

Article

A Diagnosis Method for the Compound Fault of Gearboxes Based on Multi-Feature and BP-AdaBoost

Yangyang Zhang ¹, Yunxian Jia ^{1,*}, Weiyi Wu ¹, Zhonghua Cheng ¹, Xiaobo Su ¹
and Aqiang Lin ² 

¹ Shijiazhuang Campus, Army Engineering University, Shijiazhuang 050003, China; pga_zhang@163.com (Y.Z.); weiyi1019@163.com (W.W.); 13143648622@163.com (Z.C.); giantsu030700@sina.com (X.S.)

² College of Power and Energy Engineering, Harbin Engineering University, Harbin 150001, China; linaqiang@hrbeu.edu.cn

* Correspondence: yunxian_jia@126.com

Received: 11 February 2020; Accepted: 3 March 2020; Published: 14 March 2020



Abstract: Gearbox is an important structure of rotating machinery, and the accurate fault diagnosis of gearboxes is of great significance for ensuring efficient and safe operation of rotating machinery. Aiming at the problem that there is little common compound fault data of gearboxes, and there is a lack of an effective diagnosis method, a gearbox fault simulation experiment platform is set up, and a diagnosis method for the compound fault of gearboxes based on multi-feature and BP-AdaBoost is proposed. Firstly, the vibration signals of six typical states of gearbox are obtained, and the original signals are decomposed by empirical mode decomposition and reconstruct the new signal to achieve the purpose of noise reduction. Then, perform the time domain analysis and wavelet packet analysis on the reconstructed signal, extract three time domain feature parameters with higher sensitivity, and combine them with eight frequency band energy feature parameters obtained by wavelet packet decomposition to form the gearbox state feature vector. Finally, AdaBoost algorithm and BP neural network are used to build the BP-AdaBoost strong classifier model, and feature vectors are input into the model for training and verification. The results show that the proposed method can effectively identify the gearbox failure modes, and has higher accuracy than the traditional fault diagnosis methods, and has certain reference significance and engineering application value.

Keywords: gearboxes; compound fault diagnosis; multi-feature; BP-AdaBoost

1. Introduction

Energy consumption [1], energy saving [2,3], fault diagnosis [4], and their problem-solving method is getting more and more attention for rotating machinery [5–7]. As a key transmission component for adjusting speed and transmitting power, gearbox is an important component of rotating machinery [8], which is widely used in industrial, civil and military fields, such as lathe, wind turbine, automobile, construction machinery, helicopter, etc. [9–14]. However, the gearbox is one of the mechanical devices with a high failure rate because of its complex structure and long-term operation in the complex and harsh environment such as high speed and heavy load [15]. Once a fault occurs, it will lead to unplanned downtime, huge economic losses, and even serious disaster consequences [16–19]. Therefore, it is very important to monitor the health of the gearbox and find out the faults as early as possible so as to plan the shutdown and maintenance properly, so as to ensure the safe operation of the mechanical equipment, improve the production efficiency, and increase the economic benefits [20,21].

At present, many scholars have been devoted to the research of gearbox fault diagnosis methods, and have achieved certain results [22–25]. These fault diagnosis methods can be divided into two

types: model-based and data-driven [26,27]. The model-based fault diagnosis method needs to use the relevant professional knowledge of a mechanical failure mechanism to build a physical model. Although the diagnosis accuracy is relatively high, the modeling is difficult and the model does not have universality [28]. However, the data-driven fault diagnosis method does not depend on the knowledge of the professional field, the algorithm is relatively simple and has good universality, and has been widely used in gearbox fault diagnosis [29,30]. Among them, because the vibration sensor is cheap and easy to install and use, the vibration signal acquisition process is simple and the signal contains rich state information, so the data-driven fault diagnosis method based on the vibration signal has attracted more researchers' attention [31,32].

With the rapid development of artificial intelligence, the emergence of artificial neural network provides a new idea for fault diagnosis. It is a new trend for the development of fault diagnosis technology of rotating machinery to use an artificial neural network method, machine learning algorithm, and other intelligent methods for fault pattern recognition [33]. Zhang et al. [34] proposed a transfer learning method for fault diagnosis using artificial neural network, which can realize fault diagnosis under different working conditions. He et al. [35] presented an approach that preprocesses sensor signals using short-time Fourier transform (STFT). Based on a simple spectrum matrix obtained by STFT, an optimized deep learning structure, and a large memory storage retrieval (LAMSTAR) neural network is built to diagnose the bearing faults. Baraldi et al. [36] developed a diagnosis system based on K-nearest neighbours classifiers, which is used to detect the degradation and fault classification of bearings. Wu et al. [37] realized the diagnosis of bearing faults in rotating machinery by selecting some features as the inputs of a support vector machine (SVM) classifier. In addition, a convolutional neural network is widely used in fault diagnosis [38,39].

Gearboxes usually include gears, shafts, bearings, and other components with symmetry. Due to the influence of manufacturing processes, lubrication conditions and long-term direct impact of variable loads, the failure proportion of gear in various parts of a gearbox is the largest [40,41]. Single gear may suffer from pitting, wear, and tooth breakage [42,43]. However, there may be multiple gears that fail at the same time in the gearbox under actual circumstances, which is the compound fault of gearbox mentioned in this paper. Although the above data-driven fault diagnosis method has made some progress and effect, the diagnosis accuracy of the gearbox compound fault is not very high, and there is still some room for improvement in the following aspects:

1. The vibration signal of the gearbox is a non-stationary, nonlinear signal and contains a certain amount of noise [44]. In the process of signal processing, the traditional signal processing methods such as Fourier transform and its improved algorithms, short-time Fourier transform, and wavelet transform have some limitations in processing the vibration signals of the gearbox.
2. Traditional fault diagnosis methods rely on expert knowledge and experience to select feature parameters, and mostly extract feature parameters from a single time domain, frequency-domain, or time-frequency-domain analysis. Although some research reveals some fault features, there is still not enough to distinguish different fault modes accurately.
3. At present, there is little research on the compound fault diagnosis of gearbox, but, in actual situations, there may be different faults of multiple gears. Although some researchers have made some progress in the diagnosis method of compound faults of gearboxes, the diagnosis accuracy is not high enough, and the algorithm is relatively complex.

In the context of the above problems and research, in order to use the vibration signal of gearboxes to diagnose the compound fault of gearboxes more effectively, this paper proposes a precise diagnosis method of the compound fault of gearboxes based on multi-feature and BP-AdaBoost. Firstly, empirical mode decomposition (EMD) is introduced to deal with the non-stationary and nonlinear signals in gearbox faults, so as to achieve the purpose of noise reduction; then, time domain analysis and wavelet packet analysis are carried out for the new fault signals after noise reduction processing. In time domain analysis, the sensitivity analysis method can be used to screen out the time domain feature

parameters with high sensitivity, and the band energy feature vectors can be obtained through wavelet packet analysis. The two together constitute the state feature vectors that can effectively reflect the running state of the gearbox; finally, AdaBoost algorithm is introduced into the BP neural network model, and the state feature vectors are input into the constructed BP-AdaBoost fault diagnosis model, and finally the precise diagnosis of compound fault of gearbox is realized.

The flow of the proposed method is shown in Figure 1, which includes the following steps:

1. Collect vibration signals of the gearbox under different operating conditions, decompose the original vibration signal of gearbox by EMD, and select the appropriate intrinsic mode functions (IMFs) to reconstruct the new signal;
2. First perform time domain analysis on the reconstructed signal, calculate time domain statistical parameters, and select sensitive time domain feature parameters through sensitivity analysis;

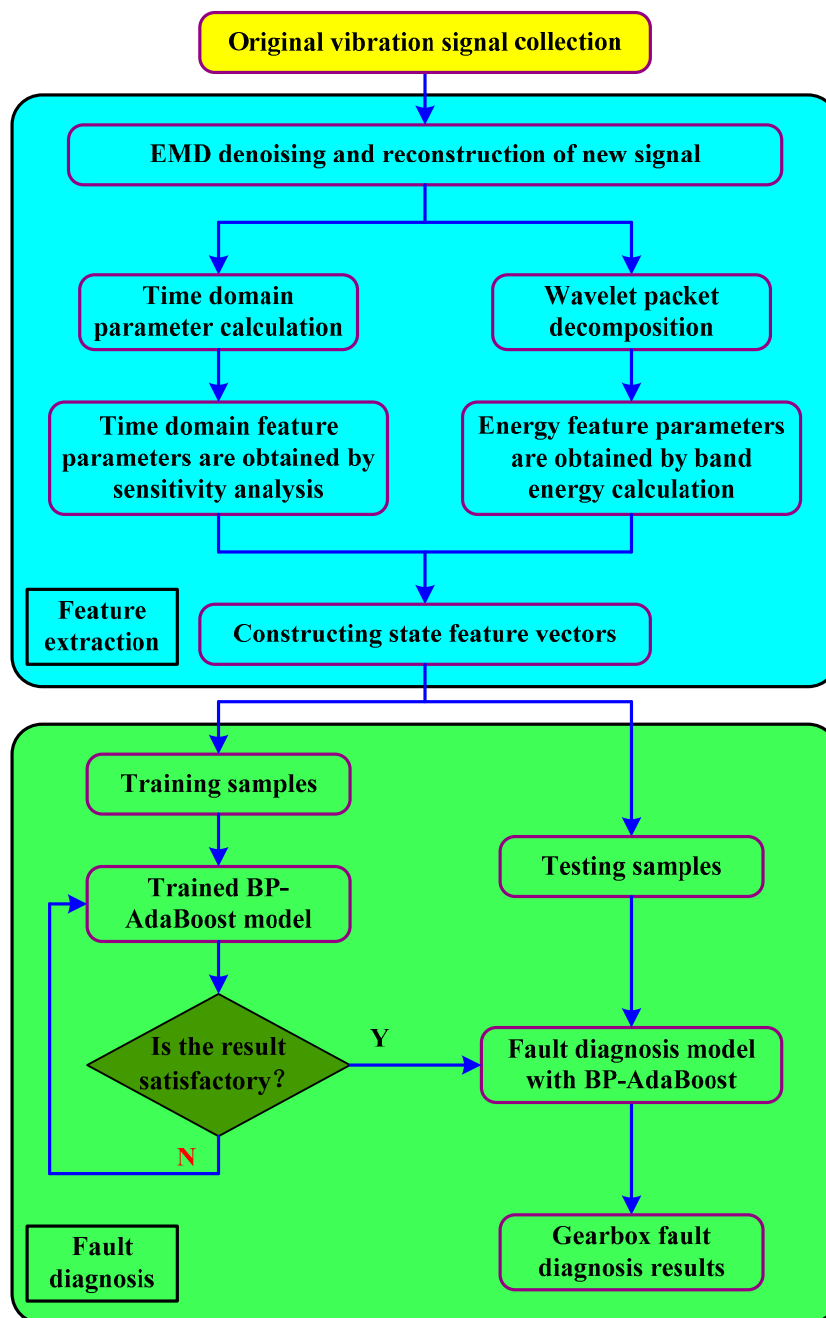


Figure 1. The flow of the proposed method.

3. Then, perform wavelet packet analysis on the reconstructed signal, calculate the waveband energy of each frequency band by wavelet packet decomposition, construct band energy feature vectors, and finally form gearbox state feature vectors with the selected time domain feature parameters;
4. Establish the BP-AdaBoost model, select training samples and test samples, and use the training samples to train the model;
5. Input the test samples into the trained BP-AdaBoost model to obtain the fault diagnosis results.

The main contributions of the work are summarized as follows:

1. By combining BP-AdaBoost with sensitivity analysis and wavelet packet analysis, a novel intelligent fault diagnosis method for the compound fault of gearboxes is proposed, which can effectively diagnose the compound fault of gearboxes.
2. In order to improve the accuracy of fault diagnosis, we improve the traditional fault diagnosis method in three aspects. Firstly, EMD decomposition and reconstruction are used to denoise the signal, then sensitivity analysis is used to select the time domain feature parameters with high sensitivity and combine them with wavelet packet energy feature parameters to form the state feature vector. Finally, BP-AdaBoost is used to build the fault diagnosis model.
3. The proposed method has certain reference significance and engineering application value for the compound fault of gearboxes as well as other types of faults and machinery.

The rest of this paper is organized as follows. Section 2 briefly introduces the basic theory of proposed method including EMD, time domain analysis, wavelet packet analysis, and BP-AdaBoost algorithm. In Section 3, the effectiveness of the proposed method is verified by experiments. Comparative experiments to verify the advantages of the proposed method are presented in Section 4. Finally, conclusions are drawn in Section 5.

2. Basic Principle of the Proposed Method

2.1. EMD

The signal of gearbox fault is generally a non-stationary signal, so it is helpful to improve the accuracy of fault diagnosis to adopt the appropriate method to deal with a non-stationary signal. In the process of traditional wavelet analysis, once the wavelet function is selected, it can not be changed according to the actual situation. From this point of view, wavelet transform is not adaptive. Empirical mode decomposition (EMD) is completely adaptive, which does not need to determine the decomposition basis in advance, and is suitable for processing non-stationary and nonlinear signals [45]. Therefore, the EMD method is used to decompose and reconstruct the original signal to achieve the purpose of denoising.

The EMD method was proposed by Norden E. Huang in 1998. Now, it has been widely used in mechanical fault diagnosis, earthquake research, and medical analysis [46]. The basic idea of EMD is to decompose the complex signals into many simple ones for the convenience of analysis and research. Its essence is to obtain the intrinsic mode function through the characteristic time scale of data, and then decompose the data. The specific steps of EMD are as follows.

Step 1: determine all local extremum points of the original signal $x(t)$, connect all maximum points and all minimum points with cubic spline curves respectively, and get the upper and lower envelope lines of $x(t)$, so that all data points of the signal are between the two envelope lines.

Step 2: calculate the mean value $m(t)$ of the upper and lower envelope, and subtract $m(t)$ from the original signal $x(t)$ to get

$$h_1(t) = x(t) - m(t). \quad (1)$$

Check whether $h_1(t)$ satisfies the conditions of intrinsic mode function. If not, repeat the above steps with $h_1(t)$ replacement $x(t)$ until it is an intrinsic mode function, and record it as the first IMF of the original signal.

Step 3: after decomposing the first intrinsic mode function $c_1(t)$ from the original signal $x(t)$, subtract $c_1(t)$ from $x(t)$ to get the residual value sequence $r_1(t)$

$$r_1(t) = x(t) - c_1(t) \quad (2)$$

Step 4: take $r_1(t)$ as the new signal to be processed, repeat the above steps, get n IMFs of the original signal $x(t)$ in turn, record it as $c_1(t), c_2(t), \dots, c_n(t)$, and then leave the remainder $r_n(t)$ of the original signal.

In this way, the original signal $x(t)$ is decomposed into the sum of several intrinsic mode functions and a remainder:

$$x(t) = \sum_{i=1}^n c_i(t) + r_n(t) \quad (3)$$

In order to make the final signal contain the most useful information and the least noise, n IMFs need to be selectively utilized. Add the useful IMFs to get the final reconstruction signal $y(t)$ as the signal to be used, which can be expressed as

$$y(t) = \sum_{i=1}^c c_c(t) \quad (4)$$

2.2. Time Domain Analysis

The time domain diagram of vibration signal records the trend of signal size changing with time in the operation of mechanical system. The time domain signal contains a large amount of information, which is intuitive and easy to understand. It is the original basis of mechanical fault diagnosis. However, the vibration signals collected by the gearbox test-bed are usually non-stationary and non-periodic signals. It is difficult to see the difference between them and normal signals directly from the time domain waveform. Therefore, it is necessary to carry out time domain statistical analysis on the signals, calculate or estimate various time domain parameters and indexes of dynamic signals. By studying and selecting the appropriate time domain feature parameters, we can make accurate judgment for different types of faults.

The feature parameter extraction based on a time domain index refers to the use of the statistical parameters of vibration signals in time domain to characterize the gearbox state information. The advantage of time domain parameters in the field of fault identification is that they can directly describe the gearbox state. Common time domain statistical parameters include: mean (P1), root mean square (P2), square root amplitude (P3), absolute mean amplitude (P4), mean square (P5), maximum (P6), minimum (P7), peak-to-peak (P8), waveform index (P9), peak index (P10), pulse index (P11), margin index (P12), skewness index (P13), and kurtosis index (P14).

In practical application, all kinds of feature indexes have different emphasis on the expression of a fault. Therefore, the above indexes can be used to describe the healthy state of gearbox, which can complement each other and take sensitivity and stability into account.

In the process of time domain analysis, multiple time domain feature parameters will be extracted. The sensitivity of the selected feature parameters to the fault determines the difficulty and accuracy of fault diagnosis. Therefore, it is very important to select the appropriate feature parameters to form the fault feature vector [47]. Therefore, it is necessary to analyze the sensitivity of these parameters after extracting the time domain feature parameters, and screen out the feature parameters which can effectively represent the state of gearbox. The calculation formula of sensitivity analysis is

$$u = \left| \frac{X_i - X}{X} \right| \quad (5)$$

where u is the sensitivity, X_i is the parameter to be analyzed, and X is the reference parameter. Usually, the time domain feature parameters of normal state are taken as reference parameters, and the time

domain feature parameters of fault state are taken as parameters to be analyzed. Note that sensitivity analysis is a normalization process although it is used as a means for parameter selection.

2.3. Wavelet Packet Analysis

Wavelet transform is one of the most popular time–frequency transform methods that has the ability to characterize signal characteristics in both time and frequency domain. At present, wavelet transform is widely used in the field of rotating machinery fault diagnosis [48]. However, the decomposition method adopted by wavelet decomposition has poor time resolution in the low frequency band and poor frequency resolution in the high frequency band. In order to solve this problem, wavelet packet analysis provides the way to solve the problem; it is the extension of wavelet analysis and is a more precise method of signal analysis. It not only inherits the advantages of time-frequency localization of wavelet transform, but also further decomposes the high frequency band that are not re-decomposed by the wavelet transform, so as to improve the frequency resolution. The comparison figure of frequency band division between wavelet decomposition and wavelet packet decomposition is shown in Figure 2. Obviously, wavelet packet decomposition has symmetrical characteristics. S represents the low frequency part of a vibration signal and D represents the high frequency part of a signal.

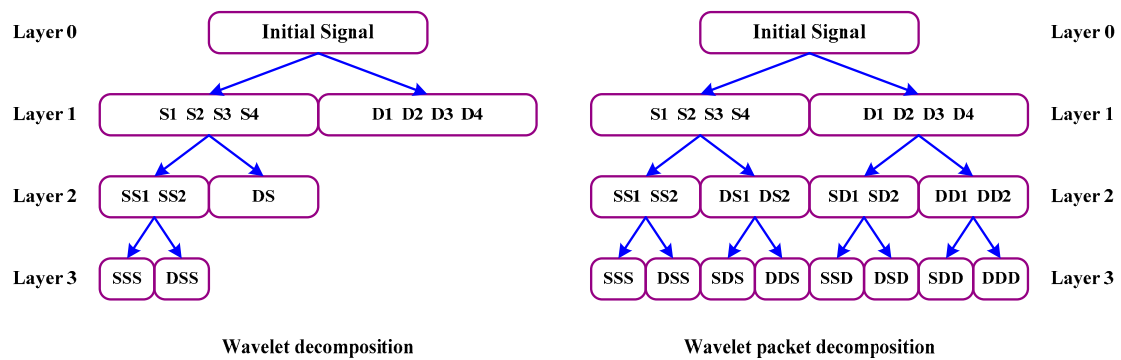


Figure 2. The comparison figure of frequency band division between wavelet decomposition and wavelet packet decomposition.

For the key components such as gearbox, wavelet packet analysis can fully capture the useful information in the vibration signal, which is conducive to accurate diagnosis, so wavelet packet decomposition is more reliable and feasible. The principle of wavelet packet decomposition is as follows.

If the original signal collected is $x_l^n(t) \in U_l^n$ (subspace formed by orthogonal sum of wavelet subspace and scale subspace), then $x_l^n(t)$ can be represented by wavelet packet decomposition:

$$x_l^n(t) = \sum_l d_l^{j,n} u_n(2^j t - l) \tag{6}$$

where j is scale, l is translation factor, $u_n(2^j t - l)$ is basis function in subspace U_l^n , $d_l^{j,n}$ is coefficient of wavelet packet decomposition, and $\{d_l^{j+1,n}\}$ can be decomposed into $\{d_l^{j,2n}\}$ and $\{d_l^{j,2n+1}\}$ by a formula, that is,

$$\begin{aligned} \{d_l^{j,2n}\} &= \sum_k (a_k - 2) d_k^{j+1,n} \\ \{d_l^{j,2n+1}\} &= \sum_k (b_k - 2l) d_k^{j+1,n} \end{aligned} \tag{7}$$

where $a_k - 2$ is the scale in the decomposition process and $b_k - 2l$ is the wavelet coefficient.

After the vibration signal of gearbox is decomposed by a frequency band in the above way, each node will arrange the final result according to the frequency band, but it is difficult to judge the operation state of gearbox intuitively through the sub-band information of wavelet

packet decomposition, and the energy value of the vibration signal in different frequency bands is different, which can effectively extract the information contained in each frequency band of the signal. Therefore, the failure mode of the gearbox can be identified by calculating the energy of each frequency band of the signal.

Suppose that the highest frequency of the vibration signal $x(t)$ is f_s . By performing wavelet packet decomposition of J -layer on $x(t)$, we can get 2^J groups of wavelet packet coefficients, which are $\omega_i^J, i = 0, 1, \dots, 2^J - 1$, respectively. The corresponding frequency band of 2^J groups of coefficients is $\left[\frac{i}{2^J}f_s, \frac{i+1}{2^J}f_s\right]$.

According to Parseval's theorem, the energy calculated in a time domain is consistent with that calculated in a frequency domain, so the energy calculation formula of each frequency band is

$$E_i^J = \|\omega_i^J\|^2 = \int_{-\infty}^{+\infty} |\omega_i^J(t)|^2 dt \quad (8)$$

Taking E_i^J as the element can form an energy feature vector such as $T = [E_0^J, E_1^J, \dots, E_{2^J-1}^J]$. In practical applications, the energy of each node is usually normalized, that is, the proportion of the energy of each node to the total energy is taken. Let $E = \sum_i E_i^J$, and the energy feature vector can be normalized as:

$$R = \frac{1}{E} [E_0^J, E_1^J, \dots, E_{2^J-1}^J] \quad (9)$$

2.4. BP-AdaBoost Algorithm

A back propagation neural network (BP neural network, BPNN) has excellent nonlinear approximation characteristics, and has been widely used in many fields, but there are also some defects, such as being easy to fall into local minima and not being able to get the global optimal solution, resulting in a low recognition rate of mechanical equipment failure mode. The AdaBoost algorithm is introduced into a BP neural network model to build a BP-AdaBoost fault diagnosis model, which can effectively improve the diagnosis accuracy of the original single BP neural network.

The AdaBoost (Adaptive Boosting) lifting method is one of the most classical methods in the boosting algorithm family. It reduces the error rate through the adaptive learning method [49]. The main idea of the algorithm is: first, each sample is initialized to equal weight, and then iterates n times with a weak classifier. After each iteration, the weight is updated according to the classification results. For the failed samples, a larger weight is given, and more attention is paid to the next iteration. Weak classifiers get a sequence of classification functions through multiple iterations. Each classification function will give a certain weight according to the classification error, and finally combine multiple weak classifiers to form a strong and stable classifier with small classification errors. The algorithm flow of BP-AdaBoost is shown in Figure 3, and its main steps are as follows.

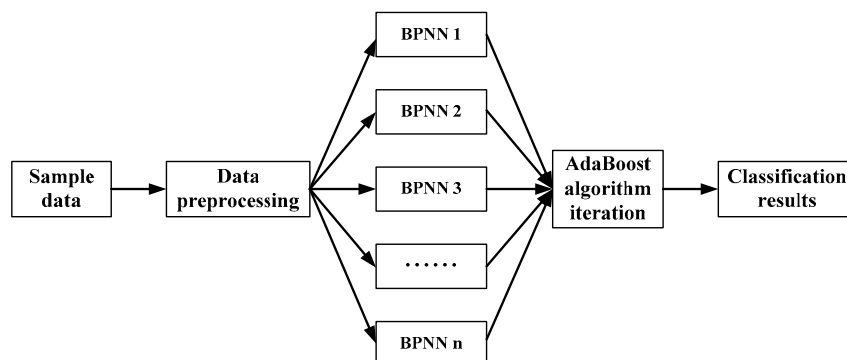


Figure 3. The algorithm flow of BP-AdaBoost.

Step 1: data selection and network initialization. Select m group training samples from the sample space and initialize the sample weight $D_t(i) = 1/m$. According to the dimensions of sample input and output, the structure of neural network is determined, and the weight and threshold of BP neural network are initialized.

Step 2: weak classifier prediction. When training the t -th weak classifier, the BP neural network is trained with the training data, and the output of the training result is predicted. The prediction sequence $g(t)$ and the prediction error sum e_t are obtained:

$$e_t = \sum_i D_t(i) \quad i = 1, 2, \dots, m (g(t) \neq y) \quad (10)$$

where $g(t)$ and y are predicted classification results and expected classification results, respectively.

Step 3: calculate the weight a_t of the prediction sequence according to the prediction error e_t . The formula is

$$a_t = \frac{1}{2} \ln\left(\frac{1 - e_t}{e_t}\right) \quad (11)$$

Step 4: update the training sample weight. Adjust the weight of training samples in the next round of training according to the prediction sequence weight a_t , and the adjustment formula is

$$D_{t+1}(i) = \frac{D_t(i)}{B_t} \cdot \exp[-a_t y_i g_t(x_i)] \quad i = 1, 2, \dots, m \quad (12)$$

where B_t is the normalization factor, the purpose is to keep the weight proportion unchanged and the sum of the weights equal to 1.

Step 5: synthesize a strong classifier. After training T rounds, we get T groups weak classification function $f(g_t, a_t)$, and combine them to get a strong classification function:

$$h(x) = \text{sign}\left[\sum_{t=1}^T a_t f(g_t, a_t)\right] \quad (13)$$

3. Experimental Verification and Analysis

In order to verify the effectiveness of the proposed method, the gearbox preset fault experiment is specially set for data verification. A QPZZ-II rotating machinery vibration analysis and fault diagnosis experimental platform system is adopted in the experiment, and the structural diagram of the experimental platform is shown in Figure 4. The gearbox used in the experiment is mainly composed of a pair of meshing gears. The big gear is the driving wheel, the number of teeth is 75, the small gear is the driven wheel, the number of teeth is 55, and the module is 2. The speed range of the motor is 75–1450 rpm, and the maximum workload is 5 N·m. The acquisition software of the system can realize the signal acquisition and storage.

According to the analysis of common failure modes of gearbox, this experiment simulates six operation states of gearbox, and the specific experimental settings are shown in Table 1.

Table 1. Gearbox preset fault experiment.

Mode	Preset Fault	Big Gear	Small Gear	Speed	Workload
1	None (Healthy)	Healthy	Healthy		
2	Tooth pitting	Tooth pitting	Healthy		
3	Tooth wear	Healthy	Tooth wear	880 r/min	0.2 A
4	Broken tooth	Broken tooth	Healthy		
5	Tooth pitting & wear	Tooth pitting	Tooth wear		
6	Broken tooth & wear	Broken tooth	Tooth wear		

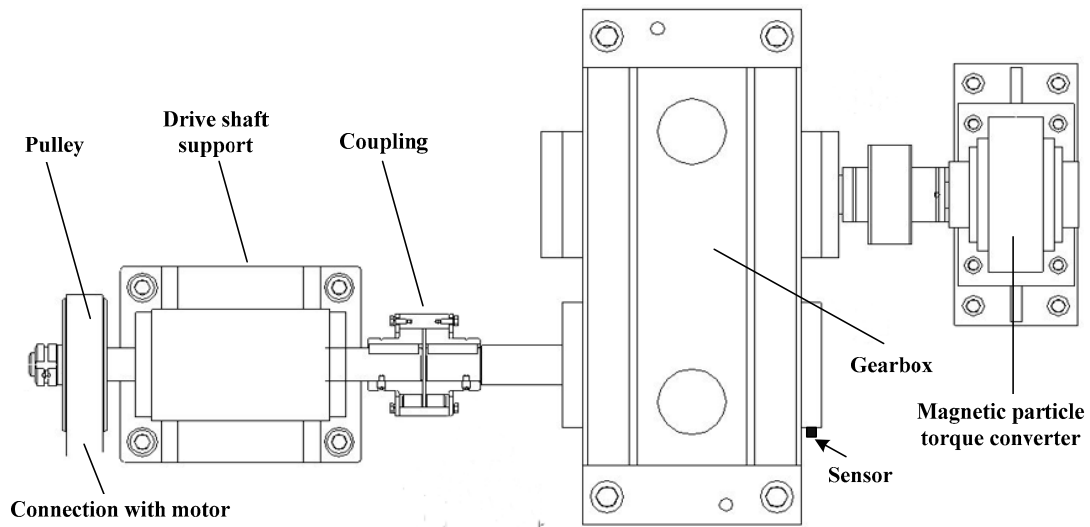


Figure 4. Structural diagram of the gearbox experimental platform.

During the experiment, the acceleration monitoring data in the y -direction on the load side of a gearbox input shaft is collected by an acceleration sensor, and the sampling frequency is 5120 Hz. The signal with a length of 5120 points each time is taken as a sample, so the time length of this signal is 1 s, and the sample contains the state information of the gearbox within 1 s, and the rotational speed of the gearbox is 880 rpm, so this signal can effectively contain all the information of the gearbox within one operation cycle.

3.1. Denoising with EMD

The original signal collected is decomposed by the EMD method, and the decomposition result is shown in Figure 5. In the figure, 12 IMFs of the signal are successively from top to bottom.

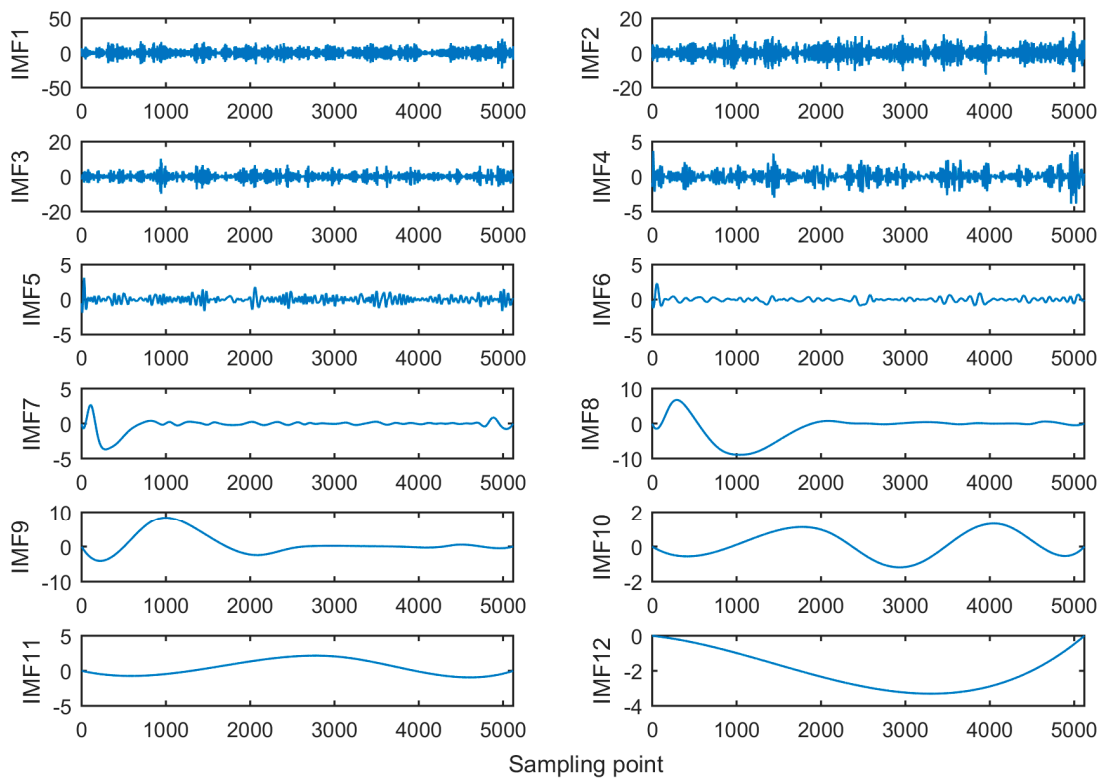


Figure 5. EMD decomposition results.

From Ref. [45], if the higher the ratio of energy of IMF obtained by EMD to that of an original signal, the higher the similarity between IMF and original signal, the more useful information it contains.

After analysis and calculation, it can be known that, among the 12 IMFs obtained by the above decomposition, the first four IMFs occupy a relatively high energy of the original signal, so these four IMFs are selected as the IMFs of the reconstructed signals. The first four IMFs are added and reconstructed to obtain a new fault signal. The time domain diagram and spectrum diagram of the reconstructed signal are shown in Figure 6. At the same time, the time domain diagram and spectrum diagram of the original signal and the error diagram of the reconstructed signal and the original signal are made, as shown in Figures 7 and 8, respectively.

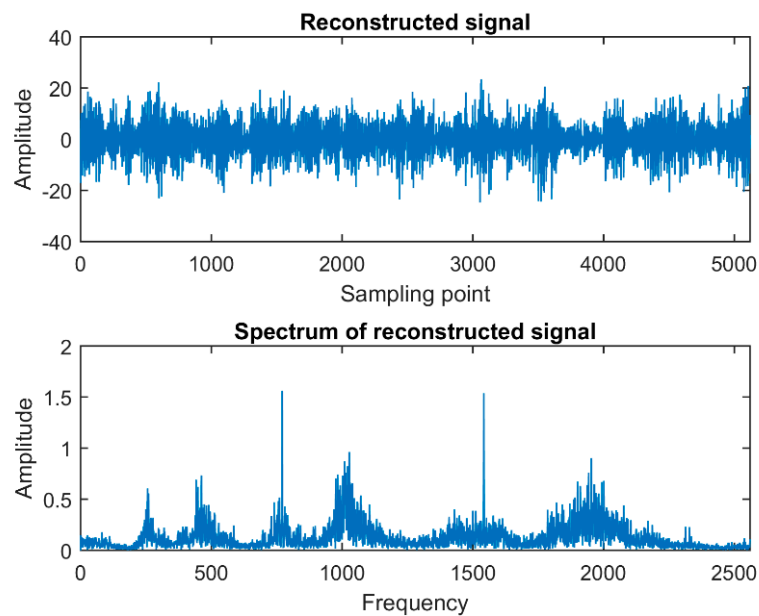


Figure 6. The time domain diagram and spectrum diagram of the reconstructed signal.

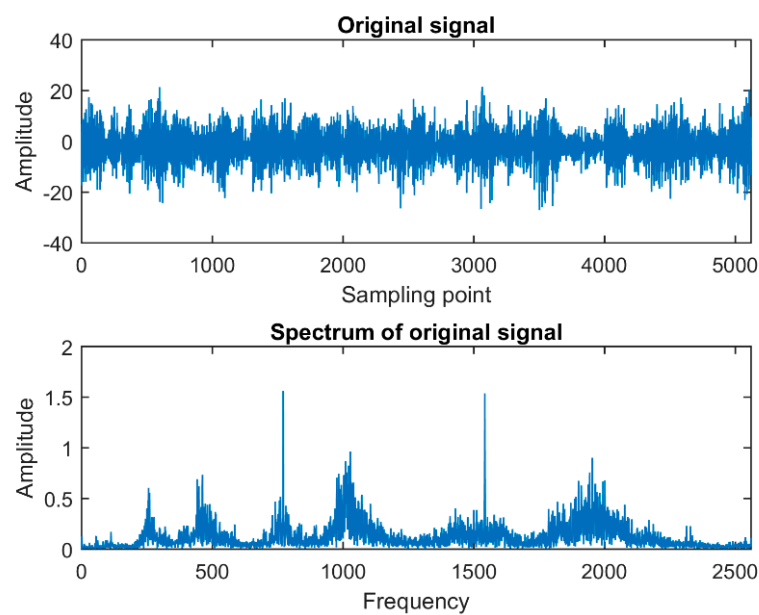


Figure 7. The time domain diagram and spectrum diagram of the original signal.

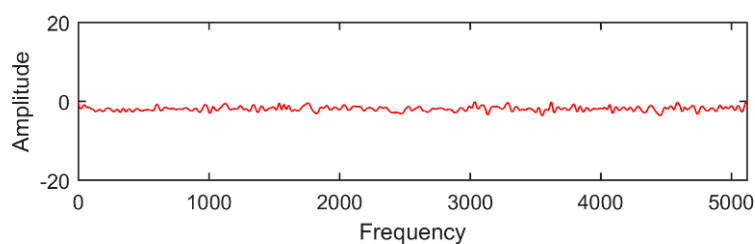


Figure 8. Error between reconstructed signal and original signal.

Through comparison, it can be found that the similarity between the new signal reconstructed by the first four IMFs and the original signal is very high, which can hardly be seen by human eyes. This is also verified in the reconstruction error diagram, which shows that the reconstructed signals retain most of the useful fault information of the original signals, and achieve the purpose of removing noise. Therefore, the reconstructed signal can be used as a new fault signal to extract fault feature parameters.

3.2. Time Domain Analysis of a Gearbox Vibration Signal

In order to extract fault information more comprehensively and effectively, 14 common time domain statistical parameters are selected as the candidate parameters. Extract the time domain feature parameters of the above new fault signals processed by the EMD method, and get the feature parameters of normal state and five fault states, respectively, as shown in Table 2.

Table 2. Time domain statistical parameters of six states.

Parameters	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5	Mode 6
P1	0.0121	0.0267	−0.2018	−0.0233	0.2207	0.0723
P2	4.2915	3.5691	20.0055	6.3319	15.9963	19.2551
P3	2.6146	2.0551	12.9985	4.0900	9.3023	12.3354
P4	3.1970	2.5351	15.5513	4.9034	11.5979	14.8461
P5	18.4172	12.7383	400.2220	40.0932	255.8814	370.7581
P6	21.4670	24.6709	82.7674	24.5781	60.6313	86.1785
P7	−19.4246	−29.4615	−69.0499	−33.4569	−91.3556	−80.3758
P8	40.8916	54.1324	151.8173	58.0350	151.9869	166.5542
P9	1.3424	1.4079	1.2864	1.2913	1.3792	1.2970
P10	5.0022	6.9124	4.1372	3.8816	3.7903	4.4756
P11	6.7148	9.7318	5.3222	5.0125	5.2278	5.8048
P12	8.2103	12.0045	6.3674	6.0093	6.5179	6.9863
P13	−0.0542	−0.0886	0.0576	−0.1437	−0.5402	−0.1214
P14	4.5205	8.4587	3.4973	3.7180	5.1208	3.7584

Each feature parameter can reflect certain fault information. The above 14 candidate feature parameters reflect different fault information from different aspects, but their sensitivity to different operating states is different. Because there is a certain correlation between the parameters, and too many parameters will inevitably lead to a large increase in the amount of calculation due to the “dimension effect”, it is necessary to carry out sensitivity analysis on 14 candidate feature parameters to screen out the feature parameters with better sensitivity, which is more conducive to the recognition of the working state of the gearbox.

Sensitivity analysis is carried out for 14 candidate feature parameters, and the analysis results are shown in Figure 9.

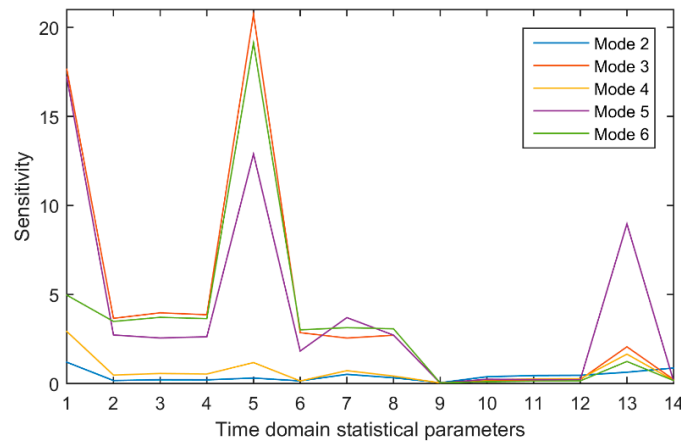


Figure 9. Sensitivity analysis results.

In Figure 9, each polyline represents the average sensitivity of each feature parameter in a fault state. For the sensitivity of a feature parameter, the larger the value in a certain failure mode, the more the feature parameter is conducive to distinguishing between normal status and failure mode. When selecting the appropriate feature parameters, not only the ability of the feature parameters to distinguish between normal and fault states, but also the ability of the feature parameters to identify different fault modes.

It can be seen from Figure 9 that the first feature parameter (P1), the fifth feature parameter (P5), and the 13th feature parameter (P13) are highly sensitive, and the differences between different fault modes are relatively large. The ability to recognize different states of the gearbox is better. Therefore, the three feature parameters of mean (P1), mean square (P5), and the skewness index (P13) are selected as the time domain feature parameters of gearbox fault.

3.3. Wavelet Packet Analysis of a Gearbox Vibration Signal

Time domain indicators can characterize part of the gearbox fault information. However, using only time domain feature parameters as feature vectors, errors in gearbox fault diagnosis results are usually very large. Therefore, it is necessary to use the signal processing technology to analyze the fault signal to obtain new fault feature parameters, and together with the fault feature parameters in the time domain to form the gear box fault feature vector.

According to relevant theories and many times of practice, compared with other wavelet functions, the wavelet basis functions based on Daubechies (db wavelet) are more suitable for analyzing the vibration signals of gearboxes, and their performance is best when the number of wavelet $n = 5$. Therefore, this paper selects Daubechies5 (called db5 wavelet) as the wavelet basis function for gearbox signal analysis, and then uses the db5 wavelet basis to perform three-layer wavelet packet decomposition on the data, and decomposes to obtain eight subbands. The wavelet packet decomposition tree structure and decomposition results are shown in Figures 10 and 11, respectively.

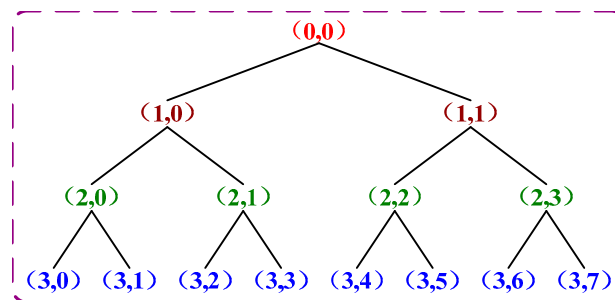


Figure 10. The wavelet packet decomposition tree structure.

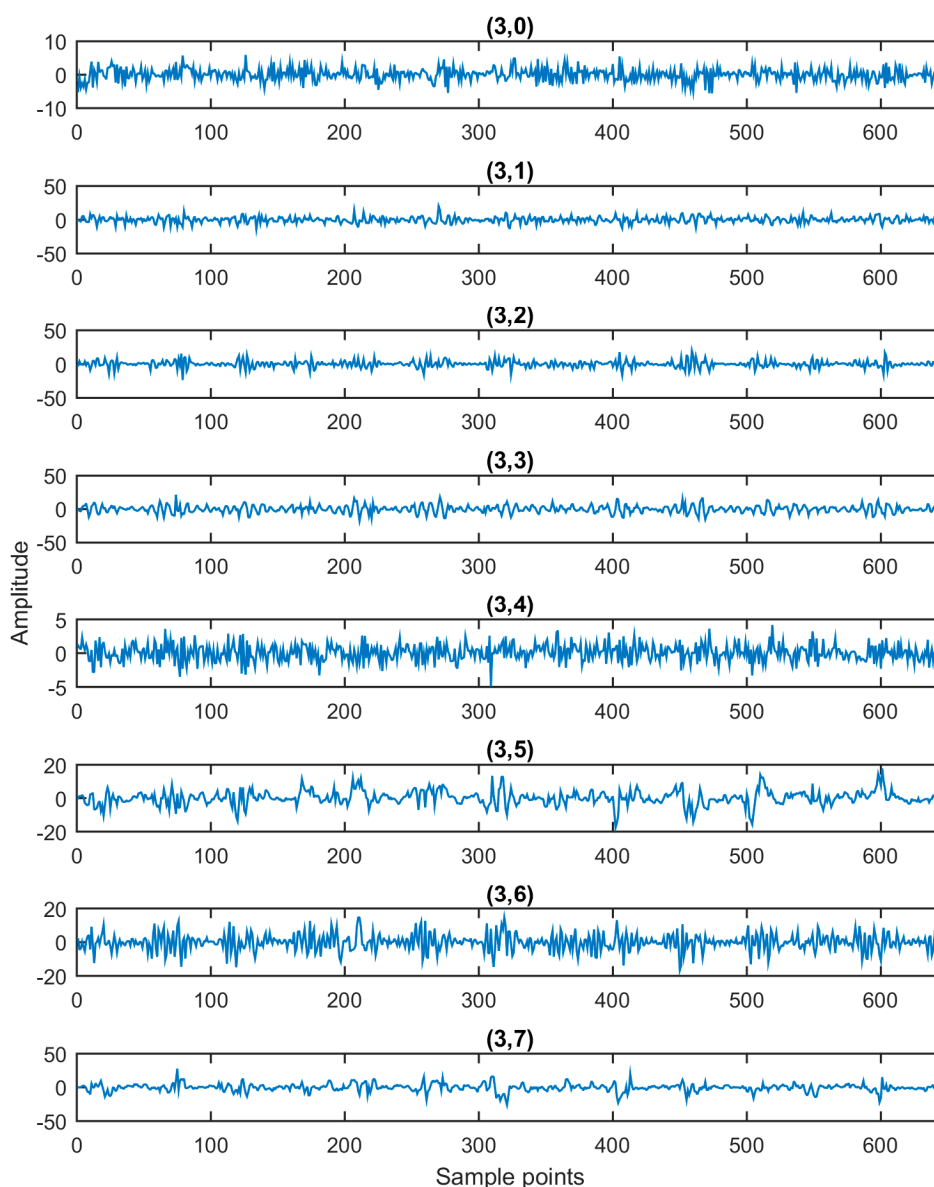


Figure 11. Wavelet packet decomposition results.

After the wavelet packet decomposition results are obtained, the energy of each frequency band in six states of gearbox is calculated, and the normalized energy values of each frequency band are used to form the energy characteristics. The frequency band energy distribution in each state is shown in Figure 12.

As can be seen from Figure 12, under different operating conditions of the gearbox, the energy characteristics of each frequency band of the vibration signal have significant differences, so it can provide an effective basis for identifying the operating state of the gearbox.

Combining the eight wavelet packet energy features with the three-time domain feature parameters obtained in the time domain analysis, the 11-dimensional gearbox state feature vector is constructed. In this experiment, 160 sets of data are collected for each state of gearbox. Therefore, after the above feature extraction, 160 sets of 11-dimensional gearbox state eigenvectors are obtained for each state. Some of the state feature vector parameter values are shown in Table 3.

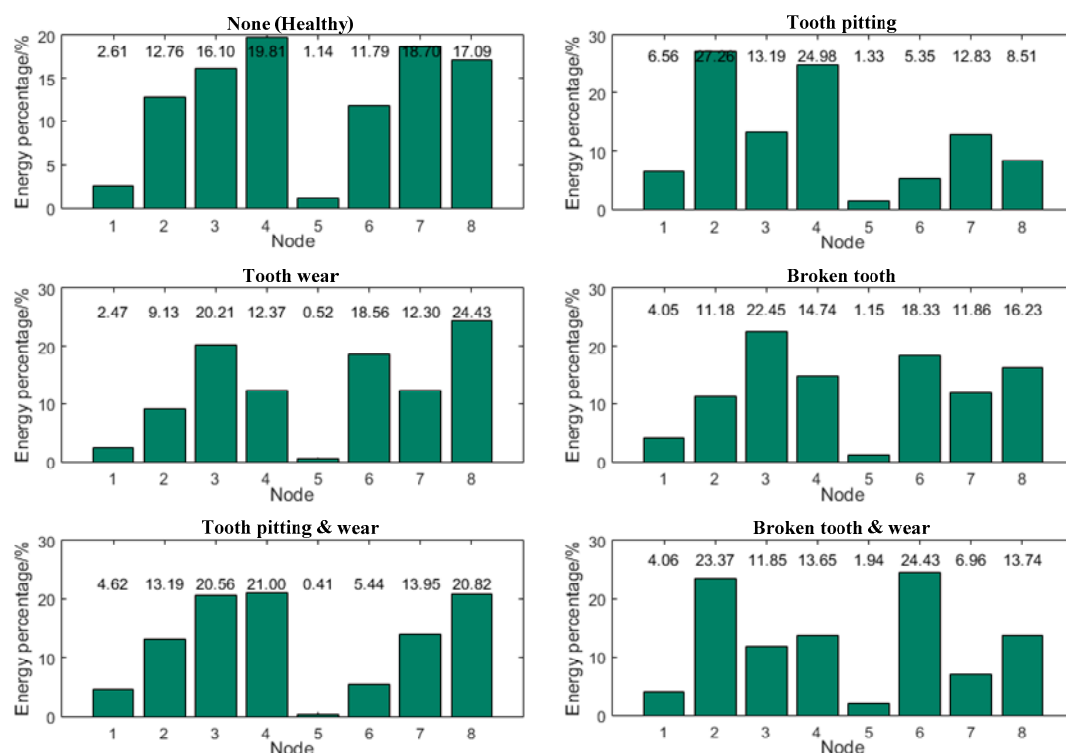


Figure 12. Histogram of frequency band energy distribution in each state.

Table 3. State feature vector.

Mode	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11
1	0.0261	0.1276	0.1610	0.1981	0.0114	0.1179	0.1870	0.1709	0.0574	18.8602	−0.0838
2	0.0656	0.2726	0.1319	0.2498	0.0133	0.0535	0.1283	0.0851	−0.0250	12.5383	−0.0711
3	0.0247	0.0913	0.2021	0.1237	0.0052	0.1856	0.1230	0.2443	−0.1034	369.1129	−0.0236
4	0.0405	0.1118	0.2245	0.1474	0.0115	0.1833	0.1186	0.1623	−0.0190	36.2753	−0.1532
5	0.0462	0.1319	0.2056	0.2100	0.0041	0.0544	0.1395	0.2082	0.1430	255.7761	−0.5698
6	0.0406	0.2337	0.1185	0.1365	0.0194	0.2443	0.0696	0.1374	0.0159	352.1008	−0.0651

3.4. Gearbox Fault Diagnosis Based on BP-AdaBoost

Firstly, the BP-AdaBoost fault diagnosis model is established. According to the dimension of data and the general rule of determining the number of neurons in the hidden layer, the structure of BP neural network is 11-6-1. At the same time, considering the requirements of calculation time and diagnosis accuracy, the number of weak classifiers is 10, and then 10 weak classifiers are used to form a strong classifier to classify the operation state of the gearbox.

Among the 160 sets of state feature vectors of each of the above states, 100 groups are randomly selected as training samples, and the remaining 60 groups are used as test samples. The training samples of each mode are input into the established BP-AdaBoost fault diagnosis model in turn, and the BP-AdaBoost model is trained. After training is completed, input the test samples and then get the test results as shown in Figure 13. The test results are statistically calculated, and the diagnostic accuracy of the BP-AdaBoost fault diagnosis model is calculated as shown in Table 4.

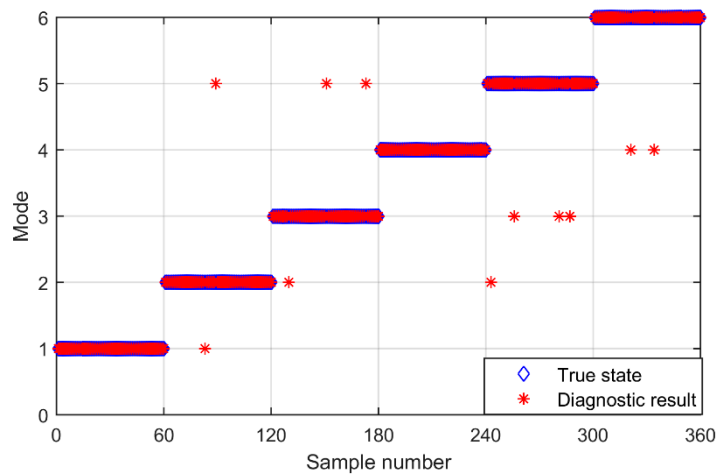


Figure 13. Test results of the BP-AdaBoost.

Table 4. Statistics of diagnosis results.

Item	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5	Mode 6	Total
Test samples	60	60	60	60	60	60	360
Correct number	60	58	57	60	56	58	349
Accuracy	100.00%	96.67%	95.00%	100.00%	93.33%	96.67%	96.94%

From the above diagnosis results, it can be seen that, on the basis of extracting the multiple features of the compound fault of gearboxes, the recognition accuracy of the gearbox fault mode using BP AdaBoost fault diagnosis model is relatively high, and two groups have achieved all the correctness, and the overall accuracy is close to 97%, which verifies the feasibility and effectiveness of the method proposed in this paper.

4. Comparison Experiment and Discussion

In order to further verify the superiority of the proposed method in gearbox compound fault diagnosis, comparative experiments are needed. In the first round of comparative experiments, the original signal is not processed by the EMD method, and the time domain analysis and wavelet packet analysis are directly performed to obtain the gearbox state feature vectors. Then, the same fault diagnosis model is used. The comparison chart is shown in Figure 14.

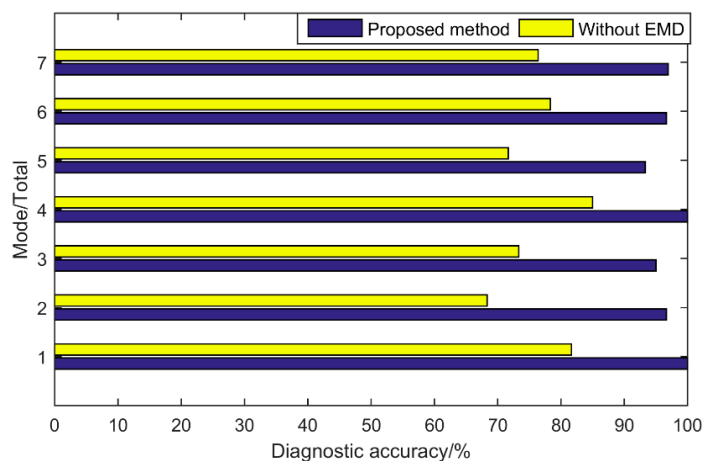


Figure 14. Comparison of proposed method and without EMD.

Through comparative analysis, it can be clearly found that the accuracy rate of the feature parameters extracted from the reconstructed vibration signal after the original vibration signal is decomposed and reconstructed by EMD is significantly higher than the accuracy rate of feature parameters extracted for fault diagnosis without EMD processing, which shows that, in the process of processing the vibration signal of the gearbox, EMD achieves the purpose of denoising, and, at the same time, the reconstructed signal contains as much useful information as possible, which can improve the signal-to-noise ratio and thereby facilitate fault diagnosis.

In the second round of comparative experiments, based on the EMD decomposition of the original vibration signal and the reconstruction signal, different feature parameters are extracted, including the time domain feature parameters, the frequency domain feature parameters, and the frequency band energy feature parameters of wavelet packet decomposition. These parameters are composed of 3D time domain feature vectors, 3D frequency domain feature vectors, and 8-dimensional energy feature vectors, respectively. Then, they are respectively input in some traditional fault diagnosis models, such as BP neural network (BPNN) and support vector machine (SVM), the final fault diagnosis results are summarized as shown in Table 5.

Table 5. Summary of diagnosis results.

Diagnostic Method	Feature Vector	Total Accuracy
BPNN	Time domain feature vectors	80.28%
BPNN	Frequency domain feature vectors	45.28%
BPNN	Energy feature vectors	81.11%
BPNN	Multi-feature vectors	89.17%
SVM	Time domain feature vectors	76.11%
SVM	Frequency domain feature vectors	50.28%
SVM	Energy feature vectors	85.28%
SVM	Multi-feature vectors	91.11%
Proposed method	Multi-feature vectors	96.94%

A number of meaningful conclusions can be clearly drawn from Table 5. First of all, for the same fault diagnosis method, the diagnosis accuracy rate is the highest when multi-feature vectors are used. When only frequency domain feature vectors are used, the accuracy is only about 50%. This is because the characteristic frequencies between compound faults of the gearbox are relatively close. It is difficult to identify the compound fault mode of the gearbox based on the frequency domain characteristics. Therefore, this paper uses a combination of sensitive time domain feature parameters and energy feature parameters. In addition, when multi-feature vectors are used at the same time, the diagnostic accuracy of the proposed method is higher than that of traditional fault diagnosis methods. Taking BPNN as an example, due to the existence of local minimum points, it can not guarantee that the network will eventually converge to the global minimum point. Therefore, the overall classification recognition effect is relatively low, and the diagnostic accuracy of the similar compound fault mode is the lowest. It should also be noted that the initialization of weights may affect the final convergence of the network, resulting in some results with strong randomness, and the diagnostic results differ greatly each time. However, BP-AdaBoost can fully take advantage of the adaptive learning method in the AdaBoost lifting method to reduce the error rate. The strong classifier formed by integrating multiple weak classifiers has small and stable classification errors.

5. Conclusions

In this paper, a diagnosis method for the compound fault of gearboxes based on the combination of multi-feature vectors composed of sensitive time domain characteristic parameters and band energy characteristics by wavelet packet decomposition and BP AdaBoost algorithm is proposed. The following conclusions can be obtained through the experiments and simulation results. By collecting and analyzing the vibration signal of the gearbox, the running state of the gearbox can be effectively

monitored. Decomposing the original vibration signal of gearbox by EMD and selecting the appropriate IMFs to reconstruct the new signal can achieve the purpose of denoising the original signal. Multiple time domain statistical parameters are calculated from the time domain analysis. Through the sensitivity analysis, the time-domain characteristic parameters that can effectively reflect the fault state can be selected, so as to avoid the overlapping or interference of multiple parameters and reduce the calculation amount. Among them, the sensitivity of mean, mean square value, and the skewness index is higher, and the ability to identify different operation states of gearbox is better. Wavelet packet decomposition is an effective feature extraction method of a gearbox vibration signal, which can extract the feature information of a vibration signal in each frequency band, and the feature vector composed of energy in each frequency band can well characterize the operation state of gearbox. The combination of the time domain feature parameters and the frequency band energy feature is used as the state feature vector of the gearbox, and the accuracy of diagnosis is close to 97% when it is input into the BP-AdaBoost fault diagnosis model, which proves the effectiveness of the method proposed in this paper. Compared with the traditional fault diagnosis methods, the advantages of the proposed method can be found. Therefore, the proposed method has great potential for online or offline fault diagnosis of a gearbox compound fault, which can effectively avoid serious accidents, ensure the safe operation of equipment, and improve economic benefits. However, it should be noted that the method proposed at present is only for the typical fault of the gear in the gearbox, but the gearbox also contains the bearing and shaft and other components, which are likely to fail, so future work will focus on solving the problem of compound fault diagnosis of multiple components in the gearbox.

Author Contributions: Methodology, Y.Z.; software, W.W.; formal analysis, Y.J.; data curation, X.S.; writing—original draft preparation, Y.Z.; writing—review and editing, Y.J.; visualization, A.L.; funding acquisition, Z.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China (No. 71871220).

Acknowledgments: The authors would like to sincerely thank all the anonymous reviewers for their valuable comments that greatly helped improve the manuscript.

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

References

1. Lin, A.; Sun, Y.; Zhang, H.; Lin, X.; Yang, L.; Zheng, Q. Fluctuating characteristics of air-mist mixture flow with conjugate wall-film motion in a compressor of gas turbine. *Appl. Therm. Eng.* **2018**, *142*, 779–792. [[CrossRef](#)]
2. Jiang, X.; Lin, A.; Ma, H.; Li, X.; Li, Y. Minimizing the thermal bridge through the columns in a refrigeration room. *Appl. Therm. Eng.* **2020**, *165*, 114565. [[CrossRef](#)]
3. Lin, A.; Zhou, J.; Fawzy, H.; Zhang, H.; Zheng, Q. Evaluation of mass injection cooling on flow and heat transfer characteristics for high-temperature inlet air in a MIPCC engine. *Int. J. Heat Mass Tran.* **2019**, *135*, 620–630. [[CrossRef](#)]
4. Irfan, M.; Saad, N.; Alwadie, A.; Awais, M.; Sheikh, M.; Glowacz, A.; Kumar, V. An Automated Feature Extraction Algorithm for Diagnosis of Gear Faults. *J. Fail. Anal. Prev.* **2018**, *19*, 98–105. [[CrossRef](#)]
5. Lin, A.; Zheng, Q.; Fawzy, H.; Luo, M.; Zhou, J.; Zhang, H. Effect of water injection cooling on flow field characteristics in the cooling section of precooled turbine-based combined cycle engine. *Int. J. Heat Mass Tran.* **2019**, *141*, 615–626. [[CrossRef](#)]
6. Liu, Z.; Liu, Y.; Shan, H.; Cai, B.; Huang, Q. A Fault Diagnosis Methodology for Gear Pump Based on EEMD and Bayesian Network. *PLoS ONE* **2015**, *10*, e0125703. [[CrossRef](#)] [[PubMed](#)]
7. Lin, A.; Zheng, Q.; Jiang, Y.; Lin, X.; Zhang, H. Sensitivity of air/mist non-equilibrium phase transition cooling to transient characteristics in a compressor of gas turbine. *Int. J. Heat Mass Tran.* **2019**, *137*, 882–894. [[CrossRef](#)]
8. Liang, X.; Zuo, M.J.; Feng, Z. Dynamic modeling of gearbox faults: A review. *Mech. Syst. Signal Process.* **2018**, *98*, 852–876. [[CrossRef](#)]

9. Yu, J. A selective deep stacked denoising autoencoders ensemble with negative correlation learning for gearbox fault diagnosis. *Comput. Ind.* **2019**, *108*, 62–72. [[CrossRef](#)]
10. Han, Y.; Tang, B.; Deng, L. An enhanced convolutional neural network with enlarged receptive fields for fault diagnosis of planetary gearboxes. *Comput. Ind.* **2019**, *107*, 50–58. [[CrossRef](#)]
11. Chen, R.; Huang, X.; Yang, L.; Xu, X. Intelligent fault diagnosis method of planetary gearboxes based on convolution neural network and discrete wavelet transform. *Comput. Ind.* **2019**, *106*, 48–59. [[CrossRef](#)]
12. Cheng, Y.; Zhu, H.; Wu, J.; Shao, X. Machine Health Monitoring Using Adaptive Kernel Spectral Clustering and Deep Long Short-Term Memory Recurrent Neural Networks. *IEEE Trans. Ind. Informat.* **2019**, *15*, 987–997. [[CrossRef](#)]
13. Zhang, M.; Wang, K.; Wei, D.; Zuo, M.J. Amplitudes of characteristic frequencies for fault diagnosis of planetary gearbox. *J. Sound Vib.* **2018**, *432*, 119–132. [[CrossRef](#)]
14. Figlus, T. A method for diagnosing gearboxes of means of transport using multi-stage filtering and entropy. *Entropy* **2019**, *21*, 441. [[CrossRef](#)]
15. Wu, C.; Jiang, P.; Ding, C.; Feng, F. Intelligent fault diagnosis of rotating machinery based on one-dimensional convolutional neural network. *Comput. Ind.* **2019**, *108*, 53–61. [[CrossRef](#)]
16. Sun, R.; Yang, Z.; Chen, X.; Tian, S.; Xie, Y. Gear fault diagnosis based on the structured sparsity time-frequency analysis. *Mech. Syst. Signal Process.* **2018**, *102*, 346–363. [[CrossRef](#)]
17. Li, Z.; Wang, Y.; Wang, K. A deep learning driven method for fault classification and degradation assessment in mechanical equipment. *Comput. Ind.* **2019**, *104*, 1–10. [[CrossRef](#)]
18. Wu, J.; Wu, C.; Cao, S.; Or, S.W. Degradation Data-Driven Time-To-Failure Prognostics Approach for Rolling Element Bearings in Electrical Machines. *IEEE Trans. Ind. Electron.* **2018**, *66*, 529–539. [[CrossRef](#)]
19. Ma, S.; Chu, F. Ensemble deep learning-based fault diagnosis of rotor bearing systems. *Comput. Ind.* **2019**, *105*, 143–152. [[CrossRef](#)]
20. Mohammed, O.D.; Rantatalo, M.; Aidanpää, J.-O. Dynamic modelling of a one-stage spur gear system and vibration-based tooth crack detection analysis. *Mech. Syst. Signal Process.* **2015**, *54–55*, 293–305. [[CrossRef](#)]
21. Zhao, D.; Wang, T.; Chu, F. Deep convolutional neural network based planet bearing fault classification. *Comput. Ind.* **2019**, *107*, 59–66. [[CrossRef](#)]
22. Lei, Y.; Lin, J.; Zuo, M.J.; He, Z. Condition monitoring and fault diagnosis of planetary gearboxes: A review. *Measurement* **2014**, *48*, 292–305. [[CrossRef](#)]
23. Ma, H.; Zeng, J.; Feng, R.; Pang, X.; Wang, Q.; Wen, B. Review on dynamics of cracked gear systems. *Eng. Fail. Anal.* **2015**, *55*, 224–245. [[CrossRef](#)]
24. Feng, Z.; Zuo, M.J.; Chu, F. Application of regularization dimension to gear damage assessment. *Mech. Syst. Signal Process.* **2010**, *24*, 1081–1098. [[CrossRef](#)]
25. Feng, Z.; Zuo, M.J. Fault diagnosis of planetary gearboxes via torsional vibration signal analysis. *Mech. Syst. Signal Process.* **2013**, *36*, 401–421. [[CrossRef](#)]
26. Benmoussa, S.; Djeziri, M.A. Remaining useful life estimation without needing for prior knowledge of the degradation features. *IET Sci. Meas. Technol.* **2017**, *11*, 1071–1078. [[CrossRef](#)]
27. Tang, J.; Shi, B.; Li, Z. Asymmetric delay feedback stochastic resonance method based on prior knowledge particle swarm optimization. *Chin. J. Phys.* **2018**, *56*, 2104–2118. [[CrossRef](#)]
28. Guo, J.; Shi, Z.; Li, H.; Zhen, D.; Gu, F.; Ball, A.D. Early Fault Diagnosis for Planetary Gearbox Based Wavelet Packet Energy and Modulation Signal Bispectrum Analysis. *Sensors* **2018**, *18*, 2908. [[CrossRef](#)]
29. Li, Y.; Ding, K.; He, G.; Jiao, X. Non-stationary vibration feature extraction method based on sparse decomposition and order tracking for gearbox fault diagnosis. *Measurement* **2018**, *124*, 453–469. [[CrossRef](#)]
30. Bartelmus, W. New focus on gearbox condition monitoring for failure prevention technology. *Key Eng. Mater.* **2013**, *588*, 184–191. [[CrossRef](#)]
31. Adamczak, S.; Stepień, K.; Wrzochal, M. Comparative study of measurement systems used to evaluate vibrations of rolling bearings. *Procedia Eng.* **2017**, *192*, 971–975. [[CrossRef](#)]
32. Guo, Y.; Zhao, L.; Wu, X.; Na, J. Vibration separation technique based localized tooth fault detection of planetary gear sets: A tutorial. *Mech. Syst. Signal Process.* **2019**, *129*, 130–147. [[CrossRef](#)]
33. Jing, L.; Zhao, M.; Li, P.; Xu, X. A convolutional neural network based feature learning and fault diagnosis method for the condition monitoring of gearbox. *Measurement* **2017**, *111*, 1–10. [[CrossRef](#)]
34. Zhang, R.; Tao, H.Y.; Wu, L.F.; Guan, Y. Transfer learning with neural networks for bearing fault diagnosis in changing working conditions. *IEEE Access* **2017**, *5*, 14347–14357. [[CrossRef](#)]

35. He, M.; He, D. Deep learning based approach for bearing fault diagnosis. *IEEE Trans. Ind. Appl.* **2017**, *53*, 3057–3065. [[CrossRef](#)]
36. Baraldi, P.; Cannarile, F.; Di Maio, F.; Zio, E. Hierarchical k-nearest neighbours classification and binary differential evolution for fault diagnostics of automotive bearings operating under variable conditions. *Eng. Appl. Artif. Intell.* **2016**, *56*, 1–13. [[CrossRef](#)]
37. Wu, S.-D.; Wu, C.-W.; Wu, T.-Y.; Wang, C.-C. Multi-Scale Analysis Based Ball Bearing Defect Diagnostics Using Mahalanobis Distance and Support Vector Machine. *Entropy* **2013**, *15*, 416–433. [[CrossRef](#)]
38. Wang, S.; Xiang, J.; Zhong, Y.; Zhou, Y. Convolutional neural network-based hidden Markov models for rolling element bearing fault identification. *Knowl. Based Syst.* **2018**, *144*, 65–76. [[CrossRef](#)]
39. Xia, M.; Li, T.; Xu, L.; Liu, L. Fault Diagnosis for Rotating Machinery Using Multiple Sensors and Convolutional Neural Networks. *IEEE/ASME Trans. Mechatron.* **2018**, *23*, 101–110. [[CrossRef](#)]
40. Chaari, F.; Baccar, W.; Abbas, M.S.; Haddar, M. Effect of spalling or tooth breakage on gearmesh stiffness and dynamic response of a one-stage spur gear transmission. *Eur. J. Mech. A Solids* **2008**, *27*, 691–705. [[CrossRef](#)]
41. Gelman, L.; Zimroz, R.; Birkel, J.; Leigh-Firbank, H.; Simms, D.; Waterland, B.; Whitehurst, G. Adaptive vibration condition monitoring technology for local tooth damage in gearboxes. *Insight Non-Destr. Test. Cond. Monit.* **2005**, *47*, 461–464. [[CrossRef](#)]
42. Liu, X.; Yang, Y.; Zhang, J. Investigation on coupling effects between surface wear and dynamics in a spur gear system. *Tribol. Int.* **2016**, *101*, 383–394. [[CrossRef](#)]
43. Sun, X.W.; Li, J.H.; Wang, S.P. Nonlinear dynamics for gear fault level. *Open Mech. Eng. J.* **2014**, *8*, 487–496.
44. Jiang, G.Q.; He, H.B.; Yan, J.; Xie, P. Multiscale Convolutional Neural Networks for Fault Diagnosis of Wind Turbine Gearbox. *IEEE Trans. Ind. Electron.* **2019**, *66*, 3196–3207. [[CrossRef](#)]
45. 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. Lond. A Math. Phys. Eng. Sci.* **1998**, *454*, 903–995. [[CrossRef](#)]
46. Lei, Y.; Li, N.; Lin, J.; Wang, S. Fault diagnosis of rotating machinery based on an adaptive ensemble empirical mode decomposition. *Sensors* **2013**, *13*, 16950–16964. [[CrossRef](#)]
47. Merainani, B.; Rahmoune, C.; Benazzouz, D.; Ould-Bouamama, B. A novel gearbox fault feature extraction and classification using Hilbert empirical wavelet transform, singular value decomposition, and SOM neural network. *J. Vib. Control* **2018**, *24*, 2512–2531. [[CrossRef](#)]
48. Parey, A.; Singh, A. Gearbox fault diagnosis using acoustic signals, continuous wavelet transform and adaptive neuro-fuzzy inference system. *Appl. Acoust.* **2019**, *147*, 133–140. [[CrossRef](#)]
49. Freund, Y.; Schapire, R.E. *Experiments with a New Boosting Algorithm, Thirteen International Conference on Machine Learning*; Morgan Kaufmann: Burlington, MA, USA, 1996; pp. 148–156.



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).