

Article

Compound Fault Diagnosis of Rolling Bearing Based on ACMD, Gini Index Fusion and AO-LSTM

Jie Ma * and Xinyu Wang 

Mechanical Electrical Engineering School, Beijing Information Science & Technology University, Beijing 100192, China; 2020020044@bistu.edu.cn

* Correspondence: mjbeijing@163.com

Abstract: Due to the symmetry of the rolling bearing structure and the rotating operation mode, it will cause the coupling modulation phenomenon when it is damaged in multiple places at the same time, which makes it difficult to accurately identify all kinds of faults. For such problems, a compound fault diagnosis method based on adaptive chirp mode decomposition (ACMD), Gini index fusion and long short-term memory (LSTM) neural network optimized by Aquila Optimizer (AO) is proposed. Firstly, a series of IMF components are obtained by decomposing the vibration signal by means of ACMD, and the required components are selected by using the correlation coefficient method. Then, the Gini index of the square envelope (GISE) and the Gini index of the square envelope spectrum (GISES) of each component are calculated, respectively, and they are fused to construct a highly dimensional feature matrix. Then, with the aim of solving the problem of difficult selection of LSTM hyperparameters, the AO-LSTM model is constructed. Finally, the feature matrix is divided into a training set and a test set. The training set is input into the model for training, and then the training network is used to predict the test set, and outputs diagnostic results. The simulation and experimental results show that the proposed method can achieve higher accuracy and stronger robustness, compared with the existing intelligent diagnosis methods for bearing compound faults.

Keywords: rolling bearing; compound fault diagnosis; ACMD; Gini index; AO-LSTM



Citation: Ma, J.; Wang, X. Compound Fault Diagnosis of Rolling Bearing Based on ACMD, Gini Index Fusion and AO-LSTM. *Symmetry* **2021**, *13*, 2386. <https://doi.org/10.3390/sym13122386>

Academic Editor: Zhiwen Chen, Hao Luo and Chao Cheng

Received: 15 November 2021
Accepted: 8 December 2021
Published: 10 December 2021

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



Copyright: © 2021 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

Rotating machinery is widely used in modern industry, and a rolling bearing is an important part of most rotating machinery and electrical equipment. Its running state directly affects the performance of the whole machine [1]. If the rolling bearing fails, it may cause heavy economic losses and serious casualties. Therefore, the accurate and automatic diagnosis of various bearing faults that may occur in rotating machinery is of great significance [2]. However, in the actual operation process of rotating machinery equipment, due to the action of external load and the degradation of bearing performance, the fault often does not appear alone, and often appears as a compound fault of mutual coupling and cross influence between signal components [3]. Compared with a single fault, a compound fault is more harmful to mechanical equipment and is more difficult to diagnose. The diagnosis method of a single fault is difficult to adapt to the diagnosis problem under the condition of a compound fault [4,5]. Therefore, the research on the intelligent diagnosis method of compound faults of rolling bearings is of great practical significance to ensure the normal and stable operation of mechanical equipment.

In order to detect the compound fault characteristics of rolling bearings under complex conditions such as high speed, heavy load and strong impact, advanced signal decomposition and filtering methods can be used to enhance and extract the compound fault signal. In recent years, more and more efficient and reliable signal processing algorithms have emerged, providing more choices for the diagnosis methods of bearing compound faults [6]. Cui et al. [7] effectively separated the compound faults of bearings by using a

combination of the second-generation wavelet and EMD demodulation. Chen et al. [8] proposed a method based on minimum entropy deconvolution and Teager energy operator to analyze the compound fault signal of a helicopter rolling bearing, and effectively extracted weak compound fault features despite strong background noise. Qi et al. [9] used adaptive signal sparse resonance decomposition (ARSD) and multipoint kurtosis optimal minimum entropy deconvolution correction (MK-MOMEDA) to effectively extract various fault features from compound faults. Wan Shuting et al. [10] used the combination of variational modal decomposition (VMD) and maximum correlation kurtosis deconvolution (MCKD) to extract fault features, which effectively realized the diagnosis of bearing compound faults. The above methods achieved certain results, providing a new idea for the compound fault diagnosis of rolling bearings. However, these methods need to know a lot of a priori information in advance, which is greatly hindered in practical engineering practice. Chen et al. [11] introduced an easy-to-handle adaptive chirp mode decomposition (ACMD) method. ACMD is a non-stationary signal decomposition method based on a variational modal decomposition (VMD) framework. It can estimate each signal component independently, solve the problem of the poor effect of VMD in processing multi-component signals with overlapping frequencies, and there is no need to set the number of signal components in advance, so it is more flexible in practical applications.

In addition, the proposal and improvement of mechanical fault feature extraction methods largely depend on the evaluation index of feature information. As an excellent sparse index, the Gini index (GI) has strong robustness to the interference of random impulse noise. Therefore, it is very appropriate to take it as the evaluation index of the component decomposed by the ACMD method. In 2017, GI was first introduced for the feature extraction of mechanical faults, and then GI was widely used in the field of fault diagnosis of gears and bearings [12]. For example, Miao et al. [13] used GI instead of the kurtosis coefficient to verify the robustness of GI to random impulse noise for the first time. Albezzawy and Nassef et al. [14,15] believe that GI is a more effective estimator than kurtosis, which is used in wavelet decomposition and variational modal decomposition to extract the best modulus. These studies further strengthen the theoretical support for GI in the field of feature extraction technology.

Since the vibration signals of various bearings are always nonlinear, non-stationary and accompanied by high background noise, it is a complex task to determine the most accurate diagnosis results from the characteristic signals only by manual experience [16,17]. Chen [18] used a combination of wavelet packet decomposition (WPD), overall average empirical mode decomposition (EEMD) and information entropy to extract the multiple features of the signal, and used the multi-classifier group composed of support vector machine (SVM) sub-classifiers to identify the multiple features, so as to realize the diagnosis and recognition of compound faults of rolling bearings. Zhang et al. [19] used the AdaBoost algorithm and the back propagation (BP) neural network to effectively identify the compound fault mode of gearbox. However, SVM and the BP neural network used in pattern recognition are shallow networks, the number of hidden layers and the ability of feature learning and expression are limited, and the training can easily fall into the local extremum. In recent years, Han et al. [20] used multi-wavelet transform (MWT) to process the bearing vibration signal, obtained multi-wavelet coefficient branches, constructed the feature map, and realized the intelligent diagnosis of rolling bearing compound faults by using a CNN classifier. Shi et al. [21] used the improved Hilbert–Huang transform (IHHT) to extract the time-frequency characteristics of fault signals, and inputted the generated time-frequency map into a convolutional neural network (CNN) for the intelligent identification of compound faults of rolling bearings.

As a typical deep learning network model, a CNN contains multiple hidden layers and can realize feature learning and expression through feature transfer layer by layer. However, in the application of fault identification, it needs a large number of two-dimensional data, and the network structure is complex, which requires more training time. A recurrent neural network (RNN) solves this kind of problem well. An RNN is a framework for

processing sequential data. It remembers the previous information through the connection structure of nodes between each layer, and uses this information to affect the output of subsequent nodes, which can fully mine the temporal information in sequential data. This method has more in-depth expression ability in processing complex feature data [22]. As a special RNN, a long short-term memory network (LSTM) can effectively avoid the gradient explosion problem. At present, LSTM has been preliminarily applied in the field of fault diagnosis, and has improved the level and efficiency of fault diagnosis. For example, Cao et al. [23] used the improved 1d-CNN-LSTM model to classify and identify six different working states of rolling bearings, and achieved high recognition accuracy. Elsewhere, Zhang et al. [24] used MCKD to denoise the vibration signal, and then used an LSTM network to train the calculated Teager energy sequence to realize bearing fault diagnosis. These studies were aimed at completing the bearing fault diagnosis of a single fault, but there is little research on applying LSTM to compound fault intelligent diagnosis.

In this paper, an LSTM network is used to identify the characteristic matrix obtained by ACMD and the GI fusion method, and the Aquila optimizer (AO) [25] is used to optimize the super parameters of the LSTM network for the intelligent diagnosis of compound faults of rolling bearings. As a new population-based optimization method, AO can make better use of the global solution to break away from the local minimum, and can converge to the optimal solution only by several iterations of the objective function. Compared with the ant colony algorithm and the particle swarm optimization algorithm, it does not easily fall into local optimization, and the convergence accuracy is higher. The design method proposed in this paper describes three main steps: firstly, the best feature vector is obtained by ACMD and the GI fusion method, and it is divided into training set and test set. Secondly, the characteristic matrix is divided into a training set and a test set, the LSTM network suitable for characteristic data is constructed, and the parameters of the LSTM network are adjusted by the AO algorithm. Finally, the training set is input into the AO-LSTM model for training, and the training network is used to classify and predict the test set. Compared with the existing intelligent diagnosis methods of bearing compound faults, this method avoids the need for diagnosis experience and can achieve higher fault identification accuracy.

2. Theoretical Basis of the Proposed Method

2.1. Adaptive Chirp Mode Decomposition (ACMD)

ACMD is a more easily processed form of VNCMD, and VNCMD is a new method to decompose chirp signals. A chirp signal consisting of a K chirp modulus can be modeled as:

$$s(t) = \sum_{i=1}^K s_i(t) = \sum_{i=1}^K A_i(t) \cos\left(2\pi \int_0^t f_i(\tau) d\tau + \theta_i\right) \quad (1)$$

where $A_i(t) > 0$, $f_i(t) > 0$ and θ_i represent the instantaneous amplitude (IA), instantaneous frequency (IF) and initial phase of i th component $s_i(t)$, respectively.

Using demodulation technology, Equation (1) can be rewritten as:

$$s(t) = \sum_{i=1}^K a_i(t) \cos\left(2\pi \int_0^t \tilde{f}_i(\tau) d\tau\right) + b_i(t) \sin\left(2\pi \int_0^t \tilde{f}_i(\tau) d\tau\right) \quad (2)$$

in which

$$\begin{aligned} a_i(t) &= -A_i(t) \cos\left(2\pi \int_0^t (f_i(\tau) - \tilde{f}_i(\tau)) d\tau + \theta_i\right), \\ b_i(t) &= -A_i(t) \sin\left(2\pi \int_0^t (f_i(\tau) - \tilde{f}_i(\tau)) d\tau + \theta_i\right). \end{aligned} \quad (3)$$

where $\tilde{f}_i(t)$ is the frequency function of $\cos(2\pi \int_0^t \tilde{f}_i(\tau) d\tau)$ and $\sin(2\pi \int_0^t \tilde{f}_i(\tau) d\tau)$; $a_i(t)$ and $b_i(t)$ are demodulated signals, their IF is $f_i(t) - \tilde{f}_i(t)$, and IA is:

$$A_i(t) = \sqrt{a_i^2(t) + b_i^2(t)} \quad (4)$$

From the above analysis, it can be seen that when VNCMD estimates the signal components, the number K of components should be known. However, it is difficult to obtain this information. The difference between ACMD and VNCMD is that it is inspired by the matching pursuit method and uses the greedy algorithm to estimate the signal components one by one [11]. For example, for the i th signal component, the following problems shall be solved:

$$\min_{a_i(t), b_i(t), \tilde{f}_i(t)} \left(\|a''_i(t)\|_2^2 + \|b''_i(t)\|_2^2 + \alpha \|s(t) - s_i(t)\|_2^2 \right), \tag{5}$$

$$s_i(t) = a_i(t) \cos\left(2\pi \int_0^t \tilde{f}_i(\tau) d\tau\right) + b_i(t) \sin\left(2\pi \int_0^t \tilde{f}_i(\tau) d\tau\right).$$

where $\|s(t) - s_i(t)\|_2^2$ represents the residual energy after removing the current estimated component and $\alpha > 0$ represents the weighting coefficient.

As with matching tracking, ACMD finds the signal component that can take the most energy from the input signal. Assuming that the signal is discretized at $t = t_0, \dots, t_{N-1}$, the discrete form of Equation (4) can be expressed as:

$$\min_{\mathbf{u}_i, \mathbf{f}_i} \left\{ \|\Theta \mathbf{u}_i\|_2^2 + \alpha \|\mathbf{s} - \mathbf{G}_i \mathbf{u}_i\|_2^2 \right\} \tag{6}$$

where $\Theta = \begin{bmatrix} \Omega & \\ & \Omega \end{bmatrix}$, Ω is a second-order difference matrix; $\mathbf{u}_i = [\mathbf{a}_i^T, \mathbf{b}_i^T]^T$, $\mathbf{a}_i = [a_i(t_0), \dots, a_i(t_{N-1})]^T$, $\mathbf{b}_i = [b_i(t_0), \dots, b_i(t_{N-1})]^T$; $\mathbf{s} = [s(t_0), \dots, s(t_{N-1})]^T$;

$$\mathbf{G}_i = [\mathbf{C}_i, \mathbf{S}_i] \tag{7}$$

$$\mathbf{C}_i = \text{diag}[\cos(\varphi_i(t_0)), \dots, \cos(\varphi_i(t_{N-1}))] \tag{8}$$

$$\mathbf{S}_i = \text{diag}[\sin(\varphi_i(t_0)), \dots, \sin(\varphi_i(t_{N-1}))] \tag{9}$$

where $\varphi_i(t) = 2\pi \int_0^t \tilde{f}_i(\tau) d\tau$.

Equation (5) shows that given $\tilde{f}_i(\tau)$ or \mathbf{G}_i , the optimization problem can be solved by using an iterative algorithm that alternately updates the demodulated signal and frequency function. For the j th iteration, the vector \mathbf{u}_i can be updated to:

$$\mathbf{u}_i^j = \begin{bmatrix} \mathbf{a}_i^j \\ \mathbf{b}_i^j \end{bmatrix} = \underset{\mathbf{u}_i}{\text{argmin}} \left\{ \|\Theta \mathbf{u}_i\|_2^2 + \alpha \|\mathbf{s} - \mathbf{G}_i^j \mathbf{u}_i\|_2^2 \right\} = \left(\frac{1}{\alpha} \Theta^T \Theta + (\mathbf{G}_i^j)^T \mathbf{G}_i^j \right)^{-1} (\mathbf{G}_i^j)^T \mathbf{s} \tag{10}$$

where \mathbf{G}_i^j is composed of $\tilde{f}_i^j(t)$, j represents the iteration counter, and the parameters α can adjust the smoothness of the output signal. The signal component can be estimated as:

$$\mathbf{s}_i^j = \mathbf{G}_i^j \mathbf{u}_i^j \tag{11}$$

According to Equation (3) and using the demodulated signal in Equation (10), the frequency increment can be calculated as:

$$\Delta \tilde{f}_i^j(t) = -\frac{1}{2\pi} \frac{d}{dt} \left(\arctan \left(\frac{b_i^j(t)}{a_i^j(t)} \right) \right) = \frac{b_i^j(t) \times (a_i^j(t))' - a_i^j(t) \times (b_i^j(t))'}{2\pi (a_i^j(t))^2 + 2\pi (b_i^j(t))^2} \tag{12}$$

In practical application, a low-pass filter is needed to preprocess the frequency increment to reduce noise interference. IF can eventually be updated to:

$$\mathbf{f}_i^{j+1} = \mathbf{f}_i^j + \left(\frac{1}{\beta} \Omega^T \Omega + \mathbf{I} \right)^{-1} \Delta \tilde{\mathbf{f}}_i^j \tag{13}$$

where $\mathbf{f}_i^j = [\tilde{f}_i^j(t_0), \dots, \tilde{f}_i^j(t_{N-1})]^T$; \mathbf{I} is the identity matrix; $(\frac{1}{\beta}\mathbf{\Omega}^T\mathbf{\Omega} + \mathbf{I})^{-1}$ is used as low-pass filter, $\Delta\tilde{\mathbf{f}}_i^j = [\Delta\tilde{f}_i^j(t_0), \dots, \Delta\tilde{f}_i^j(t_{N-1})]^T$, the parameter β can adjust the smoothness of IF.

Next, we can use \mathbf{f}_i^{j+1} to update matrix \mathbf{G}_i^{j+1} and repeat Equations (10)–(13) in the next iteration. The algorithm can not only estimate the signal component $\tilde{s}_i(t)$, but also estimate the IA and IF of the signal (i.e., $\tilde{A}_i(t)$ and $\tilde{f}_i(t)$). In order to find other signal components, the first estimate $\tilde{s}_1(t)$ shall be removed from the original signal as follows:

$$R_1(t) = s(t) - \tilde{s}_1(t) \tag{14}$$

where $R_1(t)$ represents the residual signal after removing the first estimated component.

Then, the second signal component $\tilde{s}_2(t)$ is subtracted from $R_1(t)$. After K iterations, the decomposition of the signal $s(t)$ is as follows:

$$s(t) = \sum_{i=1}^K \tilde{s}_i(t) + R_K(t) \tag{15}$$

In fact, if the energy of the remaining $R_K(t)$ is less than a certain threshold, the algorithm is stopped.

2.2. Gini Index (GI)

As a new concept in mechanical fault diagnosis, the Gini index (GI) has the most stable gradient characteristics and is an effective evaluation index compared with traditional indexes. Equation (16) is the GI of data \mathbf{x} with length N . $\mathbf{x} = [x_{(1)}x_{(2)} \cdots x_{(n)} \cdots x_{(N)}]$ is a sequence sorted in ascending order (i.e., $x_{(1)} \leq x_{(2)} \cdots \leq x_{(n)} \cdots \leq x_{(N)}$). This sorting process imposes some limitations on the application of GI.

$$GI = 1 - 2 \sum_{n=1}^N \frac{x_{(n)}}{\|\mathbf{x}\|_1} \left(\frac{N - n + 0.5}{N} \right) \tag{16}$$

Miao et al. [26] proposed a new index based on GI, which can be directly applied. For example, the Gini index of the square envelope (GISE) of the signal $\mathbf{s} = (s_1 s_2 \cdots s_N)$ is defined as:

$$GISE = 1 - 2 \sum_{n=1}^N \frac{SE_{(n)}}{\|\mathbf{SE}\|_1} \left(\frac{N - n + 0.5}{N} \right) \tag{17}$$

where $\|\cdot\|_1$ denotes the l_1 norm operation, and square envelope of signal \mathbf{s} is $\mathbf{SE} = |s|^2$ and its order $\mathbf{SE}_{\text{order}} = [SE_{(1)} SE_{(2)} \cdots SE_{(n)} \cdots SE_{(N)}]$, $SE_{(1)} \leq SE_{(2)} \cdots \leq SE_{(n)} \cdots \leq SE_{(N)}$.

The Gini index of the square envelope spectrum (GISES) of the signal \mathbf{s} is defined as:

$$GISES = 1 - 2 \sum_{n=1}^N \frac{SES_{(n)}}{\|\mathbf{SES}\|_1} \left(\frac{N - n + 0.5}{N} \right) \tag{18}$$

where the square envelope spectrum of signal \mathbf{s} is $\mathbf{SES} = \text{abs}(\text{FFT}(|s|^2))$, and FFT is fast Fourier transform.

2.3. Aquila Optimizer (AO)

The optimization rule of the AO algorithm starts from the population of the candidate solution (X) in Equation (19), which is randomly generated between the upper bound

(UB) and the lower bound (LB) of a given problem. The best-obtained solution, so far, is determined as the optimal solution approximately in each iteration.

$$X = \begin{bmatrix} x_{1,1} & \cdots & x_{1,j} & x_{1,Dim-1} & x_{1,Dim} \\ x_{2,1} & \cdots & x_{2,j} & \cdots & x_{2,Dim} \\ \cdots & \cdots & x_{i,j} & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{N-1,1} & \cdots & x_{N-1,j} & \cdots & x_{N-1,Dim} \\ x_{N,1} & \cdots & x_{N,j} & x_{N,Dim-1} & x_{N,Dim} \end{bmatrix} \quad (19)$$

where X denotes a set of current candidate solutions, which are generated randomly by using Equation (20), X_i denotes the decision values (positions) of the i th solution, N is the total number of candidate solutions (population), and Dim denotes the dimension size of the problem.

$$X_{ij} = rand \times (UB_j - LB_j) + LB_j, i = 1, 2, \dots, N, j = 1, 2, \dots, Dim. \quad (20)$$

where $rand$ is a random number, LB_j denotes the j th lower bound, and UB_j denotes the j th upper bound of the given problem.

AO algorithm simulates Aquila's behavior during hunting [25]. Aquila's behavior is modeled as a mathematical optimization paradigm, and the optimal solution is determined under specific constraints. Aquila's behavior in the first method (X_1) is mathematically shown in Equation (21).

$$X_1(t+1) = X_{best}(t) \times \frac{1-t}{T} + (X_M(t) - X_{best}(t) * rand) \quad (21)$$

where $X_1(t+1)$ is the solution of the next iteration of t , which is generated by the first search method (X_1). $X_{best}(t)$ is the best-obtained solution until the t th iteration; this reflects the approximate location of the prey. $\frac{1-t}{T}$ is used to control the extended exploration by the number of iterations. $X_M(t)$ denotes the mean location value of the current solutions connected at the t th iteration, which is calculated using Equation (22).

$$X_M(t) = \frac{1}{N} \sum_{i=1}^N X_i(t), \forall j = 1, 2, \dots, Dim. \quad (22)$$

where Dim is the dimension size of the problem and N is the number of candidate solutions.

The AO algorithm starts to improve the optimization process by generating a set of randomly predefined candidate solutions. According to the optimal solution obtained in the optimization process, the individual position is updated, and four strategies (i.e., expanding exploration, narrowing exploration scope, expanding development and narrowing development scope) are used to highlight the balance between exploration and development. See reference [25] for the detailed process of the other three strategies. When the end conditions are met, the exploration process of AO is terminated.

2.4. Long Short-Term Memory (LSTM) Network

A long short-term memory (LSTM) network is a special recurrent neural network (RNN). In order to solve the problems of gradient disappearance or gradient explosion of RNN, LSTM adds a forget gate and input gate to selectively retain and update information, which strengthens the learning ability of the network [27]. The structure of a LSTM network is shown in Figure 1. It is composed of multiple neural network cells.

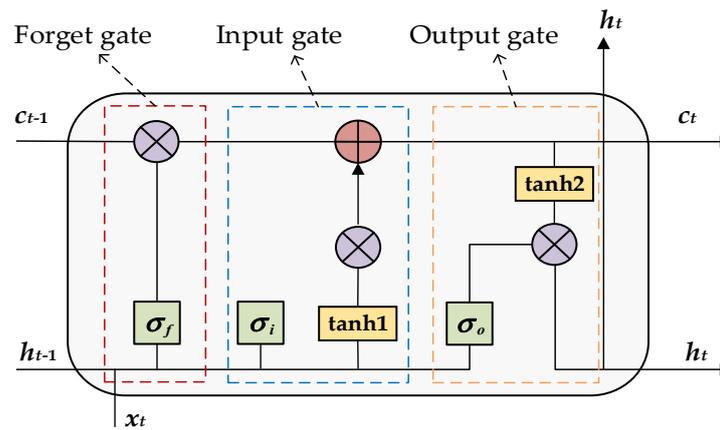


Figure 1. Structure diagram of LSTM neural network unit.

In Figure 1, c_{t-1} denotes the state of the previous cell, h_{t-1} denotes the output of the previous cell layer, and x_t denotes the input of the current cell; σ_f , σ_i and σ_o denote the forget gate, input gate and output gate, respectively; tanh1 is used to generate updated content, and tanh2 is mainly used to update the cell state of the neural network at that moment [28].

The general form of gate control is defined as:

$$g(x) = \sigma(Wx + b) \quad (23)$$

where $\sigma(x) = 1/(1 + \exp(-x))$; W and b denote the weight matrix and bias vector of the network, respectively.

The forward calculation process of LSTM can be expressed by Equations (24)–(28).

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \quad (24)$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \quad (25)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (26)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \quad (27)$$

$$h_t = o_t \odot \tanh(c_t) \quad (28)$$

where $\sigma(x)$ is the sigmoid function. It can map a real value to the interval 0~1 to describe how much information passes through. W and b represent the weight matrix and bias vector of the network. \odot represents the scalar product of two vectors. i_t , f_t , and o_t represent the input gate, forget gate and output gate, respectively. x_t and h_t represent the input and output of the hidden layer of LSTM, respectively. c_t denotes the memory unit.

When the time step is t , the input and output vectors of the hidden layer are x_t and h_t , respectively, and the memory unit is c_t . The input gate is used to control how much x_t flows into the memory unit—that is, how much can be saved to c_t . A forget gate is a key component of an LSTM unit, which can control which information should be retained and which should be forgotten, and avoid the gradient disappearance and explosion caused by gradient back propagation over time in some way. The output gate controls the influence of the memory unit on the current output value—that is, which part of the memory unit will be output in the time step.

3. Compound Fault Diagnosis Method of Rolling Bearings Based on ACMD, Gini Index Fusion and AO-LSTM

This paper proposes using ACMD to decompose the original vibration signal of a rolling bearing into several intrinsic mode functions (IMF), selecting four IMF components with large correlation coefficient, calculating the Gini index of the square envelope (GISE) and the Gini index of the square envelope spectrum (GISES) of each component,

respectively, and fusing them into a high-dimensional characteristic matrix to evaluate the characteristics of signal components more accurately. Then, the LSTM network is introduced to identify the compound fault from the characteristic matrix, and the Aquila optimizer (AO) [25] is used to optimize the super parameters of the LSTM network, so as to improve the accuracy of the intelligent diagnosis of compound faults of rolling bearings.

The specific implementation steps of the proposed method are as follows:

Step 1: Data preprocessing. The mathematical models of bearing inner ring and outer ring fault signals are established, and they are combined to form compound fault signals. The corresponding Gaussian white noise is added to these three types to form the simulation signals of three types of bearing faults. Using the public bearing dataset of Xi'an Jiaotong University, five typical bearing vibration signals are extracted from the life cycle data as experimental signals;

Step 2: Feature extraction. All kinds of fault signals are decomposed by the ACMD algorithm to obtain K IMF components, n IMF components with large correlation coefficient are selected, their GISE and GISEs are calculated, respectively, and a $2N$ dimensional characteristic matrix is constructed by fusion;

Step 3: Initializing the LSTM network. The feature matrix is divided into training set and test set, and the corresponding labels are given, respectively. The sample data in line with the LSTM network input format are constructed to eliminate the influence of differences between features. The training environment, training function and gradient descent optimization algorithm are set. In addition, the input layer dimension, the number of output layer nodes, the number of hidden layer neuron nodes and other parameters are also set.

Step 4: Selection and definition of AO optimization parameters. The LSTM is taken as the fitness function, the reconstruction error of the network is taken as the fitness value, the number of hidden layer nodes and initial learning rate are taken as the optimization objectives, and the optimization range is set. The selection of population number and maximum optimization times is based on experience.

Step 5: AO optimizing the process of the LSTM network. According to the set super parameter range, the initial parameter value is randomly generated, the LSTM network is trained, the reconstruction error output by the fitness function is used to correct the optimization objective of AO, and the LSTM network is trained again with the corrected parameters to minimize the error value. It is judged whether the maximum optimization times are reached. If so, the execution is stopped and the currently selected optimal parameter combination and the corresponding optimal fitness value are output; if not, the target parameters are updated and training continues until the maximum optimization times are reached.

Step 6: Classification and identification of signals. The optimal solution of the optimized output is used to update the parameters of the LSTM network, input the training set into the network for retraining, then use the training network to predict the training set and test set, and output the final classification diagnosis result of the bearing signal.

The detailed diagnosis flow of the method proposed in this paper is shown in Figure 2.

Based on the above research, compared with the commonly used decomposition methods such as EMD and VMD, the ACMD method has stronger robustness, simpler parameter setting, and a better effect in processing complex multi-component signals. Compared with the deconvolution method used frequently at present, the methods of ACMD and GI fusion do not need to know a lot of prior information regarding the vibration signal in advance, which will be more conducive to practical engineering application. GI, as a new sparse index, is more effective than the traditional kurtosis and correlation coefficients. In addition, the AO-LSTM model is introduced for the first time to identify the characteristics of compound faults of rolling bearing. Compared with shallow neural networks and existing deep learning network models, this method can achieve higher identification accuracy and stronger stability.

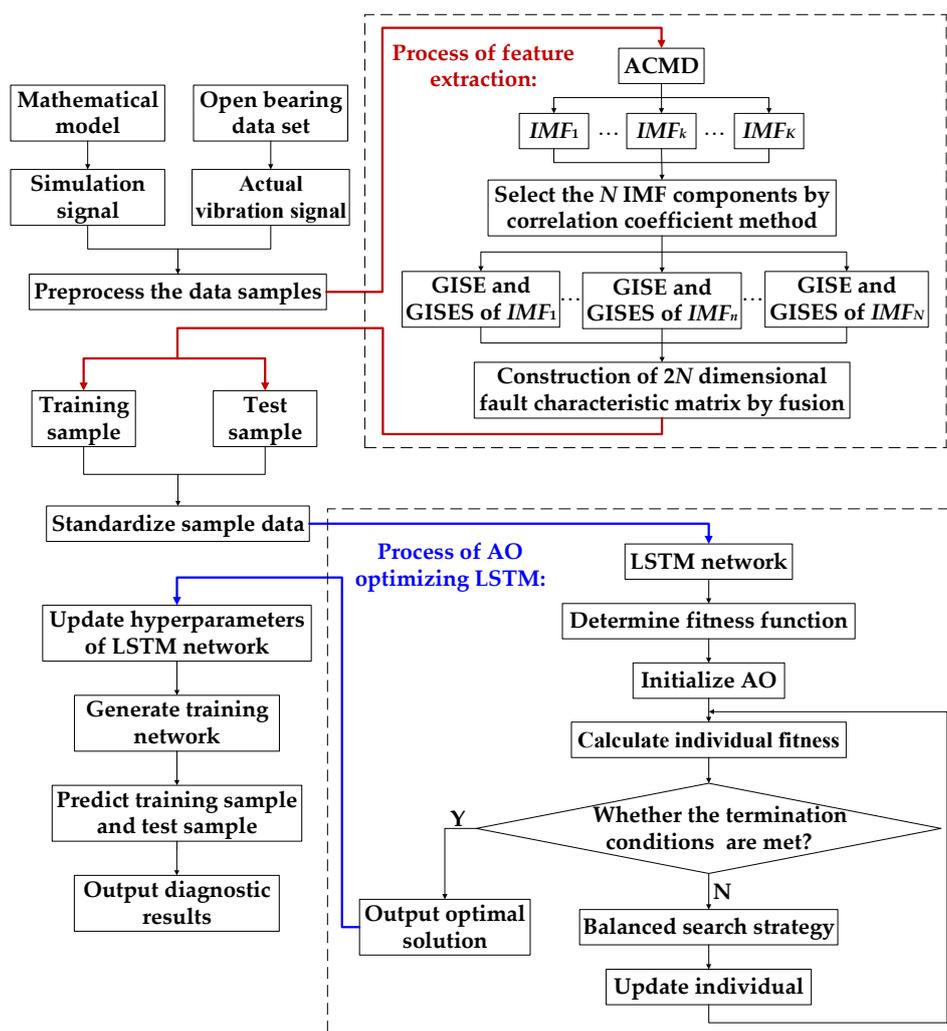


Figure 2. Flow chart of compound fault diagnosis of rolling bearing.

4. Simulation and Experimental Results of Rolling Bearing Fault Diagnosis

4.1. Simulation Verification

According to the characteristics of the failure of rolling bearings on inner rings and outer rings, the fault simulation signal $x(t)$ is constructed, and its expression is:

$$\begin{aligned}
 x(t) &= s(t) + n(t) = \sum_{i=1}^N A_i h(t - iT - \tau_i) + n(t), \\
 A_i &= A_0 \cos(2\pi f_r t + \varphi_A) + C_A, \\
 h(t) &= e^{-\beta t} \sin(2\pi f_n t + \varphi_\omega).
 \end{aligned} \tag{29}$$

where $s(t)$ is the periodic impact component, $n(t)$ is the Gaussian white noise, A_i is the modulation amplitude, and T is the impact period; f_r is the frequency conversion and C_A is the amplitude random constant; β is the attenuation index, f_n is the resonance frequency of the system, and the random variable τ_i is the small fluctuation of the i th impact relative to T , which follows the zero mean normal distribution, and the standard deviation is 0.5% of the conversion frequency.

The parameters of the inner ring fault simulation signal $x_i(t)$ are set as follows: the initial value of amplitude is $A_0 = 0.2$, the amplitude random constant $C_A = 2$, attenuation index $\beta = 700$, the natural frequency of the system $f_n = 4000$ Hz, the conversion frequency $f_r = 30$ Hz, the sampling frequency $f_s = 16,000$ Hz, and the characteristic frequency $f_i = 120$ Hz. The parameters of the outer ring fault simulation signal $x_o(t)$ are set as follows: the value of

amplitude is $A_i = 2$, $\beta = 700$, $f_s = 16,000$ Hz, $f_n = 4000$ Hz, and the characteristic frequency $f_o = 100$ Hz. In order to simulate the real compound fault signal of rolling bearing, the inner ring and outer ring simulation signals are superimposed, and then the Gaussian white noise $n(t)$ with a signal-to-noise ratio of 3 dB is added to form the compound fault simulation signal. The simulated fault signal and compound fault signal of the inner and outer ring of the bearing are shown in Figure 3.

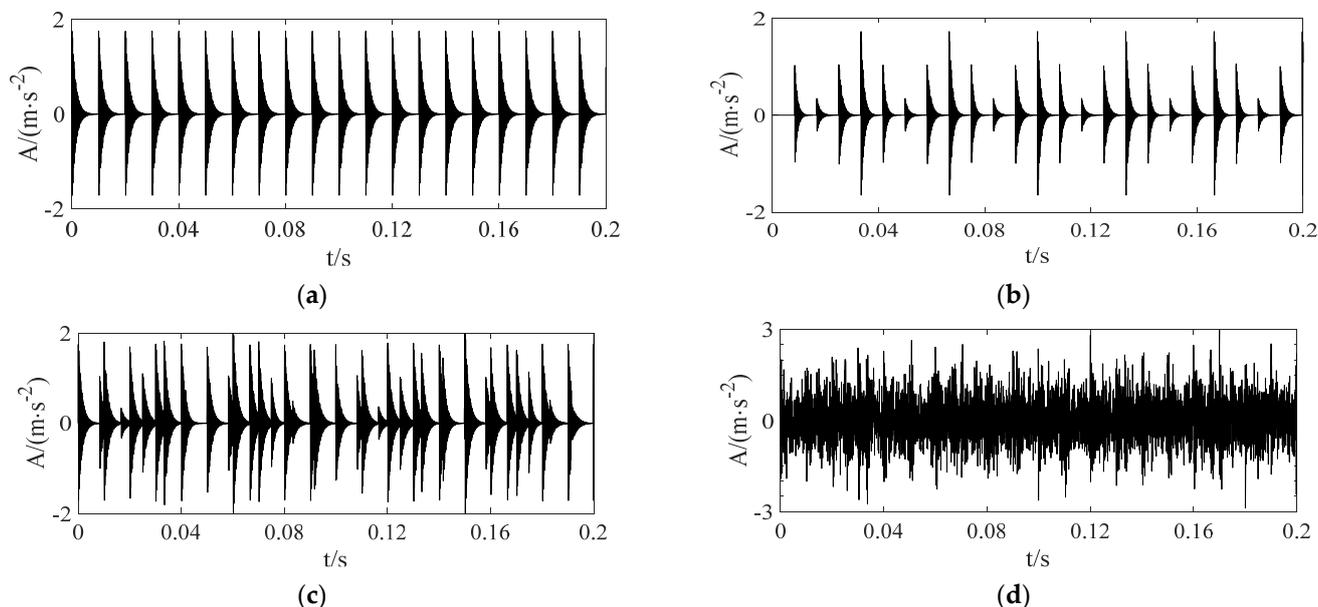


Figure 3. Time domain waveform of simulation signals. (a) Outer ring fault simulation signal; (b) inner ring fault simulation signal; (c) compound fault simulation signal; (d) compound fault simulation signal after adding 3 dB of noise.

The simulation signals of the rolling bearing inner ring, outer ring and compound fault are divided into 50 samples, respectively. A total of 150 samples are obtained in three states, in which each sample contains 4096 sampling points. Under the condition of setting the same parameters, 14 IMF components can be obtained by means of ACMD decomposition of each signal. Taking the fault signal of rolling bearing inner ring as an example, the time domain waveform of the first eight IMF components is taken, as shown in Figure 4. As can be seen from Figure 4, the IMF component still contains a large amount of noise interference. In this paper, the correlation coefficient method is used to reduce the error caused by this interference. The calculation result of taking one sample for each type of signal is shown in Figure 5.

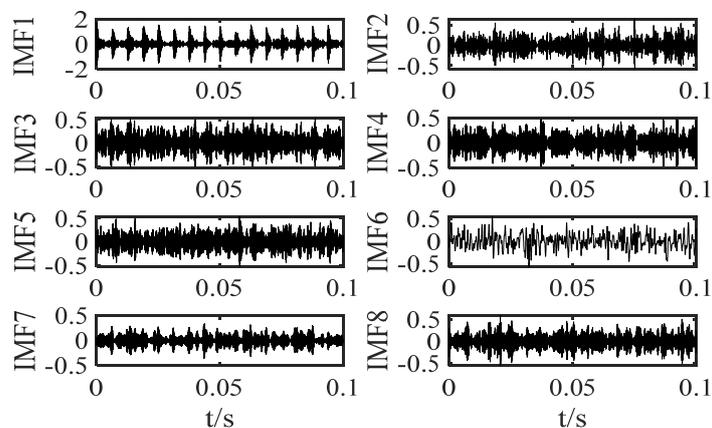


Figure 4. ACMD decomposition results.

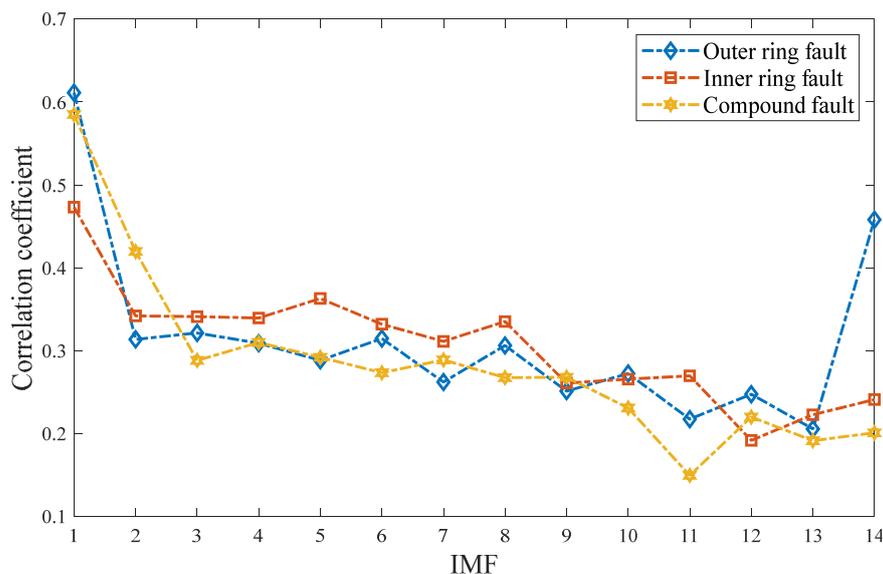


Figure 5. Correlation coefficient of each component under 3 bearing states.

Then, we select four components with large correlation coefficients, IMF1, IMF2, IMF3 and IMF4, and calculate their GISE and GISES values, respectively. The results are shown in Figure 6. It can be found that if only GISE is selected as the feature index, the features are prone to overlapping in the component IMF4; if only GISES is selected as the feature index, the features are prone to aliasing in the components IMF1 and IMF2.

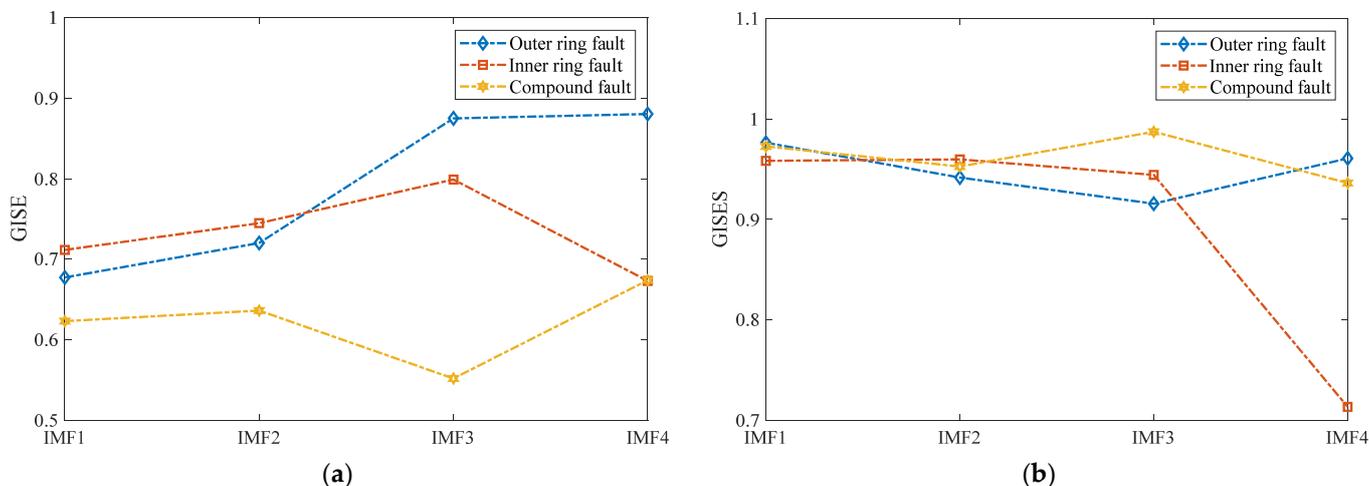


Figure 6. Comparison of the Gini index in different states. (a) GISE; (b) GISES.

This paper selects the GISE and GISES of the four IMF components as the feature index, which can effectively avoid the feature overlap of the four components, so as to ensure the accuracy of the feature information. Therefore, the GISE and GISES of four components are fused to form a vector of 1×8 , and each signal can use a characteristic matrix of 50×8 .

From the feature information containing 150 samples, 60 samples are randomly selected as the training set and 90 samples as the test set. The feature samples of the bearing inner ring, outer ring and compound fault in the training set and test set are labeled with 1, 2 and 3, respectively, and standardized. The corresponding AO-LSTM model is built, inputting the training set into the model for training, and generating the training network. Finally, the training network is used to classify and predict the divided training set and test set. The confusion matrix of the proposed method is shown in Figure 7.

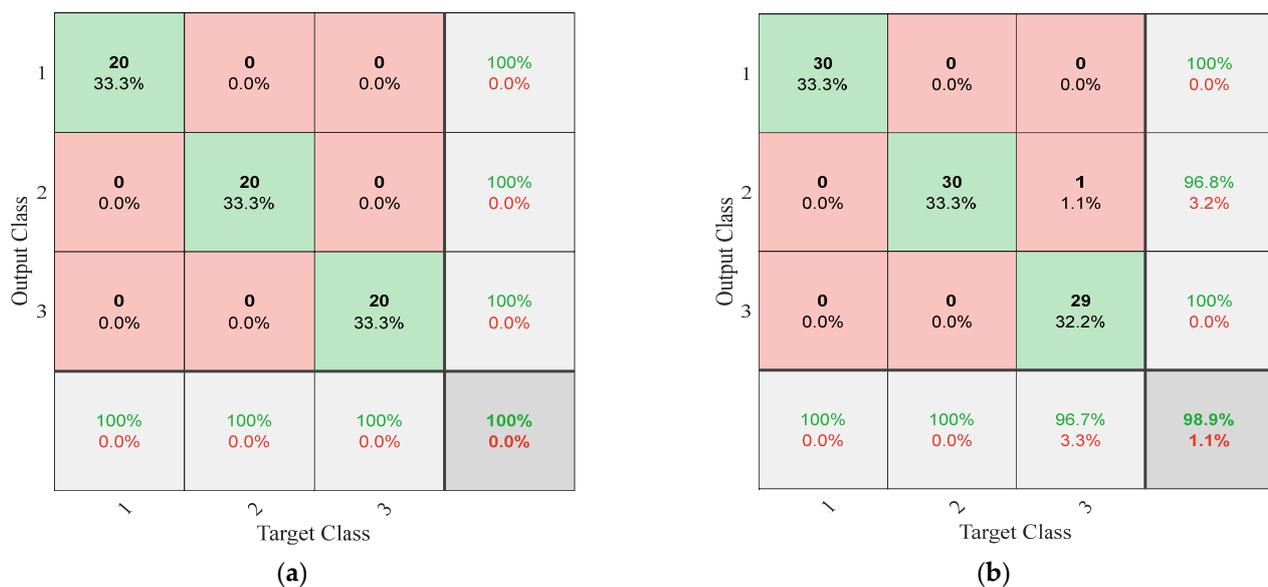


Figure 7. Confusion matrix of the proposed method. (a) Training accuracy; (b) Test accuracy.

In the confusion matrix, there is no classification error in the training samples, and the training accuracy of bearing fault signal is 100%. Only one test sample has classification error. The compound fault is diagnosed as an outer ring fault, and the test accuracy is 98.89%, of which the diagnosis accuracy of compound fault is 96.67%, which reflects the effectiveness of the proposed method to diagnose compound fault signal.

4.2. Experimental Verification

The bearing dataset of Xi'an Jiaotong University were selected to analyze and verify the proposed method, and compared with other intelligent diagnosis methods to verify the superiority of the proposed method.

4.2.1. Experimental Data Description

The test bench used to collect data is shown in Figure 8. The platform was mainly composed of a hydraulic loading system, a support bearing, a test bearing, a motor and a speed controller. The test bearing was a LDK UER204 rolling bearing, and its relevant parameters are shown in Table 1. Three types of working conditions were designed in the test, and their rotating speeds were 2100, 2250 and 2400 r/min, respectively [29]. There were five bearings under each type of working condition, in which the sampling frequency was set at 25.6 kHz, the sampling interval was 1 min, and the sampling time was 1.28 s. More details of this dataset can be found in [30].

The vibration signals collected in the experiment comprised the data of rolling bearings from a normal state to failure. In this paper, four groups of fault datasets were used to analyze and verify the proposed method. The signals of these different bearing conditions include different fault components, fault types and fault trends. Taking the compound fault of inner ring and outer ring as an example, the time domain waveform of its life cycle signal is shown in Figure 9. The fault signal of the dotted line part is intercepted. It can be seen that the characteristic amplitude of this part shows an increasing trend with the passage of operation time.

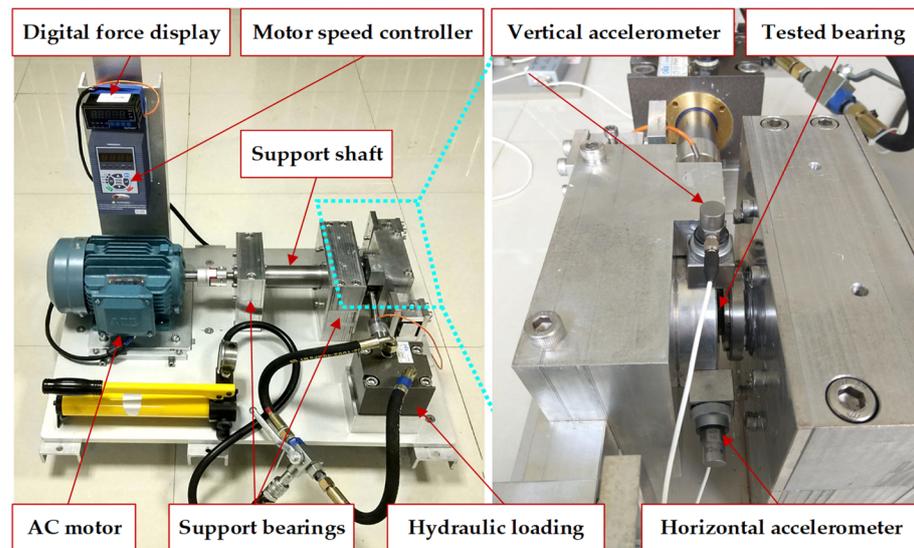


Figure 8. Rolling bearing test bench.

Table 1. LDK UER204 bearing parameters.

Parameter	Value	Parameter	Value
Inner race diameter	29.30 mm	ball diameter	7.92 mm
Outer race diameter	39.80 mm	Contact angle	0°
Bearing mean diameter	34.55 mm	Number of balls	8
Load rating	12.82 kN	Load rating	6.65 kN

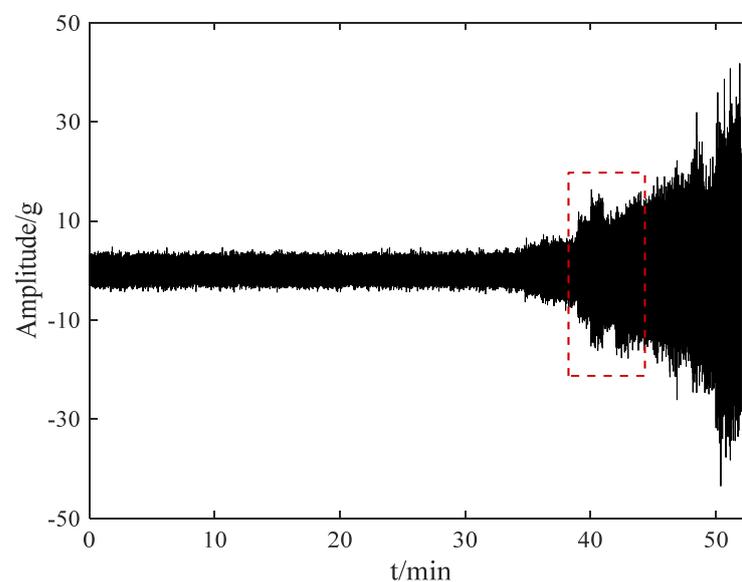


Figure 9. Life cycle vibration signal of bearing in compound fault state.

By analyzing the vibration amplitude of each bearing, several typical bearing signals such as outer ring faults, cage faults and inner ring faults could also be extracted. Fifty samples were created for each type of signal, and each sample contained 2048 sampling data points. One sample of each type of bearing state signal was taken, and their original time signal was drawn. The waveform of the five intercepted samples is shown in Figure 10. The health state signal is the health state data of the front part of the intercepted inner ring fault life cycle dataset. It can be seen that these five kinds of signals under different operating states show nonlinear and non-stationary characteristics, and the type of fault cannot be distinguished only by observing the time domain waveform.

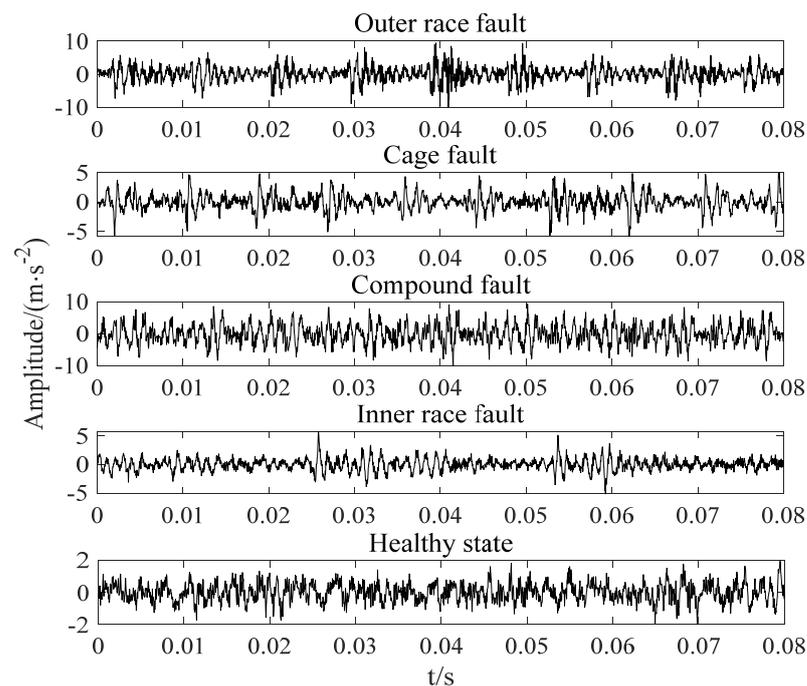


Figure 10. Time domain waveform of different state signals of rolling bearing.

4.2.2. Diagnostic Results of the Proposed Method

The accuracy of fault diagnosis is largely affected by the quality of fault features. This paper extracts effective feature vectors before intelligent diagnosis. Firstly, under the condition of setting the same parameters, ACMD is used to decompose each sample. Taking the fault signal of the outer ring of a rolling bearing as an example, one sample is taken and is decomposed to obtain 11 IMF components, and the first eight components are shown in Figure 11.

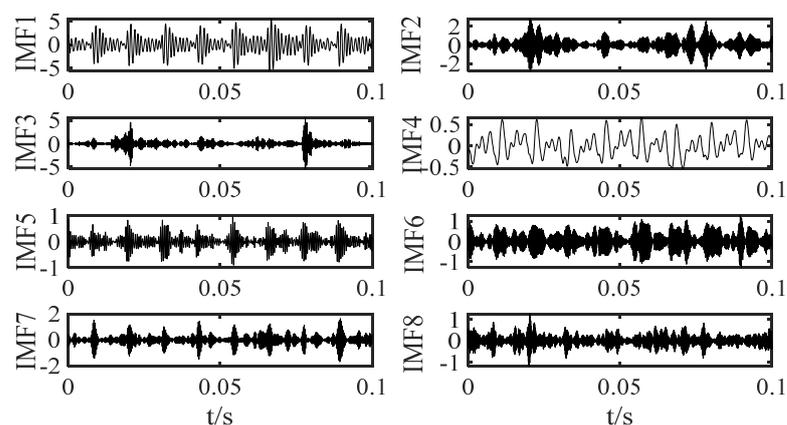


Figure 11. Decomposition results by ACMD.

It can be seen that there are still many interference components in the IMF component and taking these components as feature components will inevitably reduce the recognition accuracy of classification algorithm. Therefore, it is necessary to select the component more in line with the original signal. Here, the correlation coefficient method is used to filter the components that are less related to the original signal.

The calculation results of the correlation coefficients of each component are shown in Figure 12. Then, select four components with large correlation coefficient and calculate the GISE and GISES of these four components, respectively. The results are shown in Figure 13. It can be found that if only the GISE or GISES of each component is selected as

the feature vector, almost every component will be prone to aliasing, which greatly hinders the identification of fault features. However, if the GISE and GISES of four components are fused to construct an eight-dimensional feature vector, it can effectively avoid the overlap of more feature information and improve the accuracy of feature information. Each type of bearing signal is composed of 50 samples. Therefore, this paper uses a matrix of 50×8 to represent each type of bearing signal.

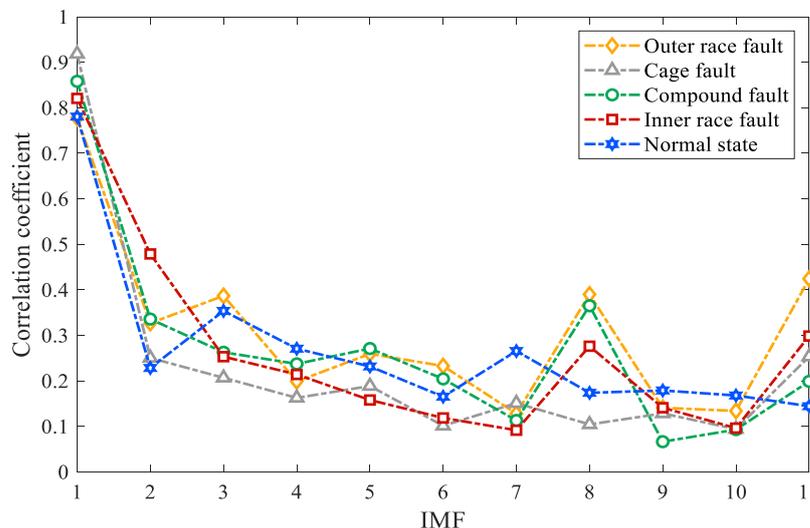


Figure 12. Correlation coefficient of each component in different states.

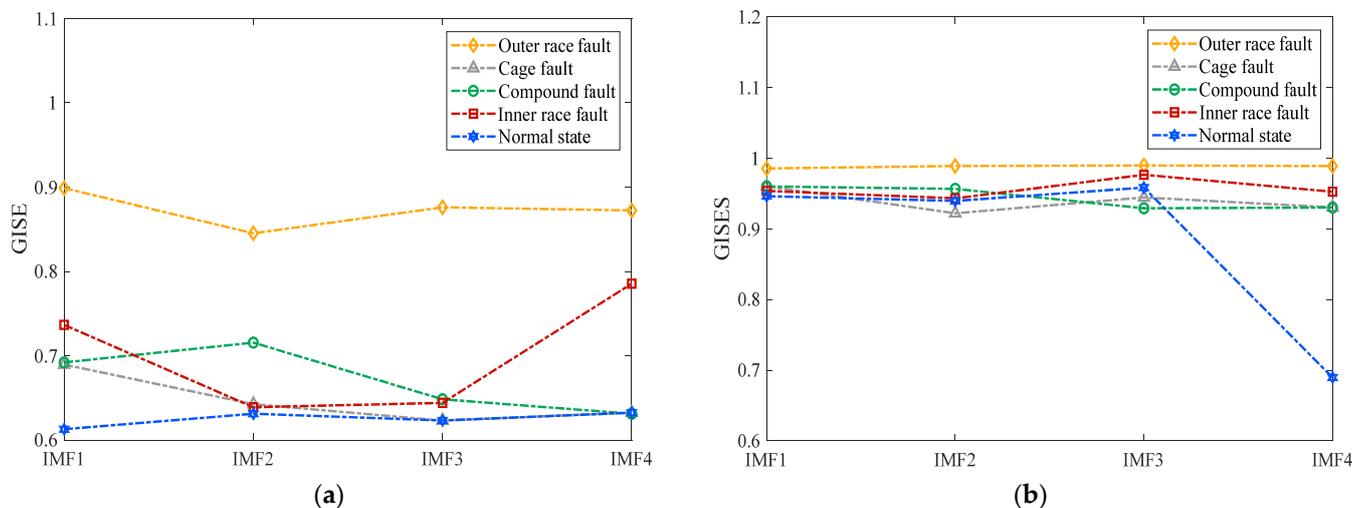


Figure 13. Gini index of selected components. (a) GISE; (b) GISES.

For each state, 20 feature samples are randomly selected for training and the remaining 30 are used for testing, and labels are attached to the corresponding sample data. More details of the selected bearing signals are described in Table 2. The LSTM is selected as the fitness function of AO, and the reconstruction error is taken as the fitness value. The learning rate and the number of hidden layer nodes of LSTM network are continuously iteratively optimized through AO algorithm, and the optimal parameters are output until the termination conditions are met.

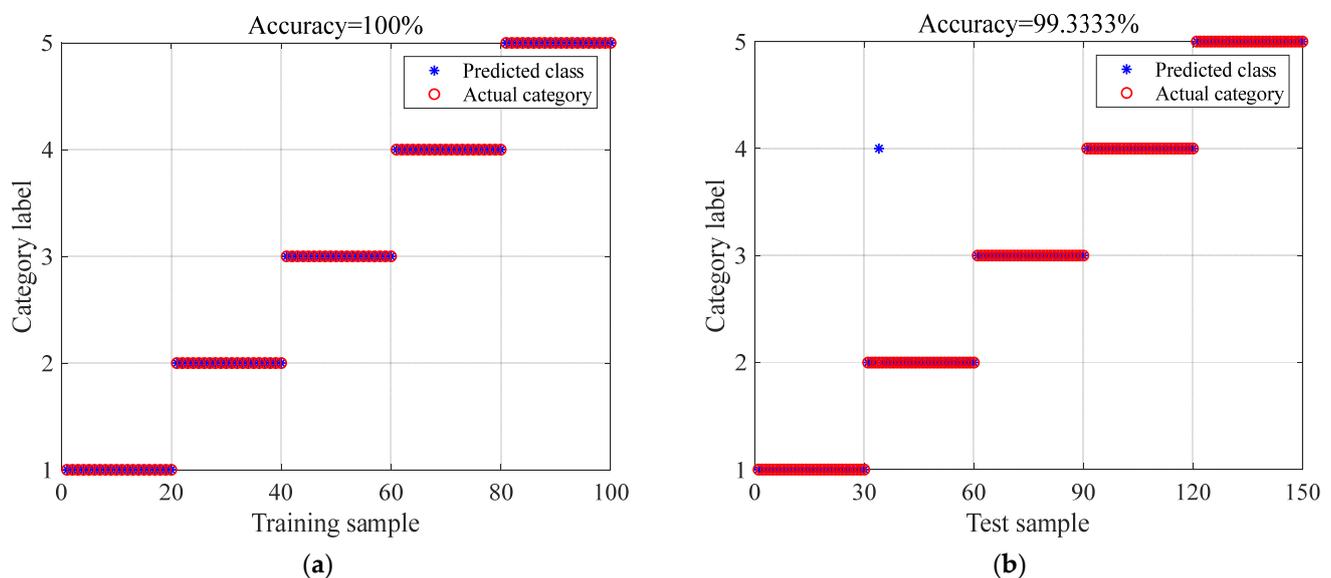
Table 2. Description of training and test samples under five bearing conditions.

Bearing Operating Conditions	Bearing Dataset	Operating Condition	Training Samples and Testing Samples	Condition Label
Outer race fault	Bearing 1_1	35 Hz/12 kN	100/150	1
Cage fault	Bearing 2_3	37.5 Hz/11 kN	100/150	2
Compound fault	Bearing 1_5	35 Hz/12 kN	100/150	3
Inner race fault	Bearing 2_1	37.5 Hz/11 kN	100/150	4
Healthy state	Bearing 2_1	37.5 Hz/11 kN	100/150	5

The main parameters of AO-LSTM model are shown in Table 3. The divided datasets are standardized, and then input into the model for training to generate a training network. The training network is used to classify and predict the experimental dataset, and the results are shown in Figure 14. The results show that all 100 training samples are classified correctly, and the training accuracy is 100%. Among the test samples, only one cage fault was misdiagnosed as an inner ring fault; 149 test samples were classified correctly, the test accuracy was 99.33%, and the compound fault diagnosis accuracy was 100%. This shows that the method based on the combination of ACMD and AO-LSTM can effectively identify a variety of fault types including compound faults.

Table 3. Description of training and test samples under five bearing conditions.

AO		LSTM	
Maximum number of iterations	10	The number of fully connected layer	5
Number of optimization parameters	2	The dimension of input sample	8
Population	5	The node number of hidden layers	10
Lower and upper limits of learning rate	[0.001, 0.1]	Learning Rate Drop Factor	0.1
Range of hidden layer node number	[5, 14]	Learning rate	0.01

**Figure 14.** Fault diagnosis results of the proposed method. (a) Training accuracy of fault diagnosis model. (b) Test accuracy of fault diagnosis model.

In order to understand the separability of extracted features more intuitively, the t-distributed stochastic neighbor embedding (t-SNE) method is used to visualize the feature information of each type of bearing by reducing the dimensionality of highly dimensional features. The feature information of each type of bearing extracted manually above is compared with the feature information obtained after classification of the AO-LSTM model. Figure 15 shows their two-dimensional visualization results, where t-SNE1 and t-SNE2 represent the first two main components of the five bearing status labels.

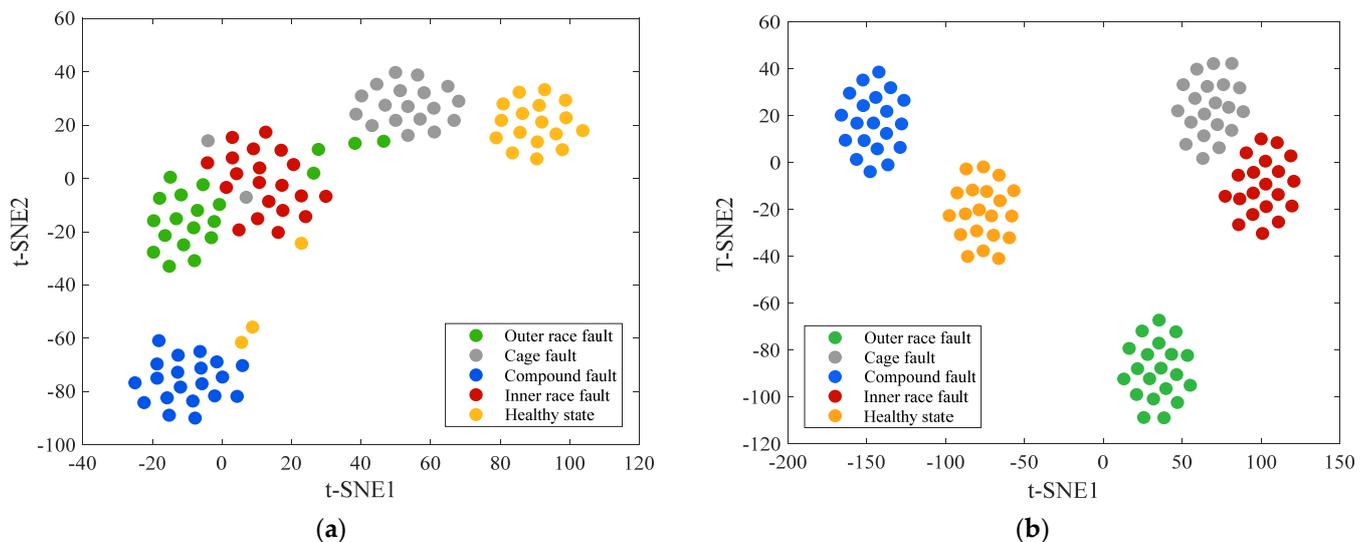


Figure 15. (a) Two-dimensional feature distribution after ACMD. (b) Two-dimensional feature distribution after ACMD and AO-LSTM.

It can be seen from the figure that in the feature information obtained by combining ACMD and the Gini coefficient, the feature distribution of the same fault type is wider, and there is feature overlap between the first and second fault types. However, after the training of the AO-LSTM model, the characteristics of these five kinds of bearing signals are clearly distinguished, and the feature aggregation of the same type is very compact. It further shows the effectiveness of the proposed method in rolling bearing compound fault diagnosis.

4.2.3. Diagnosis Results of the Proposed Method and Different Intelligent Methods

In order to verify the superiority of the AO algorithm, particle swarm optimization (PSO) and whale optimization algorithm (WOA) are used to test the LSTM network, respectively. The parameter setting of LSTM network is the same as that of AO algorithm. Figure 16 shows the comparison results of the test. It can be seen from the figure that only the AO algorithm can achieve the lowest reconstruction error by using fewer iterations, which is enough to illustrate the advantage of the AO algorithm in optimizing the hyperparameters of the LSTM network.

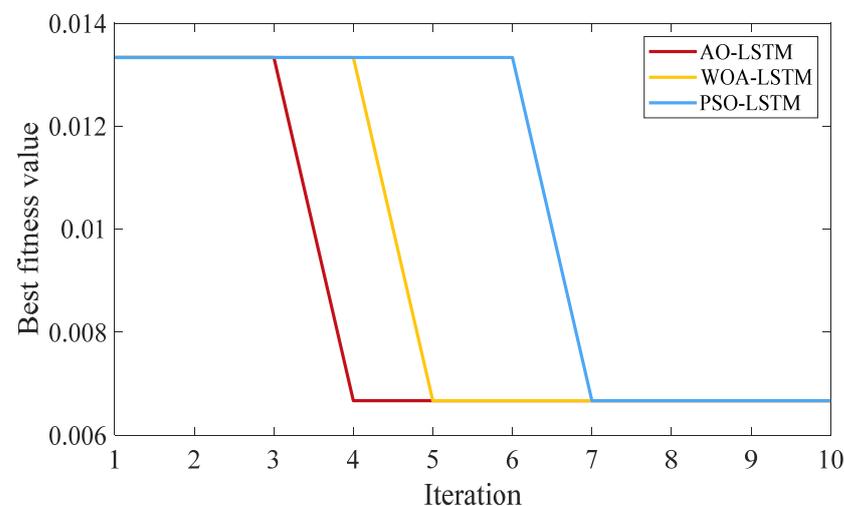


Figure 16. Convergence curves obtained by different optimization algorithms.

This paper uses the same experimental data, and the superiority of the proposed method is proved by comparing with five fault diagnosis methods. Firstly, ACMD is re-

placed by the continuous variational modal decomposition (SVMD) method, and the effects of the two decomposition methods on the classification results are compared. Secondly, the traditional characteristic index of the kurtosis coefficient is used to replace the Gini coefficient, and the classification accuracy is compared. Finally, different feature recognition methods are compared. The robustness and stability of the proposed method are verified ten times. Table 4 describes the classification results of different diagnostic methods.

Table 4. Description of classification results of various diagnostic models.

Diagnostic Model	Signal Decomposition Method	Characteristic Evaluation Index	Feature Recognition Method	Average Test Sample Accuracy (%)	Standard Deviation
The proposed model	ACMD	Gini index	AO + LSTM	98.67	0.53
Model 1	SVMD	Gini index	AO + LSTM	88.39	2.13
Model 2	ACMD	Kurtosis	AO + LSTM	85.73	2.15
Model 3	ACMD	Gini index	LSTM	91.87	1.67
Model 4	ACMD	Gini index	PSO + LSTM	94.53	1.06
Model 5	ACMD	Gini index	AO + KELM	95.67	0.94

Figure 17 shows each test result of the proposed method and other methods. In the diagnosis results, the average test accuracy is the standard to determine the fault diagnosis performance. The average diagnosis accuracy of the proposed method is 98.67%, which is significantly higher than other models. From the perspective of accuracy, it can be seen that the combination of ACMD and the Gini coefficient can extract feature information with high accuracy more effectively than the SVMD method and the kurtosis coefficient index.

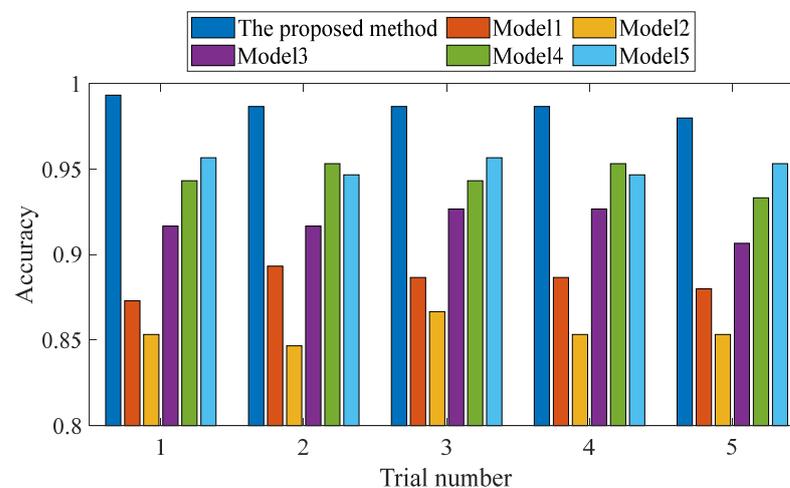


Figure 17. This is a figure. Schemes follow the same formatting.

5. Conclusions

In this paper, a bearing compound fault diagnosis model based on the combination of ACMD, Gini index fusion and AO-LSTM is proposed. It is applied to the compound fault simulation signal and the bearing experimental dataset of Xi'an Jiaotong University, respectively, to realize the intelligent identification of compound faults. Among a variety of bearing signals including compound fault, the classification accuracy of simulation signal and experimental signal test samples is 98.89% and 99.33%, respectively, of which the accuracy of compound fault diagnosis is 96.67% and 100%, respectively. Thus, the effectiveness and practicability of the proposed method are verified.

Through verification and comparative analysis, the reasons for the high accuracy of the proposed method can be summarized as follows:

1. The ACMD method uses a greedy algorithm to estimate the signal components one by one, and reduces the interference of irrelevant components to the greatest extent under the condition of ensuring that the useful feature information is not lost. In the proposed model, as the premise of feature extraction, this method can decompose the unstable vibration signal more stably and efficiently.
2. As a feature vector, the Gini coefficient is more robust to random impulse noise and other interference components because of its stable gradient characteristics and the ability to distinguish impulse and repetitive transients. It can effectively characterize the feature information and make the difference between samples more obvious.
3. The AO algorithm avoids the time-consuming and parameter uncertainty of manual parameter adjustment in optimizing the super parameters of LSTM model, ensures the accuracy and accuracy of the model, and can better monitor the running state of rolling bearing.

For the intelligent diagnosis of compound faults of rolling bearings, feature extraction can be further strengthened and improved in the future. More powerful feature learning models such as enhanced learning can be used to automatically extract the required compound fault features, realize completely unsupervised compound fault feature learning, and save feature extraction time while ensuring higher accuracy. In addition, more experimental data of compound fault of rolling bearing can be collected to analyze and verify the research method, so as to realize the universality and stability of the research method.

Author Contributions: Conceptualization, X.W.; methodology, X.W.; software, X.W.; validation, X.W.; formal analysis, J.M.; investigation, J.M.; resources, J.M.; data curation, J.M.; writing—original draft preparation, X.W.; writing—review and editing, J.M. and X.W.; 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 R&D program 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 in XJTU-SY Bearing dataset at <http://biaowang.tech/xjtu-sy-bearing-datasets>, accessed on 15 October 2021, reference number [30].

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Liu, Y.; Yan, X.S.; Zhang, C.N.; Wen, L. An ensemble convolutional neural networks for bearing fault diagnosis using multi-sensor data. *Sensors* **2019**, *19*, 5300. [[CrossRef](#)]
2. Ma, J.P.; Zhuo, S.; Li, C.W.; Zhan, L.W.; Zhang, G.Z. An enhanced intrinsic time-scale decomposition method based on adaptive lévy noise and its application in bearing fault diagnosis. *Symmetry* **2021**, *13*, 617. [[CrossRef](#)]
3. Zhang, K.; Zhou, D.H.; Cai, Y. Review of multiple fault diagnosis methods. *Control Theory Appl.* **2015**, *32*, 1143–1157.
4. Zheng, S.; Liu, T.; Liu, C.; Li, H. Compound faults diagnosis method of rolling bearing based on sparse representation of cascaded over complete dictionary. *J. Vib. Shock* **2021**, *40*, 174–179.
5. Wang, M.Y.; Wang, H.Q.; Dong, F.; Ren, B.Y.; Song, L.Y. A method of compound fault signal separation based on EVMD-LNMF. *J. Vib. Shock* **2019**, *38*, 146–152.
6. Qi, Y.S.; Fan, J.; Li, Y.T.; Gao, X.J.; Liu, L.Q. An improved deconvolution algorithm and its application in compound fault diagnosis of rolling bearing. *J. Vibration. Shock* **2020**, *39*, 140–150.
7. Cui, L.L.; Gao, L.X.; Yin, H.C.; Xu, Y.G. Research on composite fault diagnosis method based on the second generation wavelet. *China Mech. Eng.* **2009**, *20*, 442–446.
8. Chen, H.Z.; Wang, J.X.; Tang, B.P.; Li, J.C. Helicopter rolling bearing hybrid faults diagnosis using minimum entropy deconvolution and Teager energy operator. *J. Vib. Shock* **2017**, *36*, 45–50.

9. Qi, Y.S.; Liu, F.; Li, Y.T.; Gao, X.J.; Liu, L.Q. Compound fault diagnosis of wind turbine rolling bearing based on MK-MOMEDA and Teager energy operator. *Acta Energ. Sol.* **2021**, *42*, 297–307.
10. Wan, S.T.; Zhang, X.; Dou, L.J. Separation of composite rolling bearings fault features with strong noise interference. *J. Cent. South Univ.* **2018**, *49*, 1950–1959.
11. Chen, S.Q.; Yang, Y.; Peng, Z.K.; Wang, S.B.; Zhang, W.M.; Chen, X.F. Detection of rub-impact fault for rotor-stator systems: A novel method based on adaptive chirp mode decomposition. *J. Sound Vib.* **2018**, *440*, 83–89. [[CrossRef](#)]
12. Zhao, M.; Lin, J. Improvement of kurtosis-guided-grams via Gini index for bearing fault feature identification. *IEEE Trans. Ind. Electr.* **2017**, *65*, 2548–2556. [[CrossRef](#)]
13. Miao, Y.; Zhao, M.; Lin, J. Health assessment of rotating machinery using a rotary encoder. *Meas. Sci. Technol.* **2017**, *28*, 125001. [[CrossRef](#)]
14. Albezzawy, M.N.; Nassef, M.G.; Sawalhi, N. Rolling element bearing fault identification using a novel three-step adaptive and automated filtration scheme based on Gini index. *ISA Trans.* **2020**, *101*, 453–460. [[CrossRef](#)]
15. Nassef, M.G.A.; Hussein, T.M.; Mokhiamar, O. An adaptive variational mode decomposition based on sailfish optimization algorithm and Gini index for fault identification in rolling bearings. *Measurement* **2021**, *173*, 108514. [[CrossRef](#)]
16. Qu, J.L.; Yu, L.; Yuan, T.; Tian, Y.P.; Gao, F. Adaptive fault diagnosis algorithm for rolling bearings based on one-dimensional convolutional neural network. *Chin. J. Sci. Instrum.* **2018**, *39*, 134–143.
17. Chen, W.; Chen, J.X.; Jiang, Y.Q.; Song, D.L.; Zhang, W.D. Fault identification of rolling bearing based on RS-LSTM. *China Sci. Paper* **2018**, *13*, 1134–1141.
18. Chen, Y.Q. Intelligence Diagnosis System Research of Rolling Bearing Composite Faults. Master's Thesis, Yanshan University, Qinhuangdao, China, 2013.
19. Zhang, Y.Y.; Jia, Y.X.; Wu, W.Y.; Cheng, Z.H.; Su, X.B.; Lin, A.Q. A diagnosis method for the compound fault of gearboxes based on multi-feature and BP-AdaBoost. *Symmetry* **2020**, *12*, 461. [[CrossRef](#)]
20. Han, T.; Yuan, J.H.; Tang, J.; An, L.Z. An approach of intelligent compound fault diagnosis of rolling bearing based on MWT and CNN. *J. Mech. Transm.* **2016**, *40*, 139–143.
21. Shi, J.; Wu, X.; Liu, T. Bearing compound fault diagnosis based on HHT algorithm and convolution neural network. *Trans. Chin. Soc. Agric. Eng.* **2020**, *36*, 34–43.
22. Yu, P.; Cao, J. Deep learning approach and its application in fault diagnosis and prognosis. *Comput. Eng. Appl.* **2020**, *56*, 1–18.
23. Cao, Z.Z.; Ye, C.M. Application of improved CNN-LSTM model in fault diagnosis of rolling bearings. *Comput. Syst. App.* **2021**, *30*, 126–133.
24. Zhang, Q.; Jang, W.Z.; Li, H. Combined MCKD-Teager energy operator with LSTM for rolling bearing fault diagnosis. *J. Harbin Inst. Technol.* **2021**, *53*, 68–76.
25. Abualigah, L.; Yousri, D.; Abd Elaziz, M.; Ewees, A.A.; Al-qaness, M.A.; Gandomi, A.H. Aquila Optimizer: A novel meta-heuristic optimization algorithm. *Comput. Ind. Eng.* **2021**, *157*, 107250. [[CrossRef](#)]
26. Miao, Y.H.; Wang, J.J.; Zhang, B.Y.; Li, H. Practical framework of Gini index in the application of machinery fault feature extraction. *Mech. Syst. Signal Process.* **2022**, *165*, 108333. [[CrossRef](#)]
27. Shi, H.T.; Guo, L.; Tan, S.; Bai, X.T.; Sun, J. Rolling bearing initial fault detection using long short-term memory recurrent network. *IEEE Access* **2019**, *7*, 171559–171569. [[CrossRef](#)]
28. Zhang, J.F.; Song, Y.; Li, G.; Wang, C.Y.; Jiao, Y.F. A method fault diagnosis of rolling bearing of wind turbines based on long short-term memory neural network. *Comput. Meas. Control* **2017**, *25*, 16–19.
29. Wang, B.; Lei, Y.G.; Li, N.P.; Li, N.B. A hybrid prognostics approach for estimating remaining useful life of rolling element bearings. *IEEE Trans. Reliab.* **2020**, *69*, 401–412. [[CrossRef](#)]
30. Lei, Y.G.; Han, T.Y.; Wang, B. XJTU-SY rolling element bearing accelerated life test datasets: A tutorial. *J. Mech. Eng.* **2019**, *55*, 1–6.