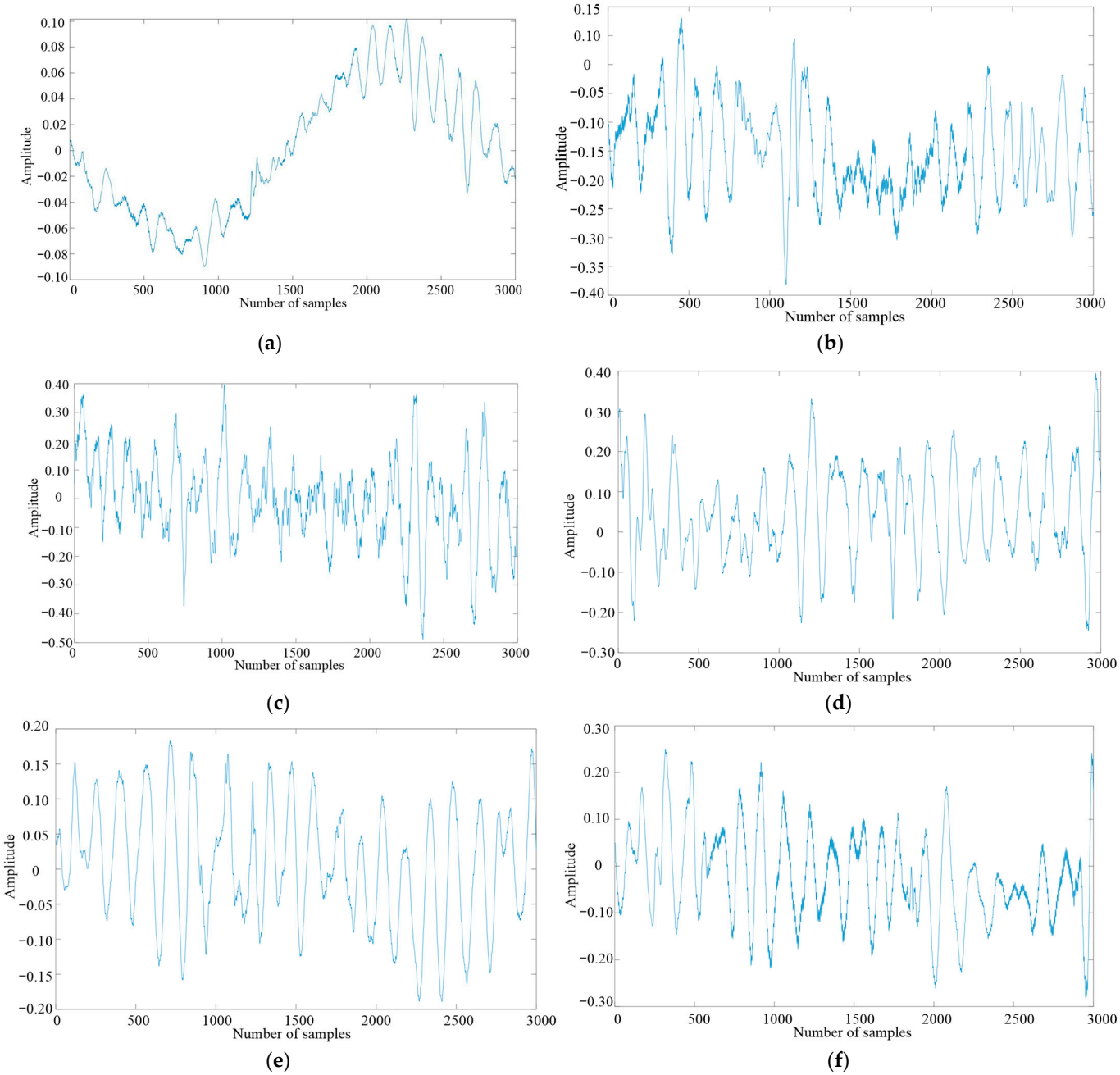
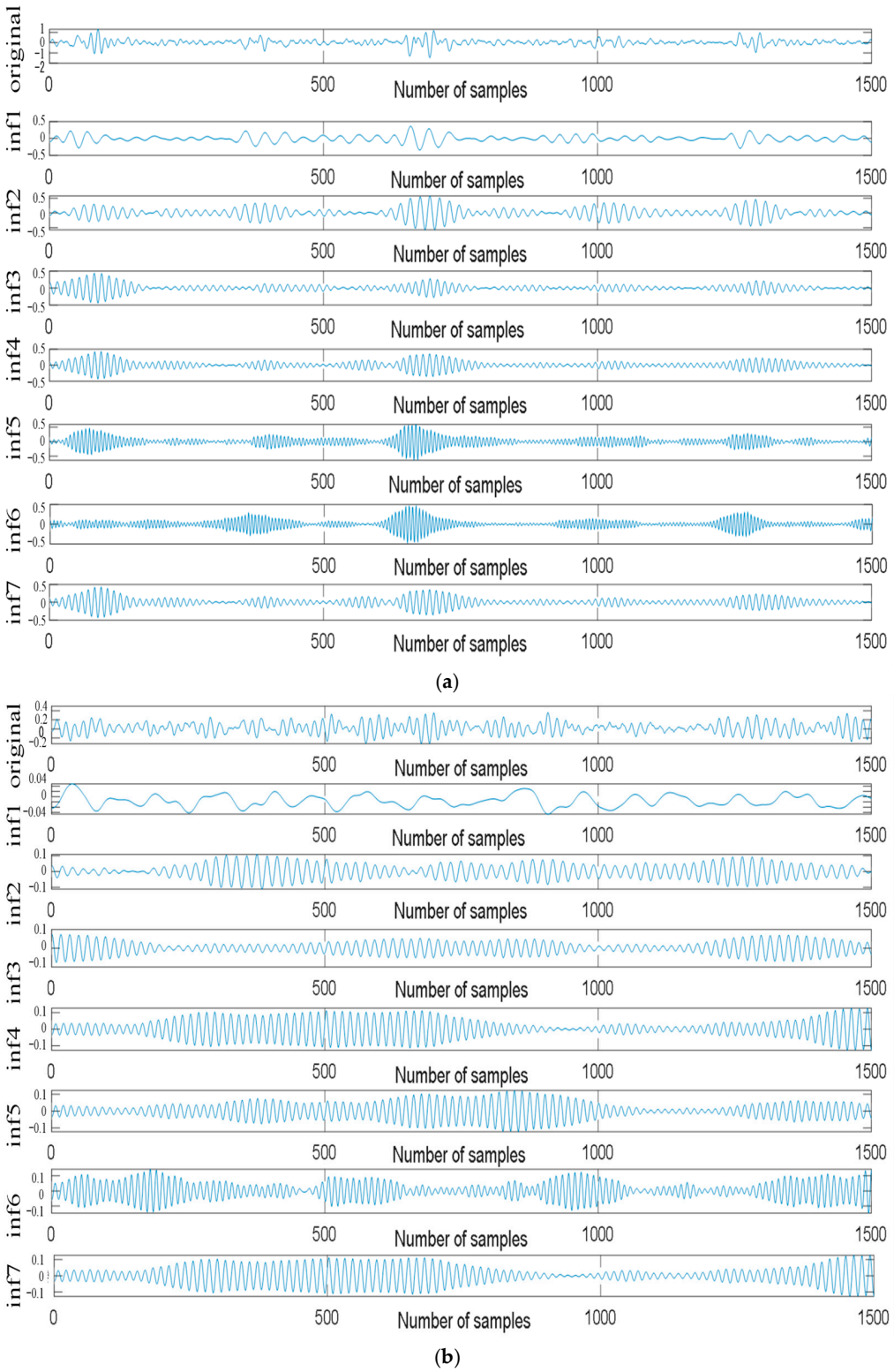


Figure 6. Fault components in our test rig: (**a**) Normal condition, (**b**) Rolling cage fault, (**c**) Rolling ball fault, (**d**) Inner-ring fault, (**e**) Outer-ring fault, (**f**) Both Inner-ring and outer-ring fault.

4.2. Experimental Results of Case 1

The vibration signals of ten fault modes F0–F9 are decomposed by RVMD and the number of decomposed modes is set to seven. These patterns are distributed across different

frequency bands. The RVMD decomposed signals of fault modes F1 and F4 are illustrated in Figure 7. It is shown that the different decomposed signals have various pieces of fault information, so the previously defined time–frequency domain features are extracted from decomposed signals of each fault mode, and we can obtain a feature set containing 203 features from the VMD decomposed signals.



(**a**)

(**b**)

Figure 7. The RVMD decomposition signals in fault modes F1 and F4. (**a**) Decomposition signals in fault mode F1. (**b**) Decomposition signals in fault mode F4.

The Relief-F feature selection method is employed to select superior time–frequency domain features for fault classification. As shown in Figure 8a, the 203 weights of the frequency domain features are obtained by Relief-F, and all features are arranged in order of weight value from large to small. The different number of the features are used as the input of RF and the classification results are shown in Figure 8b. The experimental result shows that Relief-F can not only remove redundant and irrelevant features, but also can improve the generalization ability and classification efficiency of the model. The classification result and computational complexity are considered comprehensively, and the top 40 features with a large weight value in this paper are selected to form a new feature set and are used for fault diagnosis of rolling bearing.

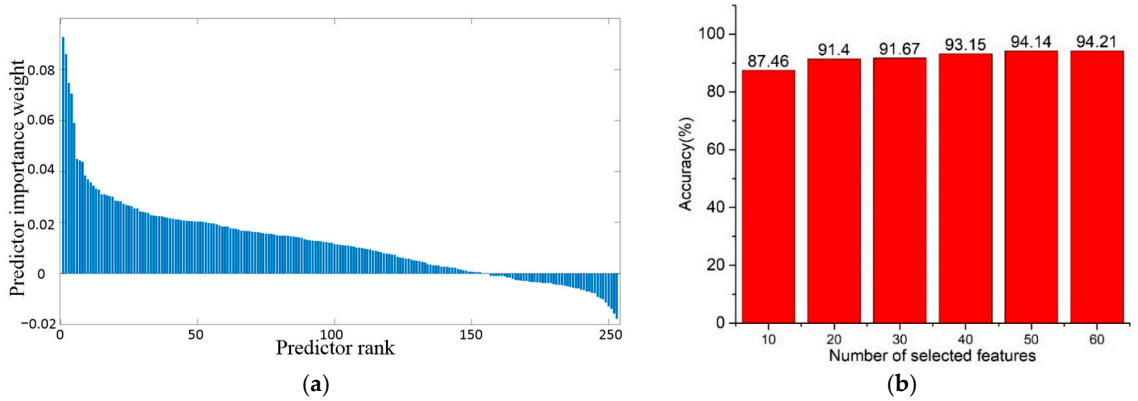


Figure 8. Feature selection results. (**a**) Weight value of all features in case 1, (**b**) Classification results of different number of the features in case 1.

Then, the classified effectiveness of the features is illustrated in Figure 9, and it can be seen that the features have different classification capabilities for different fault modes. Frequency domain feature C20 is extracted from the seventh decomposed signal in the drive end, which can be used to distinguish fault mode F0 from F1 and F6. Time domain feature C29 of the seventh decomposition signal is very useful for the classification of fault modes F0, F1, F4, F8, and F9. The two features are the discriminative features with high weight values for fault diagnosis, which can verify the effectiveness of the Relief-F method.

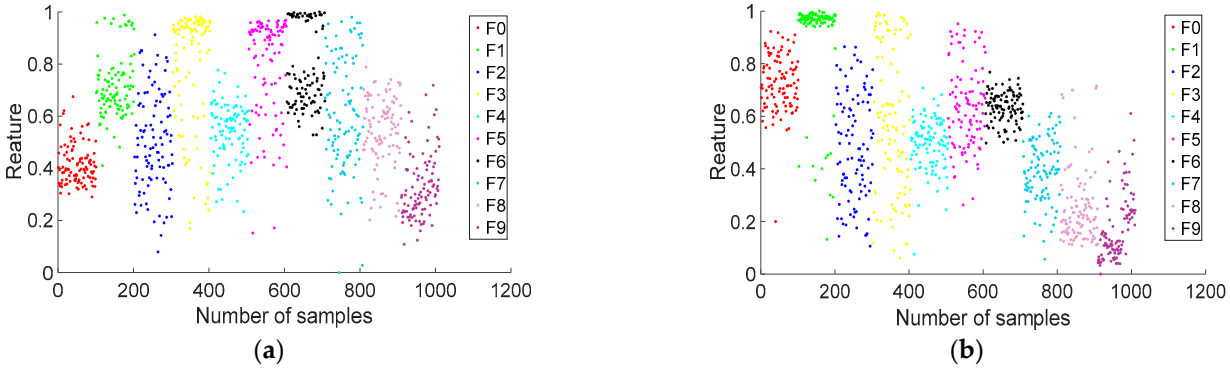


Figure 9. Classification effectiveness of the different features in Case 1. (**a**) Frequency domain feature C_{20}. (**b**) Time domain feature C_{29}.

The optimized performance of the improved artificial bee colony algorithm is shown in Figure 10. It can be seen that the proposed method has a higher classification accuracy than the artificial bee colony algorithm optimized random forest (ABC-RF) model. The iterative speed of the ABC is almost the same as that of the SBX-ABC, but the SBX-ABC is easier to find the global optimal solution and avoid falling into the local optimal solution. The confusion matrix on the test set in case 1 is shown in Figure 11. As shown in Figure 11a, the classification

accuracy of ABC-RF in case 2 is 95.49%, though the ABC-RF method has good classification results, but it has a poor performance in fault mode S3 and S6. The classification results of the proposed method in Case 2 is shown in Figure 11b, the classification performance in fault mode S3 and S6 has been improved, and the overall classification accuracy reaches 97.85%. Therefore, SBX-ABC has a better optimization ability to find the optimal key parameters, and most fault samples are accurately identified by optimized RF.

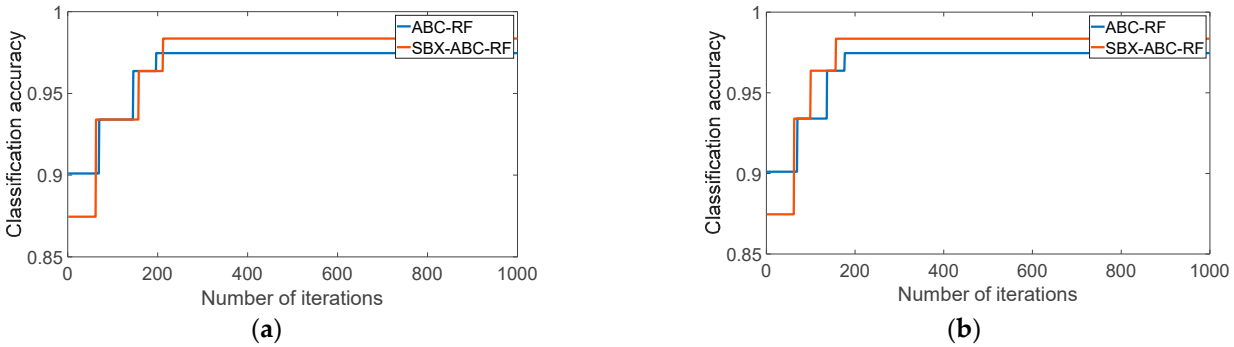


Figure 10. The optimized performance of SBX-ABC in Case 1. (**a**) The population of the swarm is 50. (**b**) The population of the swarm is 100.

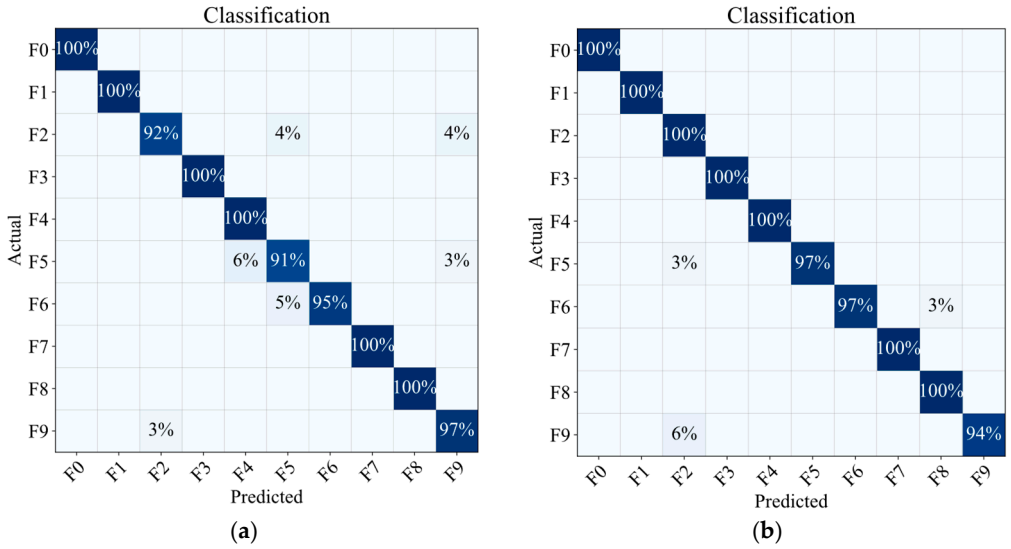


Figure 11. Classification results under different optimized algorithms in Case 1. (**a**) ABC-RF. (**b**) SBX-ABC-RF.

To verify the classification performance of the proposed method with other popular models, five different fault diagnosis methods are selected, namely 1D-BNCNN, LSTM, MLP, SVM, and KNN; for comparison, the detailed model parameter setting is given in reference [38]. The original vibration signals are directly used as the input of 1D-BNCNN and LSTM, and the other methods take the results of the time–frequency domain analysis as the input. To decrease random errors, the final diagnosis accuracy is obtained by repeating the testing procedures 10 times, and the average accuracy is taken as the final diagnosis accuracy. The classification results of these methods on the CWRU dataset are shown in Table 4. It can be seen that the classification accuracy of the SBX-ABC-RF method is 98.68%, which is better than other advanced methods. The experimental results prove the superiority of the proposed method. The proposed method has an improvement of 2.96% and 10.68% compared to 1D-BNCNN and Bi-LSTM, respectively, which shows that our method can better extract fault features. After the manual feature extraction, the accuracy of SBX-ABC-RF is significantly higher than MLP, SVM, and KNN, and it can be observed

that the optimized RF has a better classification ability. In addition, the proposed method performs classification which increases by 1.32% compared to ABC-RF, which shows that the SBX-ABC has a better optimization performance.

Table 4. Comparison results of different methods in Case 1.

Methods	Accuracy (%)
1D-BNCNN	95.72
Bi-LSTM	88.00
MLP	69.76
SVM	46.42
KNN	89.95
ABC-RF	97.36
SBX-ABC-RF	98.68

4.3. Experimental Results of Case 2

The vibration signals of each condition are decomposed into seven subset signals by RVMD, and the decomposed signals of the rolling cage fault and rolling ball fault are shown in Figure 12. To obtain the multi-scale and rich information features' useful fault information, the previously defined time–frequency domain features are extracted from each decomposed signals, and we can construct a feature set containing 203 features.

To select the highly discriminative features, the Relief-F is used to evaluate the diagnostic ability of each feature. As shown in Figure 13a, the weight of the 203 features has been sorted from largest to smallest, and the features with the largest weight are selected to identify the fault modes. The different numbers of features are selected as the input of RF and the classification results are shown in Figure 13b. The classification results and computational complexity are considered comprehensively, and the top 40 features are selected to form a new feature set and are applied for model training. The classification effectiveness of the extracted features for different fault modes is illustrated in Figure 14. The frequency domain feature C26 is extracted from the seventh decomposed signal, the time domain feature C1 is obtained from the sixth decomposed signal, and the two features can be used to distinguish fault mode S0 from the other fault modes, which have high weight values assigned by Relief-F.

The optimized performance of the SBX-ABC algorithm is illustrated in Figure 15. It can be seen that SBX-ABC has a better performance in finding the optimal solutions, and the classification accuracy with the 50 population size is lower than that when the population size is 100. However, when the population size is 100, the optimized time is too long, but the classification accuracy does not increase significantly, so the population size is set to 70 in case 2. The confusion matrixes on the test set in case 2 are shown in Figure 16. As shown in Figure 16a, the classification accuracy of ABC-RF in case 2 is 95.49%, although the ABC-RF method has good classification results in fault mode S1 and S3, and it has a poor performance in fault mode S2, S4, and S6 in which the accuracy is below 93%. The classification results of the proposed method in case 2 are shown in Figure 16b. The classification performance in fault mode S2, S4, and S6 has been improved and the overall classification accuracy reaches 97.85%. Therefore, the proposed method can effectively extract the fault features and accurately identify fault modes.

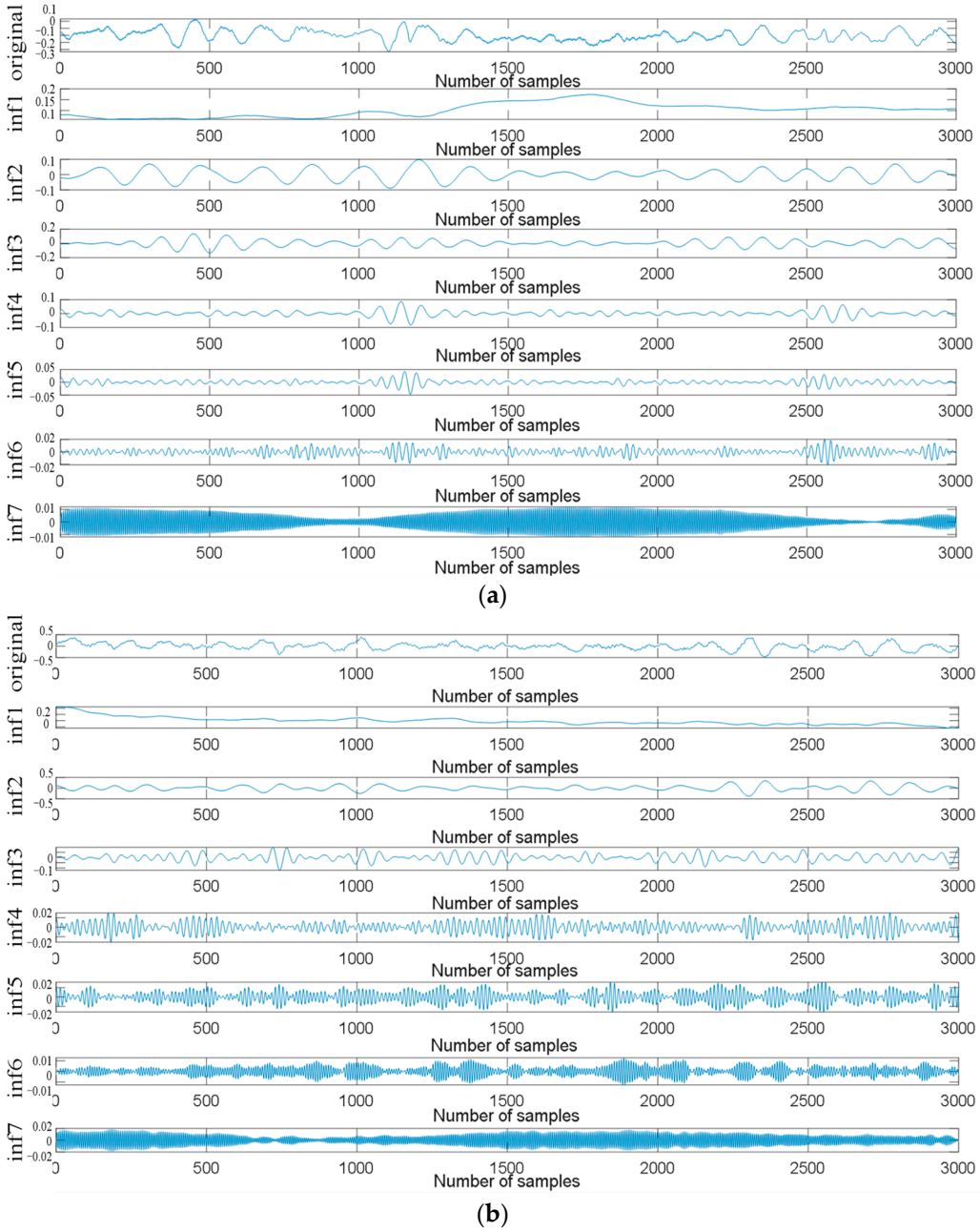


Figure 12. RVMD decomposition signals in Case 2. (**a**) Rolling cage fault. (**b**) Rolling ball fault.

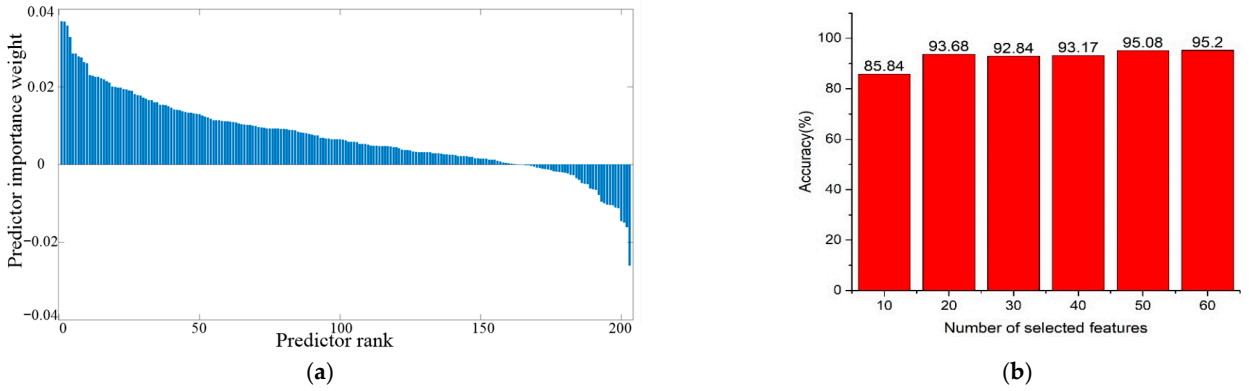


Figure 13. Feature selection results. (**a**) The weight value of 203 features in case 2, (**b**) Classification results of different numbers of the features in case 2.

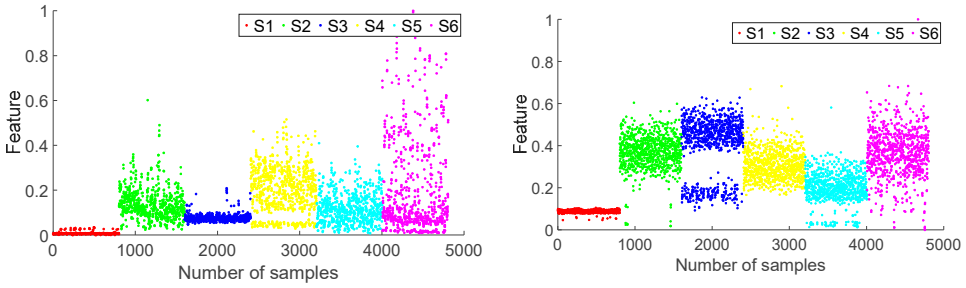


Figure 14. Classification effectiveness of the different features in case 2. (**a**) Frequency domain feature C_{26}. (**b**) Time domain feature C_1.

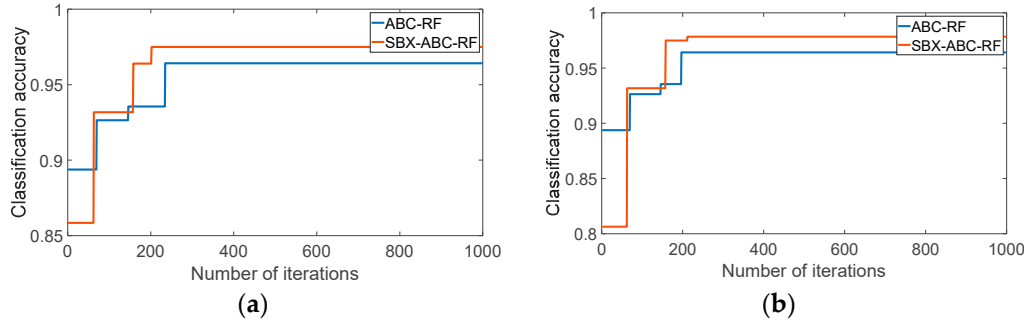


Figure 15. The optimized performance of SBX-ABC. (**a**) The population of the swarm is 50. (**b**) The population of the swarm is 100.

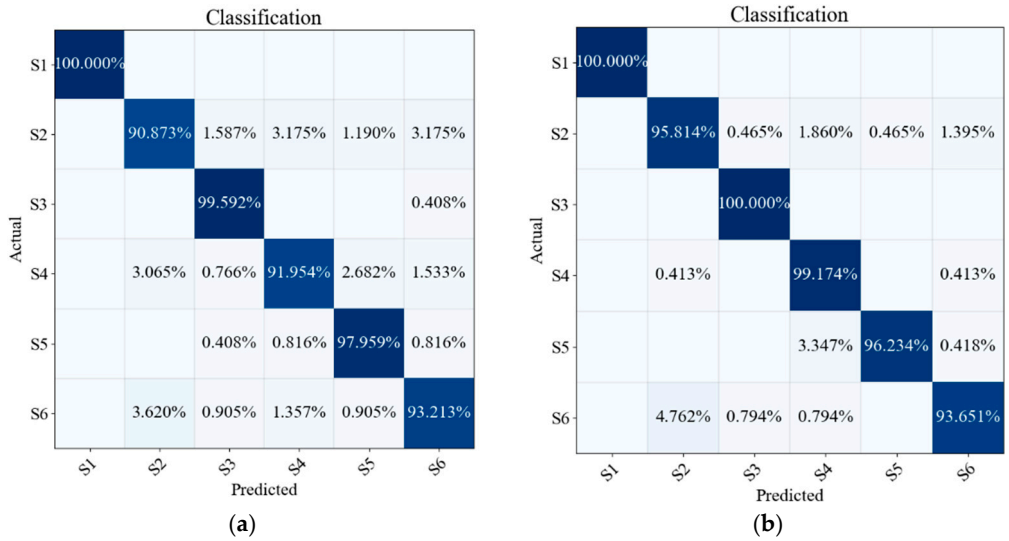


Figure 16. Classification results under different optimized algorithms in Case 2. (**a**) ABC-RF. (**b**) SBX-ABC-RF.

Furthermore, to verify the superiority of the proposed method, the previous methods are used to conduct the comparison in Case 2. The final diagnosis accuracy is obtained by repeating the testing procedures 10 times, and the average accuracy of these methods in case 2 are shown in Table 5. It can be seen that the classification accuracy of SBX-ABC-RF is 97.85%, which is better than other advanced methods. The proposed method has an improvement of 2.65% and 7.62% compared to 1D-BNCNN and Bi-LSTM, respectively, which shows that our method can better extract fault features. After the manual feature extraction, the classification accuracy of SBX-ABC-RF is higher than the results of three traditional machine learning methods and ABC-RF. It is clear that the proposed method can

effectively improve the classification performance and has the stronger ability to identify different fault modes.

Table 5. Comparison results of different methods in Case 2.

Methods	Accuracy (%)
1D-BNCNN	96.2
Bi-LSTM	90.23
MLP	89.72
SVM	83
KNN	85.02
ABC-RF	95.49
SBX-ABC-RF	97.85

5. Conclusions

This paper proposes a fault diagnosis method for rolling bearings combined with RVMD, Relief-F, and optimized RF. The original vibration signals are decomposed by RVMD to obtain decomposed signals in multiple frequency bands. Moreover, the multi-scale time–frequency domain features are extracted from different decomposed signals and the Relief-F is used to select highly sensitive features for fault diagnosis. To improve the classification performance of RF, an improved artificial bee colony algorithm is proposed based on the simulated binary crossover, which can realize the optimal key parameters of RF. Finally, several experiments are conducted on the Case Western Reserve University bearing dataset and the dataset collected from our rolling bearing fault testbed. The experimental results show that the proposed method achieves a significant improvement compared to the current state-of-the-art algorithms. In addition, compared with deep neural network methods, the proposed method has low requirements on hardware resources and has high interpretability. Therefore, the proposed method in this paper can be applied to the intelligent fault diagnosis of rolling bearings. In the future, the diagnosis of complex faults needs further study; other ensemble learning algorithms and deep learning methods may be utilized to further improve the classification ability.

Author Contributions: Formal analysis, Y.Y.; methodology, Y.Y. and J.Z.; software, J.Z.; supervision, H.W. and X.X.; validation, J.Z.; writing—original draft, J.Z.; Writing—review and editing, J.Z. and Y.Y. Data curation, J.Z., Y.H. and J.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Key Scientific Research Projects of the University in the Henan Province under Grant (24A413013), the Henan Provincial Scientific and Technological Tackling Key Problems Project (242102320215), the Henan Provincial High-end Foreign Experts Introduction Plan (HNGD2024032), and Zhongyuan University of Technology Discipline Strength Enhancement Plan Project (GG202412).

Data Availability Statement: The data generated during the current study are available from the first author or corresponding author on reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Yang, B.; Lei, Y.G.; Jia, F.; Xing, S.B. An intelligent fault diagnosis approach based on transfer learning from laboratory bearings to locomotive bearings. *Mech. Syst. Sig. Process.* **2019**, *122*, 692–706. [CrossRef]
2. Pan, H.; Yang, Y.; Li, X.; Zheng, J.; Cheng, J. Symplectic geometry mode decomposition and its application to rotating machinery compound fault diagnosis. *Mech. Syst. Sig. Process* **2019**, *114*, 189–211. [CrossRef]

3. Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.; Yen, N.C.; Tung, C.C.; Liu, H.H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **1998**, *454*, 903–955. [CrossRef]
4. Dragomiretskiy, K.; Zosso, D. Variational Mode Decomposition. *IEEE Trans. Signal Process* **2014**, *62*, 531–544. [CrossRef]
5. Li, G.Z.; Tang, G.; Luo, G.G.; Wang, H.Q. Underdetermined blind separation of bearing faults in hyperplane space with variational mode decomposition. *Mech. Syst. Sig. Process* **2019**, *120*, 83–97. [CrossRef]
6. Yan, X.A.; Jia, M.P. Application of CSA-VMD and optimal scale morphological slice bispectrum in enhancing outer race fault detection of rolling element bearings. *Mech. Syst. Sig. Process* **2019**, *122*, 56–86. [CrossRef]
7. Wu, Y.; Hu, Z.; An, W.; Li, X.; Wang, X.; Du, P.; Yu, X. Automatic diagnosis of rolling element bearing under different conditions based on RVMD and envelop order capture. *IEEE Access* **2019**, *7*, 91799–91808. [CrossRef]
8. Liu, Y.; Chai, Y.; Liu, B.; Wang, Y. Impulse signal detection for bearing fault diagnosis via residual-variational mode decomposition. *Appl. Sci.* **2021**, *11*, 3053. [CrossRef]
9. Song, J.; Zhao, J.; Zhang, X.; Dong, F.; Zhao, J.; Xu, L.; Yao, Z. Accurate demagnetization faults detection of dual-sided permanent magnet linear motor using enveloping and time-domain energy analysis. *IEEE Trans. Ind. Informat* **2020**, *16*, 6334–6346. [CrossRef]
10. Wang, M.; Feng, G.; Yan, H.; Qiu, J.; Zhang, H. Membership-function-dependent fault detection filtering design for interval type-2 T-S fuzzy systems in finite frequency domain. *IEEE Trans. Fuzzy Syst.* **2021**, *29*, 2760–2773. [CrossRef]
11. Tang, J.; Wu, J.; Qing, J.; Kang, T. Rolling Bearing Fault Monitoring for Sparse Time-Frequency Representation and Feature Detection Strategy. *Entropy* **2022**, *24*, 1822. [CrossRef]
12. Kankar, P.K.; Sharma, S.C.; Harsha, S.P. Fault diagnosis of ball bearings using machine learning methods. *Expert. Syst. Appl.* **2011**, *38*, 1876–1886. [CrossRef]
13. Han, C.; Mo, C.; Tao, L.; Ma, Y.; Bai, X. An efficient fatigue assessment model of offshore wind turbine using a half coupling analysis. *Ocean. Eng.* **2022**, *263*, 112318. [CrossRef]
14. Rauber, T.W.; de Assis Boldt, F.; Varejao, F.M. Heterogeneous Feature Models and Feature Selection Applied to Bearing Fault Diagnosis. *IEEE Trans. Ind. Electron.* **2015**, *62*, 637–646. [CrossRef]
15. De Moura, E.P.; Souto, C.D.R.; Silva, A.A.; Irmão, M.A.D.S. Evaluation of principle component analysis and neural network performance for bearing fault diagnosis from vibration signal processed by RS and DF analysis. *Mech. Syst. Signal Process* **2011**, *25*, 1765–1772. [CrossRef]
16. Zhao, X.; Jia, M. Fault diagnosis of rolling bearing based on feature reduction with global-local margin Fisher analysis. *Neurocomputing* **2018**, *315*, 447–464. [CrossRef]
17. Yue, L.; Ge, Z.Q. Weighted random forests for fault classification in industrial processes with hierarchical clustering model selection. *J. Process Control* **2018**, *64*, 62–70.
18. Kappaganthu, K.; Nataraj, C. Feature selection for fault detection in rolling element bearings using mutual information. *J. Vib. Acoust.* **2011**, *133*, 061001. [CrossRef]
19. VVakharia; Gupta, V.K.; Kankar, P.K. Efficient fault diagnosis of ball bearing using ReliefF and random forest classifier. *Mech. Sci. Eng.* **2017**, *39*, 2969–2982. [CrossRef]
20. Urbanowicz, R.J.; Meeker, M.; La Cava, W.; Olson, R.S.; Moore, J.H. Relief-based feature selection: Introduction and review. *J. Biomed. Inform.* **2018**, *85*, 189–203. [CrossRef]
21. Lu, Q.; Yang, R.; Zhong, M.; Wang, Y. An improved fault diagnosis method of rotating machinery using sensitive features and RLS-BP neural network. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 1585–1593. [CrossRef]
22. 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]
23. Rapur, J.S.; Tiwari, R. On-line time domain vibration and current signals based multi-fault diagnosis of centrifugal pumps using support vector machines. *J. Nondestr. Eval.* **2018**, *38*, 6. [CrossRef]
24. OSagi; Rokach, L. Ensemble learning: A survey. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2018**, *8*, 1249.
25. Wan, L.; Gong, K.; Zhang, G.; Yuan, X.; Li, C.; Deng, X. An Efficient Rolling Bearing Fault Diagnosis Method Based on Spark and Improved Random Forest Algorithm. *IEEE Access* **2021**, *9*, 37866–37882. [CrossRef]
26. Xu, G.; Liu, M.; Jiang, Z.; Söffker, D.; Shen, W. Bearing fault diagnosis method based on deep convolutional neural network and random forest ensemble learning. *Sensors* **2019**, *19*, 1088. [CrossRef] [PubMed]
27. Tahi, M.; Miloudi, A.; Dron, J.; Bouzouane, B. Decision tree and feature selection by using a genetic wrapper for fault diagnosis of rotating machinery. *Aust. J. Mech. Eng.* **2018**, *18*, 496–504. [CrossRef]
28. Tian, J.; Liu, L.; Zhang, F.; Ai, Y.; Wang, R.; Fei, C. Multi-domain entropy-random forest method for the fusion diagnosis of inter-shaft bearing faults with acoustic emission signals. *Entropy* **2020**, *22*, 57. [CrossRef]
29. Tang, G.; Pang, B.; Tian, T.; Zhou, C. Fault diagnosis of rolling bearings based on improved fast spectral correlation and optimized random forest. *Appl. Sci.* **2018**, *8*, 1859. [CrossRef]
30. Gao, W.-F.; Liu, S.-Y. A modified artificial bee colony algorithm. *Comput. Oper. Res.* **2012**, *39*, 687–697. [CrossRef]

31. Wang, Y.; Markert, R. Filter bank property of variational mode decomposition and its applications. *Signal Process* **2016**, *120*, 509–521. [CrossRef]
32. AMoshrefzadeh; Fasana, A. The Autogram: An effective approach for selecting the optimal demodulation band in rolling element bearings diagnosis. *Mech. Syst. Signal Process* **2018**, *105*, 294–318.
33. Sikonja, M.R.; Konoenko, I. Theoretical and empirical analysis of ReliefF and RReliefF. *Mach. Learn. J.* **2003**, *53*, 23–69. [CrossRef]
34. Chen, S.Z.; Yang, R.; Zhong, M.Y. Graph-based semi-supervised random forest for rotating machinery gearbox. *Control Eng. Pract.* **2021**, *117*, 104952. [CrossRef]
35. Yang, D.; Liu, Y.; Li, S.; Li, X.; Ma, L. Gear fault diagnosis based on support vector machine optimized by artificial bee colony algorithm. *Mech. Mach. Theory* **2015**, *90*, 219–229. [CrossRef]
36. CWRU Bearing Data Center. CWRU Rolling Bearing Dataset. Available online: https://engineering.case.edu/bearingdatacenter/download-data-file (accessed on 13 February 2025).
37. Chen, G.; Yuan, J.; Zhang, Y.; Zhu, H.; Huang, R.; Wang, F.; Li, W. Enhancing reliability through interpretability: A comprehensive survey of interpretable intelligent fault diagnosis in rotating machinery. *IEEE Access* **2024**, *12*, 103348–103379. [CrossRef]
38. Qi, B.; Li, Y.; Yao, W.; Li, Z. Application of EMD combined with deep learning and knowledge graph in bearing fault. *J. Signal Process. Syst.* **2023**, *95*, 935–954. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.