


Article

A New Fusion Fault Diagnosis Method for Fiber Optic Gyroscopes

Wanpeng Zhang, Dailin Zhang * , Peng Zhang and Lei Han

School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, China; m202070578@hust.edu.cn (W.Z.); m202170784@hust.edu.cn (P.Z.); m201970599@hust.edu.cn (L.H.)

* Correspondence: mnizhang@hust.edu.cn

Abstract: The fiber optic gyroscope (FOG) is a high precision inertial navigation device, and it is necessary to ensure its reliability for effective use. However, the extracted fault features are easily distorted due to the interference of vibrations when the FOG is in operation. In order to minimize the influence of vibrations to the greatest extent, a fusion diagnosis method was proposed in this paper. It extracted features from fault data with Fast Fourier Transform (FFT) and wavelet packet decomposition (WPD), and built a strong diagnostic classifier with a sparse auto encoder (SAE) and a neural network (NN). Then, a fusion neural network model was established based on the diagnostic output probabilities of the two primary classifiers, which improved the diagnostic accuracy and the anti-vibration capability. Then, five fault types of the FOG under random vibration conditions were established. Fault data sets were collected and generated for experimental comparison with other methods. The results showed that the proposed fusion fault diagnosis method could perform effective and robust fault diagnosis for the FOG under vibration conditions with a high diagnostic accuracy.

Keywords: fiber optic gyroscope; fault diagnosis; SAE; WPD



Citation: Zhang, W.; Zhang, D.; Zhang, P.; Han, L. A New Fusion Fault Diagnosis Method for Fiber Optic Gyroscopes. *Sensors* **2022**, *22*, 2877. <https://doi.org/10.3390/s22082877>

Academic Editors: Adam Glowacz, Jose A Antonino-Daviu and Wahyu Caesarendra

Received: 6 March 2022

Accepted: 5 April 2022

Published: 8 April 2022

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



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

1. Introduction

Fault diagnosis is aimed to identify the type, location, time, and size of a fault [1]. In this paper, fault diagnosis of FOG is defined as the ability to effectively identify the normal state of an FOG, as well as five different types of faults. With the development of machine learning, data-driven fault diagnosis methods are widely used. In these methods, the historical data of sensors becomes the main means of fault identification. The process of diagnosis can usually be divided into three steps: signal processing, feature extraction, and fault identification, as shown in Figure 1. The original data are converted into low-dimensional features with better classification performance. Then, the features are fed into the classifier to achieve the identification and classification of faults. Now, more and more intelligent algorithms have been used for fault diagnosis of industrial equipment, the integration of which has greatly improved the accuracy of diagnosis and ensures the stable operation of equipment under actual working conditions [2]. Neural networks [3,4], support vector machines [5], and random forests [6,7] are widely used for the fault diagnosis of rolling bearings. Zahoor et al. [8] proposed the information ratio principal component analysis (Ir-PCA) method, which is applied to the fault diagnosis of multistage centrifugal pumps (MCP); Nguyen et al. [9] proposed a Stacked sparse autoencoder-based deep neural network (SSA-DNN) for the diagnosis of variable speed gear faults. In more complex industrial scenarios, methods such as transfer learning [10,11], ensemble learning [12,13] wavelet packet transform [14–16], and clustering [17,18] are also widely adopted. These intelligent algorithms have played a significant role in the fault diagnosis of modern industrial equipment and sensors.

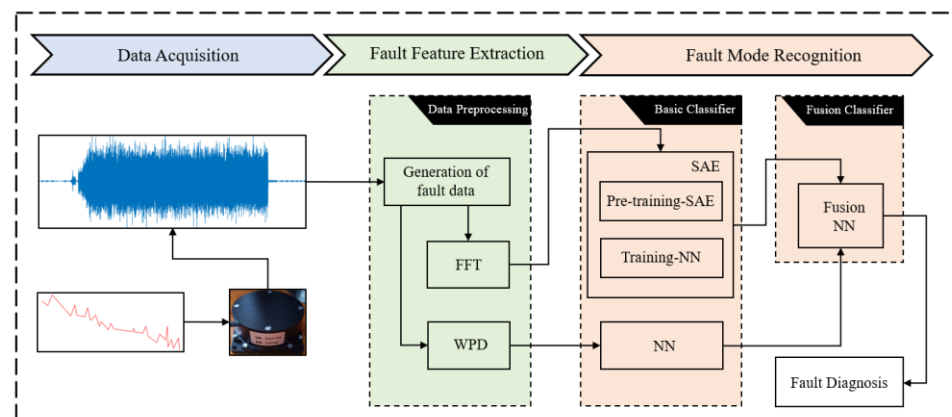


Figure 1. The framework of the proposed fusion model to FOG fault diagnosis.

The fiber optic gyroscope is an angular velocity-sensitive detection device with high measurement accuracy, which can even be used for a long time under harsh working conditions. FOG plays a complementary role for laser targeting in shield attitude measurement systems [19,20]. Therefore, it is essential to diagnose the faults of FOG and obtain its fault classes so as to ensure the stability and reliability of the sensors during construction. Yu et al. [21] proposed the genetic algorithm based on a cloud-model and radial basis function neural network for the online fault diagnosis of gyroscopes; Liu et al. [22] presented the least squares support vector machine based on wavelet packet decomposition for gyroscope fault diagnosis. In terms of FOG fault diagnosis, Chen et al. [23] put forward an improved sparrow search algorithm based on the support vector machine method, while Guan [24] proposed a CNN-based deep feed forward network.

Vibrations generated by the running shield will cause the measurement of FOG to be out of calibration or invalid [25]. Although many intelligent methods have obtained exciting results in FOG fault diagnosis, most of them collect or generate fault data under ideal conditions, and do not take into account the influence of vibration. Signal interference caused by random vibrations of various frequencies during construction [26,27] will combine the fault signal features of the fiber optic gyroscope with the vibration signal features, further reducing the extraction ability of the fault features, thereby decreasing the diagnosis accuracy [28]. To solve this problem, a new fusion fault diagnosis method for FOG is proposed in this paper, according to the outstanding contribution of neural networks in fault diagnosis. Firstly, two feature extraction methods, FFT and WPD, are adopted in this paper to extract fault features. Secondly, two classifiers with high diagnostic accuracy, FFT-based SAE and WPD-based NN, are established. Finally, a fusion fault diagnosis network is constructed with the output of the basic classifier as the feature vector. The fusion model proposed is compared with other fault diagnosis methods, including BP, SVM, CNN, and random forest (RF), etc., to verify the effectiveness of the fusion model. The contributions of this paper can be summarized as follows:

(1). In this paper, five failure modes of FOG are established. The complementary characteristics of the frequency domain features and energy statistical features of the fault data under vibration conditions are verified by normalized Wasserstein distance.

(2). In this paper, a fusion model is proposed for FOG fault diagnosis. The model combines two neural network diagnostic classifiers of different depths. This approach can improve the accuracy and robustness of fault detection under vibration conditions, compared to existing fault diagnosis methods.

2. Materials and Methods

The framework of the fusion model proposed is shown in Figure 1. The data of FOG under vibration conditions were collected in the data preprocessing stage to generate fault data. The second stage is the training of two basic classifiers, namely a sparse auto

encoder neural network based on fast Fourier transform, and a neural network based on wavelet packet decomposition. The third stage is the training of the fusion classifier. In this section, the theoretical basis and parameter settings of these three classifiers will be further explained.

2.1. Sparse Auto Encoder Neural Network Based on Fast Fourier Transform

The auto encoder is a system that reduces the dimensionality of its input. The auto encoder consists of an encoder and a decoder part, as shown in Figure 2. The main purpose of the auto encoder is to transform the input x from a high-dimensional space into a low-dimensional feature space to obtain the intermediate variable ξ . Then, ξ is reconstructed from the feature space into an \hat{x} . Then the difference between x and \hat{x} is minimized. When the dimensionality of the intermediate variable ξ (the node numbers in the hidden layer) is smaller than the dimensionality of the input x (the node numbers in the input layer), the intermediate variable ξ can be regarded as the dimensionality reduction data, compressed with minimal information loss to the input data. Since no label information is utilized in the encoding and decoding process, it can be used in an unsupervised task. Many variants of the auto encoder, such as the denoising auto encoder [29], shrinking AE [30], and SAE [31,32], are created to better learn feature reformulation and avoid overfitting. A sparse auto encoder (SAE) is actually a sparse constraint added to the encoding to penalize the activation output values of the hidden layer [33]. In an SAE, the hidden cell $\rho_k (k = 1, 2, \dots, n)$ can be calculated by Equation (1), where f_s is the activation function of Sigmoid and it will be close to zero, which means that most of the hidden layer is suppressed. The Kullback–Leibler (KL) divergence for measuring these deviations can be written as Equation (2):

$$\rho_k = \frac{1}{n} \sum_{i=1}^n [f_s(b_k + W_{ik}x_i)] \quad (1)$$

$$\text{KL}(\rho \parallel \rho_k) = \rho \log \frac{\rho}{\rho_k} + (1 - \rho) \log \frac{1 - \rho}{1 - \rho_k} \quad (2)$$

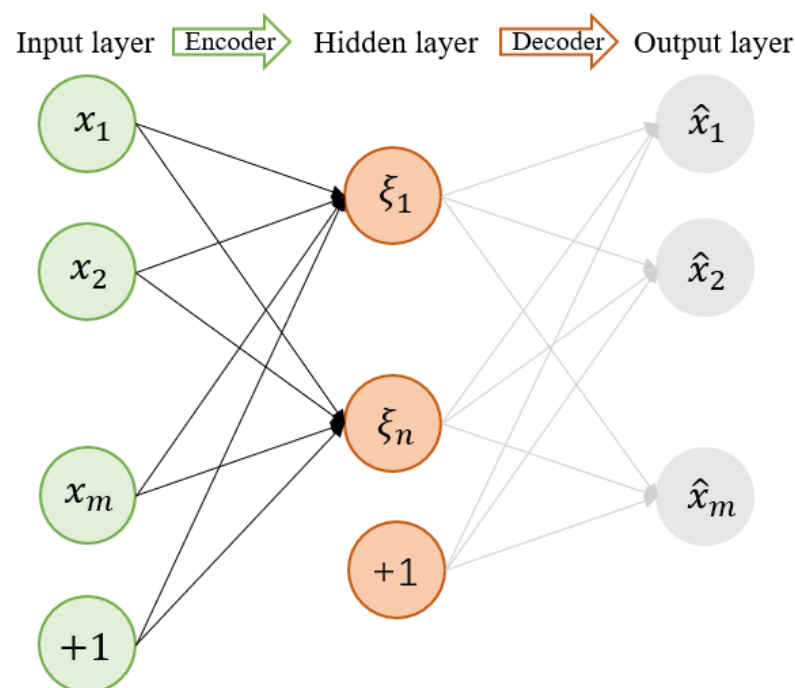


Figure 2. The schematic of the auto encoder.

In addition, the loss function needs to be regularized to prevent overfitting. β and γ are the penalty coefficients of the sparse and regular terms, respectively. The total cost function of SAE in pre-training is shown as Equation (3), where $\sum_{i=1}^n L(x_i, \hat{x}_i)$ is the mean square error (MSE) loss of decoding \hat{x}_i and the true x_i

$$J_{\text{SAE}}(\theta) = \sum_{i=1}^n L(x_i, \hat{x}_i) + \beta \text{KL}(\rho \parallel \rho_k) + \frac{\gamma}{2} \sum_{i=1}^m \sum_{k=1}^s (W_{ik}^2 + W_{ki}^2) \quad (3)$$

After pre-training, the parameters of the input and hidden layers are copied to the classifier, and the total loss function of the classifier is given by the output of the fully connected and SoftMax layers, as shown in Figure 3. In addition, the cross-entropy loss function is shown as:

$$L = \frac{1}{N} \sum_i L_i = -\frac{1}{N} \sum_i \sum_{c=1}^M y_{ic} \log(p_{ic}) \quad (4)$$

where M is the number of categories, y_{ic} is the sign function 0 or 1, N is the number of samples, and p_{ic} is the predicted probability that observation sample i belongs to class c .

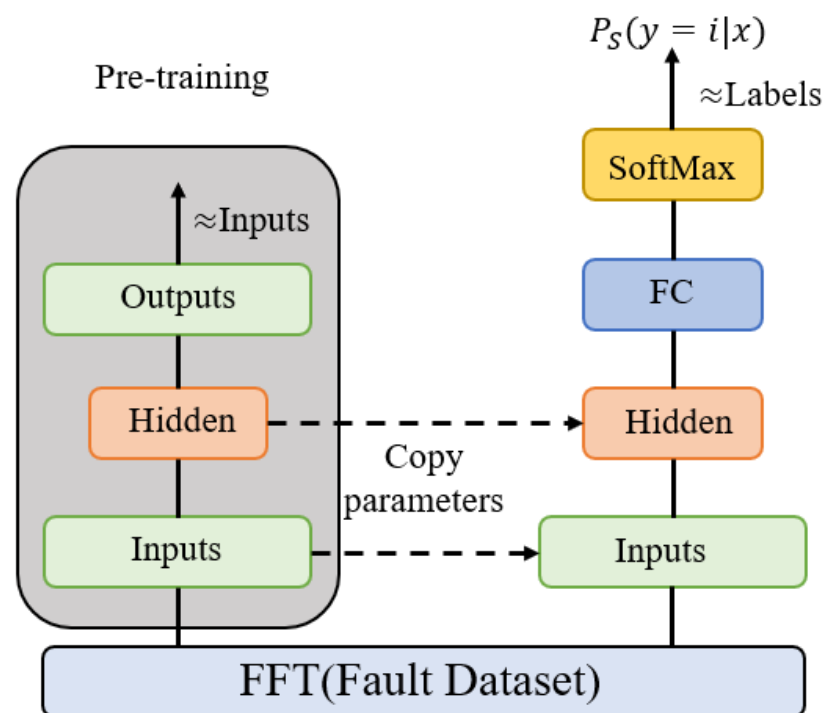


Figure 3. The sparse auto encoder neural network is based on FFT.

The input data with FFT will still be combined with fault features and vibration frequencies. At this time, the sparse auto encoder can obtain better low-dimensional features, and further improve the diagnostic accuracy.

2.2. Neural Networks Based on Wavelet Packet Decomposition

Wavelet packet decomposition (WPD) can extract the features of the full-band information of the FOG signal under vibration conditions.

Figure 4 shows the schematic of the three-layer wavelet decomposition. In the figure, A represents the approximation filtering, which can extract the low frequency part of the previous decomposition; D represents the detailed filtering, which can extract the high frequency part of the previous decomposition. The frequency and intensity of the vibration received by the FOG are different due to the different degrees of abrupt change

at fault. Thus, the db10 wavelet of the Daubechies wavelet family with good symmetry and regularity is chosen to carry out a seven-layer decomposition. It can better extract the characteristics of the FOG output signal in each frequency band.

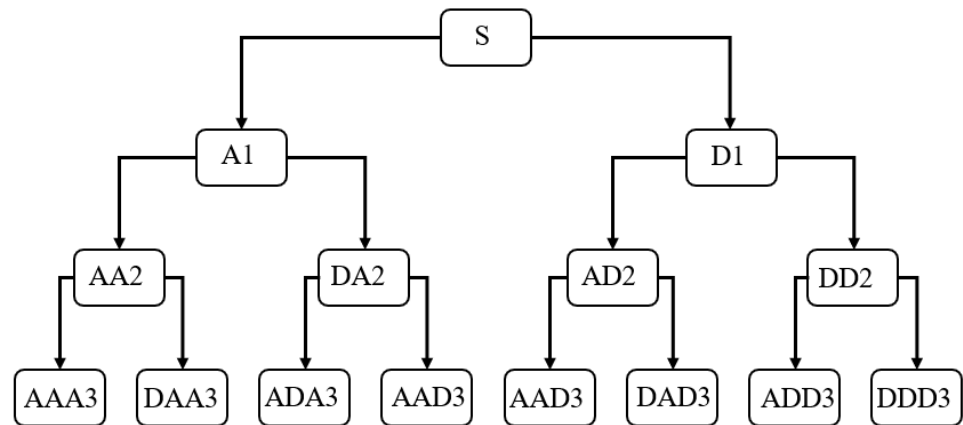


Figure 4. The schematic of wavelet packet decomposition.

For a given set of orthogonal scale transform functions and wavelet functions, the following conditions should be satisfied:

$$\begin{cases} \phi(t) = \sqrt{2} \sum_k h_k \phi(2t - k) \\ \psi(t) = \sqrt{2} \sum_g g_k \phi(2t - k) \end{cases} \quad k \in Z \quad (5)$$

$\phi(t)$ is the orthogonal scale transform function; $\psi(t)$ is the wavelet function; h_k is the orthogonal conjugate filter of $\phi(t)$; g_k is the orthogonal conjugate filter of $\psi(t)$. Make $u_0(t) = \phi(t)$, $u_1(t) = \psi(t)$:

$$\begin{cases} u_{2n}(t) = \sqrt{2} \sum_k h_k u_n(2t - k) \\ u_{2n+1}(t) = \sqrt{2} \sum_g g_k u_n(2t - k) \end{cases} \quad (6)$$

u_n is the wavelet packet function, and the corresponding wavelet packet decomposition algorithm is:

$$\begin{cases} d_j^{2n}[k] = \sum_{l \in Z} h_{l-2k} d_{j+1}^n[l] \\ d_j^{2n+1}[k] = \sum_{l \in Z} g_{l-2k} d_{j+1}^n[l] \end{cases} \quad (7)$$

d_j^{2n} is stranded wavelet packet reconstruction coefficients, which are level j , and number $2n$. Calculating the energy of each frequency band of the signal after wavelet packet decomposition, according to Parseval theorem, the energy of the signal is expressed as:

$$\|x(t)\|^2 = \int_{-\infty}^{\infty} |x(t)|^2 dt \quad (8)$$

In the wavelet packet decomposition, the following relationship exists:

$$\int_{-\infty}^{\infty} |x(t)|^2 dt = \sum_j \sum_k |d_j^k|^2 \quad (9)$$

d_j^k is wavelet packet reconstruction coefficients, as in Equation (7). Therefore, we can calculate the signal energy of each layer in the wavelet packet decomposition process through the wavelet packet reconstruction coefficients.

After the decomposition is completed, the decomposed energy characteristics are inputted to the input layer of the classifier, and the total loss function of the classifier is as in Equation (4) via the fully connected layer and the SoftMax layer, as shown in Figure 5.

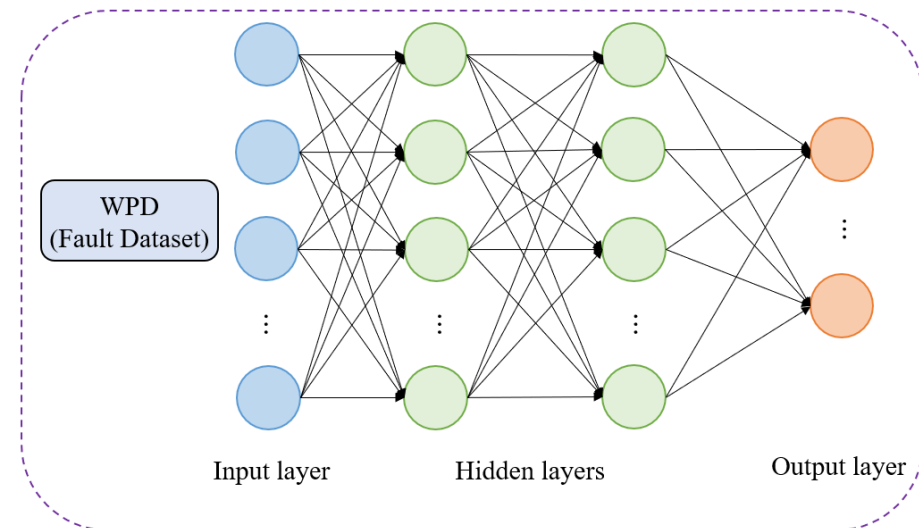


Figure 5. The neural network based on WPD.

2.3. Fusion Fault Diagnosis Model

In the vibration condition, the fault feature extraction ability and generalization of a single model are relatively weak. In this section, the advantages of the above two classifiers are integrated through model fusion to improve the prediction accuracy of the model. The voting method, a commonly used ensemble learning method, generally integrates the labeling results of multiple basic classifiers to select the one with the most output. However, individual classifiers have different detection capabilities for each fault class. Moreover, it is difficult to ensure the stability of the fusion effect when the number of base classifiers is small. For these reasons, the stacking method [34] was chosen for model fusion in this section. Stacking is a multilayer model in which two trained models are used as base classifiers, and their predicted probability results are used as a new training dataset to create a new learner that integrates the results of the two basic classifiers. The first layer of learners is defined as the basic classifier, and the second layer of learner is called the fusion classifier.

As shown in Figure 6, the two basic classifiers in this paper are FFT-based SAE and WPD-based NN. We divided the training data into basic training data and fusion training data. Firstly, the two base classifiers were trained with the basic training data. After the training was completed, the fusion training data were spliced by the 6-class output probabilities of the base classifiers to obtain the 12-dimensional data for fusion training. The fusion classifier was trained with these data to obtain the complete fusion model parameters. The parameters of each of these classifiers are listed in the Table 1.

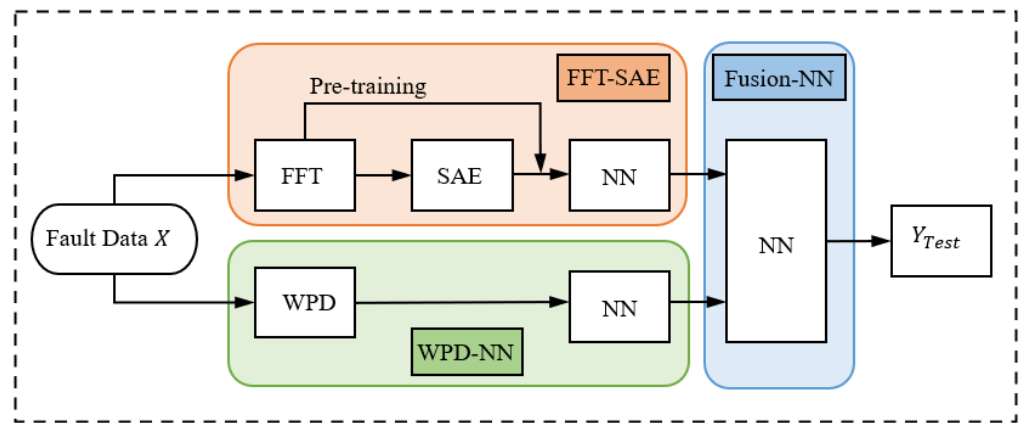


Figure 6. The schematic of the fusion model.

Table 1. Structure and hyperparameter settings of the three classifier models.

Model	Structure (Units and Activation)	Hyperparameter
FFT + SAE	Pre-training Dense (1000, 128); activation = 'Sigmoid' Dense (128, 1000)	Max epochs = 10,000; Batchsize = 6000
		Beta = 0.01
	Training	Optimizer = Adam ($lr = 0.005$)
		Dense (128, 64); activation = 'Relu'
WPD + NN	Dense (64, 32); activation = 'Relu'	Max epochs = 10,000; Batchsize = 3600
		Optimizer = Adam ($lr = 0.005$)
	Dense (32, 6); activation = 'SoftMax'	Dropout rate 0.2
	Dense (128, 64); activation = 'Relu'	Max epochs = 10,000; Batchsize = 3600
Fusion model	Dense (64, 32); activation = 'Relu'	Optimizer = Adam ($lr = 0.005$)
		Dense (32, 6); activation = 'SoftMax'
		Max epochs = 10,000; Batchsize = 1200
		Optimizer = Adam ($lr = 0.005$, weight decay = 0.0001)

3. Results

In this section, FOG vibration experiments were designed to generate fault data, and fault diagnosis experiments were conducted to verify the feasibility of the proposed method. The FOG used in the experiment is a single-axis interferometric closed-loop FOG, FOGS170A, and the technical parameters are shown in Table 2. In total, 1000 sets of data were collected under each fault class to ensure the stability of the data.

Table 2. Sensor technical parameters.

Gyroscope Type:	FOGS107A
Zero bias stability:	$<0.05^\circ/h(10s,1\sigma)$
Zero bias repeatability:	$<0.05^\circ/h(10s,1\sigma)$
Random wandering factor:	$<0.005^\circ/\sqrt{h}$
Dynamic measurement range:	$\pm 500^\circ/s$
Output method:	RS422

3.1. Data Preparation

3.1.1. Mathematical Model of FOG

In order to establish the mathematical model of FOG, the physical model needs to be studied first. The subject of the interferometric FOG is a Sagnac interferometer, which consists of a light source, coupler, light detector, integrated optical circuit, and fiber optic coil, shown as Figure 7. The light from the light source is directed into the fiber coil from the opposite direction, and, due to the Sagnac effect, when the two beams are reunited at the origin, we can measure the differential phase shift, and calculate the velocity based on the interference effect produced by the two beams [24].

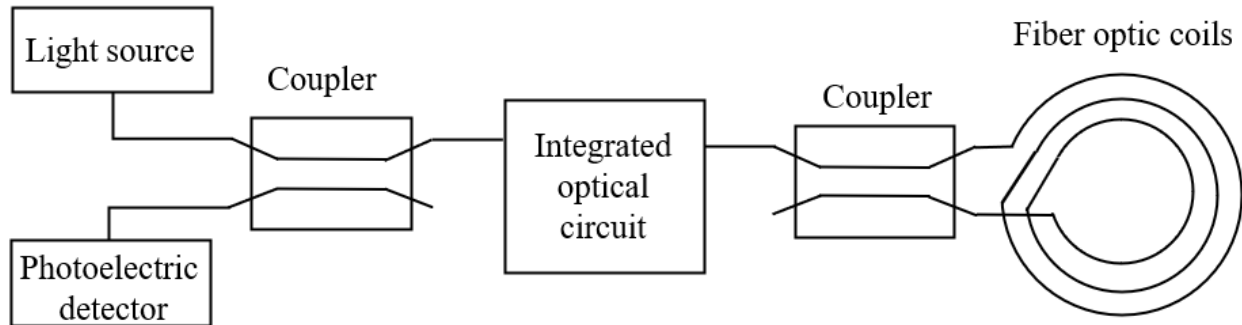


Figure 7. Physical model of FOG.

For the fiber optic gyroscope drift model [35], the effect can be described as:

$$\omega(t) = \omega_0(t) + \varepsilon(t) \quad (10)$$

where $\omega(t)$ is the output angular velocity of FOG, $\omega_0(t)$ is the actual angular velocity of FOG, and $\varepsilon(t)$ is the total drift of FOG.

$$\varepsilon(t) = \varepsilon_0 + A \sin(2\pi ft + \theta_0) + \sigma n(t) + w(t) \quad (11)$$

where ε_0 is the constant drift of FOG, A is the amplitude of the periodic component, $\sigma n(t)$ is the white noise with intensity σ , and $w(t)$ is the colored noise.

$$w(t) = \frac{4n(k-1) \sin(n(k))}{1 + n^2(k-1)} \quad (12)$$

where $n(k)$ is the sequence of white noise.

3.1.2. Generation of Fault Data

The experiment was arranged as shown in Figure 8, with the FOG placed on the shaker, and the shaker vibrating horizontally. The national military standard vibration test standard GJB150.16A-2009 Schedule C.7 vibration spectrum was used to simulate random vibration, and the frequency spectrum is shown in Figure 9. Frequency spectrum specific inflection point data can be found in Appendix A Table A1. The sampling frequency of the FOG was 100 HZ, vibration time was 600 s, and a total of 1080 s data was collected. The output curve of FOG angular velocity is shown in Figure 10.

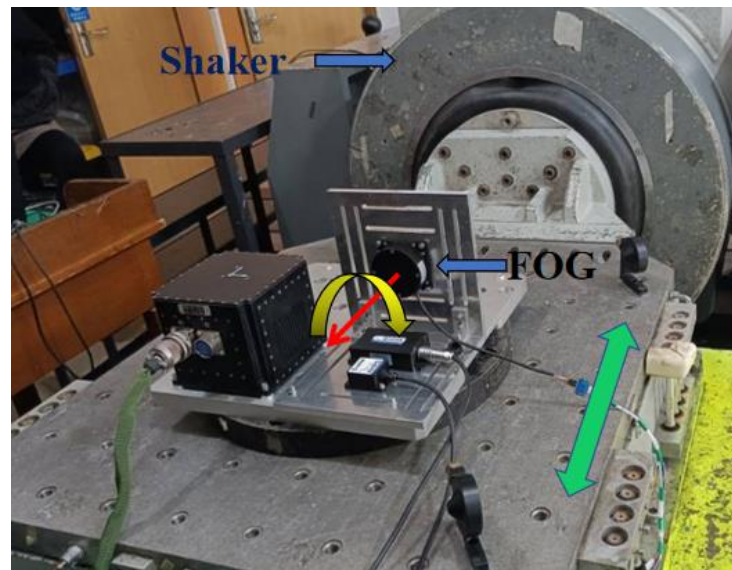


Figure 8. Experimental conditions of FOG. The green arrow represents the vibration direction of the shaker; the yellow arrow represents the positive direction of the angular velocity measured by the FOG.

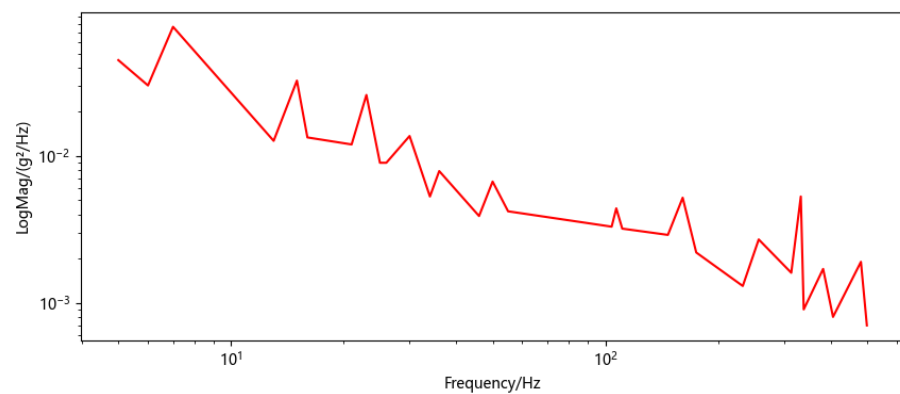


Figure 9. Random vibration power spectrum density.

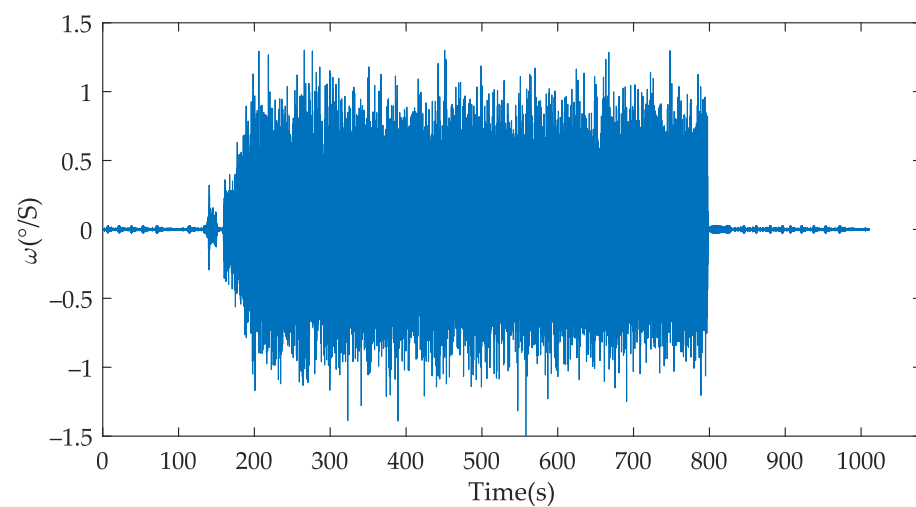


Figure 10. The output of FOG with the random vibration condition.

Since it is very difficult to obtain a large amount of fault data for FOG, the mathematical simulation method was selected in this paper to generate the fault data [35–37]. The generation of fault data was carried out based on real data. Each sample took 10 s data, 1000 points, offset 1 s, and generated 1000 groups of samples for each class under six fault classes, such as normal data and bias fault, blocking fault, drift fault, cycle fault, multiplicative fault, with a total 6000 groups of data samples.

Bias fault: FOG is subjected to a strong vibration shock. The fiber ring may be offset or deformed, resulting in a certain constant deviation of the output signal from the signal in its normal state. The mathematical model is shown in Equation (13), in which y is the output of the FOG in its normal state, x is the number of sampling points, and x_s is the number of points in the range [100, 900] when the fault occurs. k_1 is the bias amount, and the random range is $\pm[0.02, 0.2]$. The generated fault data are shown in Figure 11b.

$$y = \begin{cases} y(x) & x < x_s \\ y(x) + k_1 & x \geq x_s \end{cases} \quad (13)$$

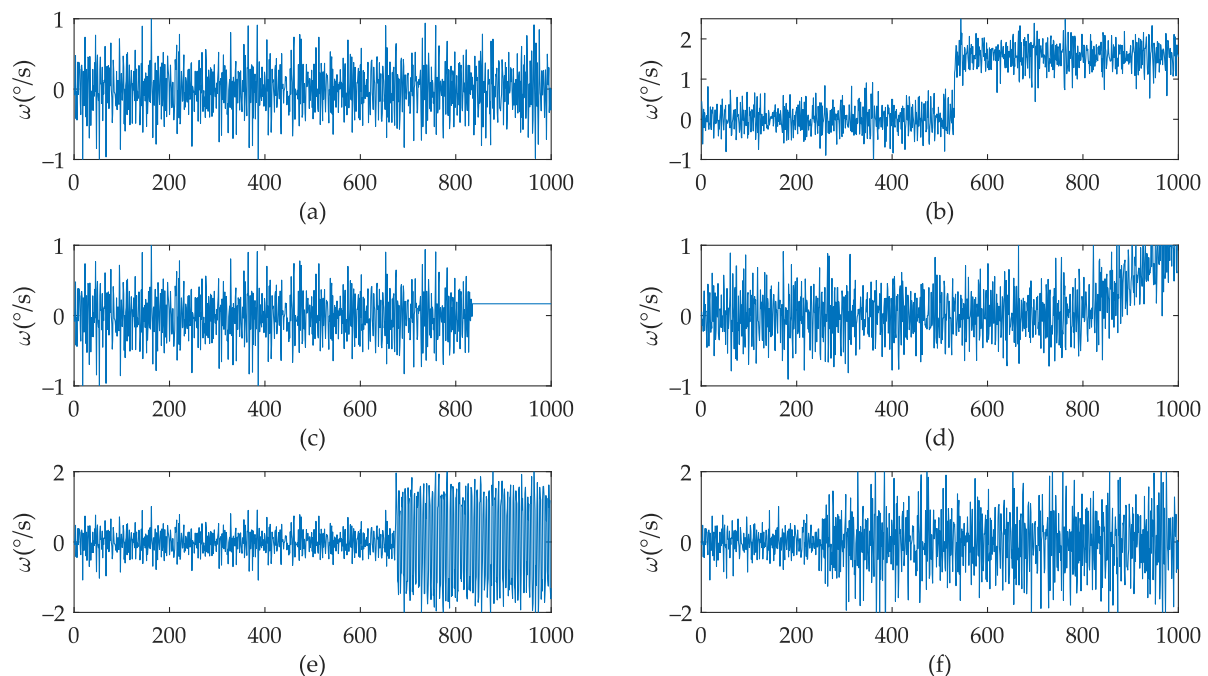


Figure 11. Fault classes of FOG. (a) normal data. (b) bias fault data. (c) blocking fault data. (d) drift fault data. (e) periodic interference fault. (f) multiplicative fault data.

Blocking fault: This refers to a point at which the gyro signal transmission is affected from a certain point onwards, keeping the same output value from the previous moment. The mathematical model is shown in Equation (14). The generated fault data are shown in Figure 11c.

$$y = \begin{cases} y(x) & x < x_s \\ y(x_s) & x \geq x_s \end{cases} \quad (14)$$

Drift fault: This is one of the most common types of failure in FOG, often due to changes in the working environment or internal parameters, such as temperature changes, causing the output of the FOG to have an increasing constant term. The output value tends to increase over time. The mathematical model is shown in Equation (15), where k_2 is the coefficient of drift, and random range is $\pm[0.002, 0.01]$. The generated fault data are shown in Figure 11d.

$$y = \begin{cases} y(x) & x < x_s \\ y(x) + k_2(x - x_s) & x \geq x_s \end{cases} \quad (15)$$

Periodic interference fault: This means that the output signal of the FOG is attached to a periodically changing signal from a certain moment. The mathematical model is shown in Equation (16), where $square(x)$ is a square wave with the same frequency as the gyroscope sampling frequency, and the amplitude random range is $\pm[0.02, 1.60]$. The generated fault data are shown in Figure 11e.

$$y = \begin{cases} y(x) & x < x_s \\ y(x) + square(x - x_s) & x \geq x_s \end{cases} \quad (16)$$

Multiplicative fault: The output signal of the FOG is multiplied by a constant from a certain point onwards due to a drastic change in the FOG scale factor, which is in turn caused by a drastic change in the operating environment. The mathematical model is shown in Equation (17), where k_3 is the multiplier factor, and random range is $[2, 8]$. The generated fault data are shown in Figure 11f.

$$y_s = \begin{cases} y(x) & x < x_s \\ k_3 y(x) & x \geq x_s \end{cases} \quad (17)$$

Figure 11 shows the time domain figure of the fault signal. In order to obtain better detection results, we need to preprocess the signal and extract more distinguishing features. Figure 12 shows the preprocessing effect of the FFT and WPD methods.

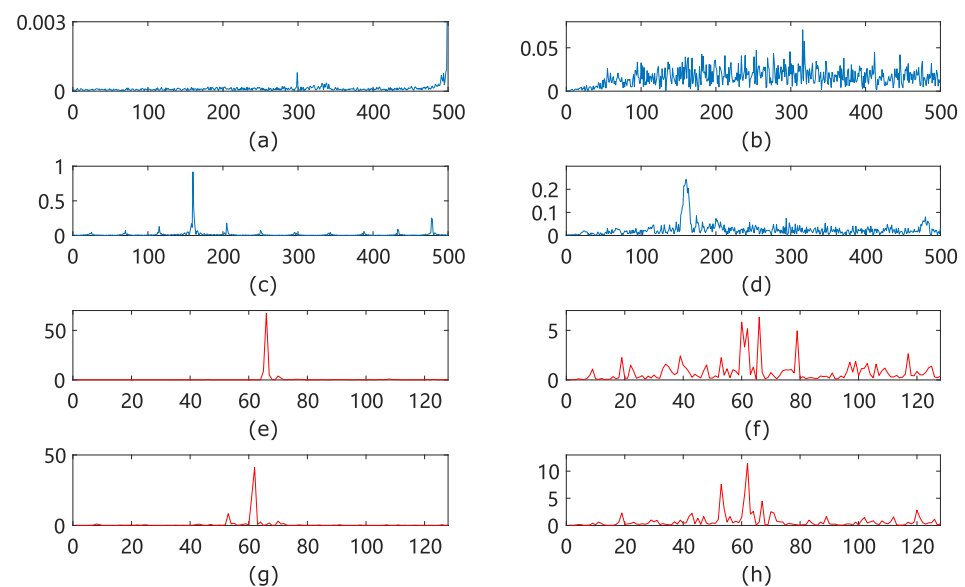


Figure 12. Preprocessing of fault data. (a) FFT of normal data. (b) FFT of normal data with vibration. (c) FFT of periodic interference fault. (d) FFT of periodic interference fault with vibration. (e) WPD of normal data. (f) WPD of normal data with vibration. (g) WPD of periodic interference fault. (h) WPD of periodic interference fault with vibration.

Fault diagnosis is a process in which the diagnostic model effectively distinguishes signals of different fault types, but the interference of vibration increases this difficulty. Figure 12a,b show the normal signal FFT processing in stationary and vibration states, respectively. Figure 12c,d show the fault signal (periodic disturbance fault) FFT processing in stationary and vibration states, respectively; Figure 12e–h show the WPD processing of the same condition with (a–d). From the figures, it can be seen that vibration is disturbing for the diagnosis of the fault classes.

To further illustrate the difference between the two processing methods, we adopt the Wasserstein distance to measure the difference of data distribution under different fault

classes, and the larger Wasserstein distance indicates that the greater difference between the two classes, the more helpful to the diagnosis of fault classes.

Figure 13 shows the normalized Wasserstein distance between different fault classes after FFT and WPD processing, with '0-1' representing the Wasserstein distance between class 0 and class 1, a symmetric value with 1-0. After the two processing methods, '0-5' and '1-3' are the smallest, respectively. This is the same as the difference between two basic classifiers diagnosed in the subsequent experimental validation. The two processing methods have different focuses for the same two classes, which indicates, to a certain degree, that the two methods have the complementary possibilities.

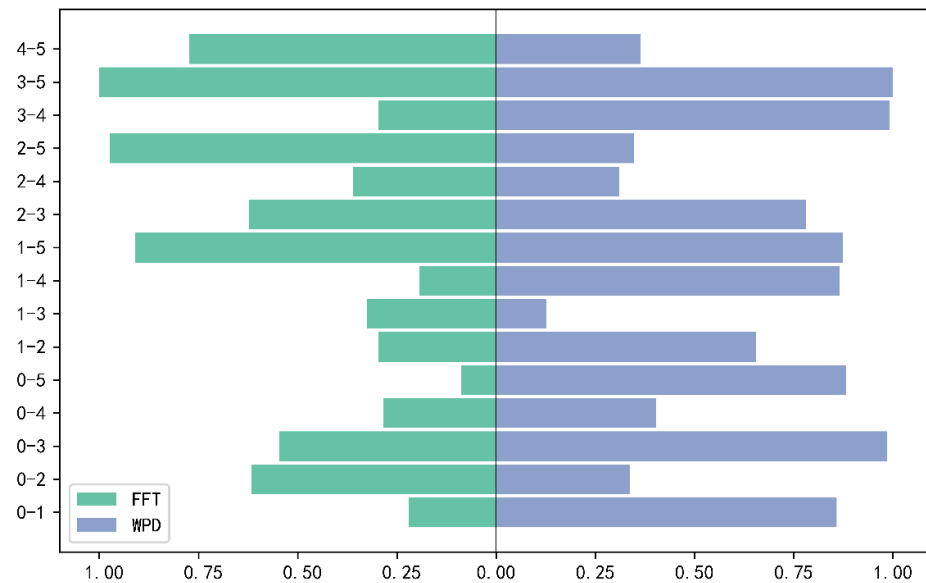


Figure 13. Normalized Wasserstein distance taper figure for two processing methods.

3.2. Experimental Results

In this section, three experiments are designed. First is the parameter selection experiment of the number of training data for the basic classifier and fusion classifier. Secondly, the stability of the model is demonstrated by five sets of cross-validation experiments. Finally, the proposed fusion model is compared with other diagnostic methods. All models in this paper are built on the Pytorch platform, and all experiments are implemented on a computing platform configured with an NVIDIA GTX 1660Ti GPU and 16 GB of RAM. The experimental results are the average of five experiments.

3.2.1. Parameter Selection Experiment

The fusion fault diagnosis method proposed in this paper is divided into two stages of classifier training. The first stage of training requires a large amount of data to ensure the performance of the basic classifier, and the second stage of training further reduces the number of input features, but still requires a sufficient amount of data to fit the distribution of test data. In this experiment, the fault dataset is divided into 80% labeled training data and 20% unlabeled test data, and the labeled data are divided into several groups with different proportions for two stages of training in order to find the most suitable proportion.

In Table 3, the ratio represents the ratio of the basic classifier and fusion classifier training data. It can be noticed from Table 3 that the accuracy of the two basic classifiers increases with the first layer training data from 93.19% and 91.94% to 95.50% and 94.44%. However, the fusion classifier is the highest at the ratio of 6:2, measuring 98.23%. It is understandable that, if the fusion classifier training data are excessive, the accuracy of the basic classifier is low, resulting in low overall accuracy; if the training data of the fusion classifier are insufficient, overfitting will reduce the accuracy, and increase the standard

deviation, resulting in poor stability. Therefore, the following experiments are conducted based on a ratio of 6:2.

Table 3. The accuracy (%) table with different parameters.

Ratio	WPD + NN	FFT + SAE	Fusion	std (%)
4:4	93.19	91.94	96.65	0.48
5:3	94.03	93.30	97.27	0.55
5.5:2.5	94.52	93.37	97.42	0.49
6:2	94.51	94.50	97.93	0.58
7:1	94.85	94.44	97.17	0.66

3.2.2. Stability Experiments of the Fusion Model

Since the fault characteristics of vibration data and stationary data are quite different, in order to verify the stability of the model and avoid the interference of specific samples, this experiment uses the cross-validation method to validate the fusion model. Specifically, all samples are divided into five parts equally, one part is used as the training data, one part is used as the test data of the model, and the rest is the training data of the basic classifier.

As shown in Figure 14, we divided the fault data into five equal parts in order, categorized as 'A'–'E'. A total of five groups were divided, and each group was selected different test samples for five experiments. The histogram was obtained, as shown in Figure 15.

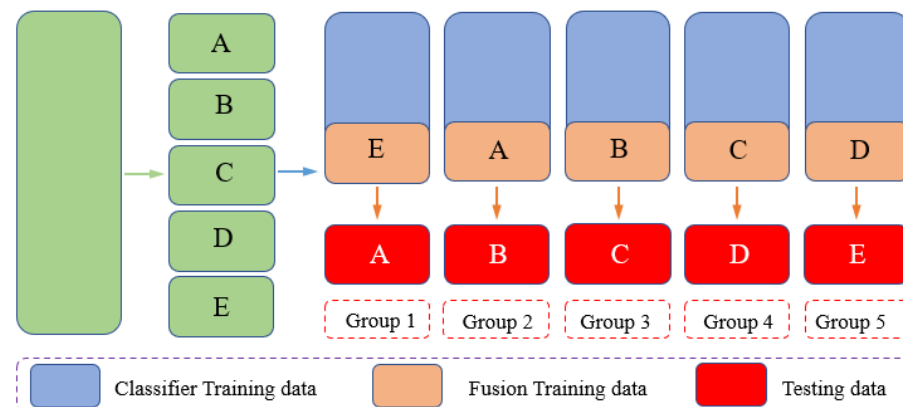


Figure 14. The diagram of cross-validation.

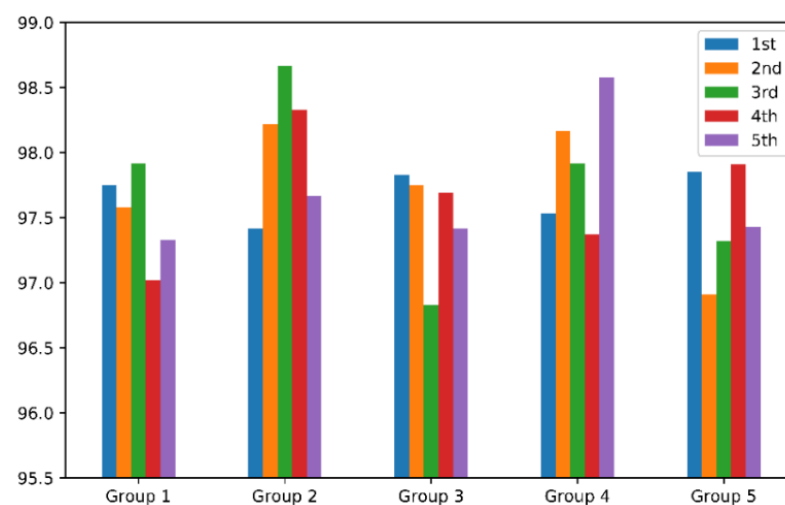


Figure 15. Bar chart of accuracy (%) with five groups cross-validation.

In Figure 15, five experiments were conducted in each group, and the average accuracy and standard deviation of each group were $97.52 \pm 0.35\%$; $97.73 \pm 0.53\%$; $97.5 \pm 0.40\%$; $97.91 \pm 0.49\%$; and $97.48 \pm 0.41\%$, respectively. The overall average accuracy was 97.63%, which was similar in each group; the standard deviation between groups was 0.19%, which was smaller than the standard deviation within each group. It indicates that the accuracy is little affected by the group and sample, and the model is relatively stable.

3.2.3. Comparisons with Other Methods

In this section, two sets of experiments are compared in order to reflect the performance of the proposed fusion algorithm. (1) Two basic classifiers are used as comparisons to demonstrate that the proposed fusion model can further improve diagnosis accuracy based on the basic classifier, (2) as compared with BP, SVM, optimized SVM, CNN, and random forest, which are popularly used as fault diagnosis algorithms. It proves that the fusion model proposed has higher accuracy.

1. Comparison with basic classifier.

As shown in Table 4, the proposed fusion model has a 3.22% improvement over FFT + SAE, and 2.31% improvement over WPD + NN, which indicates that the fusion model can further improve the diagnosis accuracy. The confusion matrices of the two basic classifiers are shown in Figures 16 and 17.

Table 4. The accuracy (%) table comparison with two basic classifiers.

Method	Accuracy	Std. Deviation
WPD + NN	95.62	0.29
FFT + SAE	94.14	0.34
Fusion module	97.93	0.58

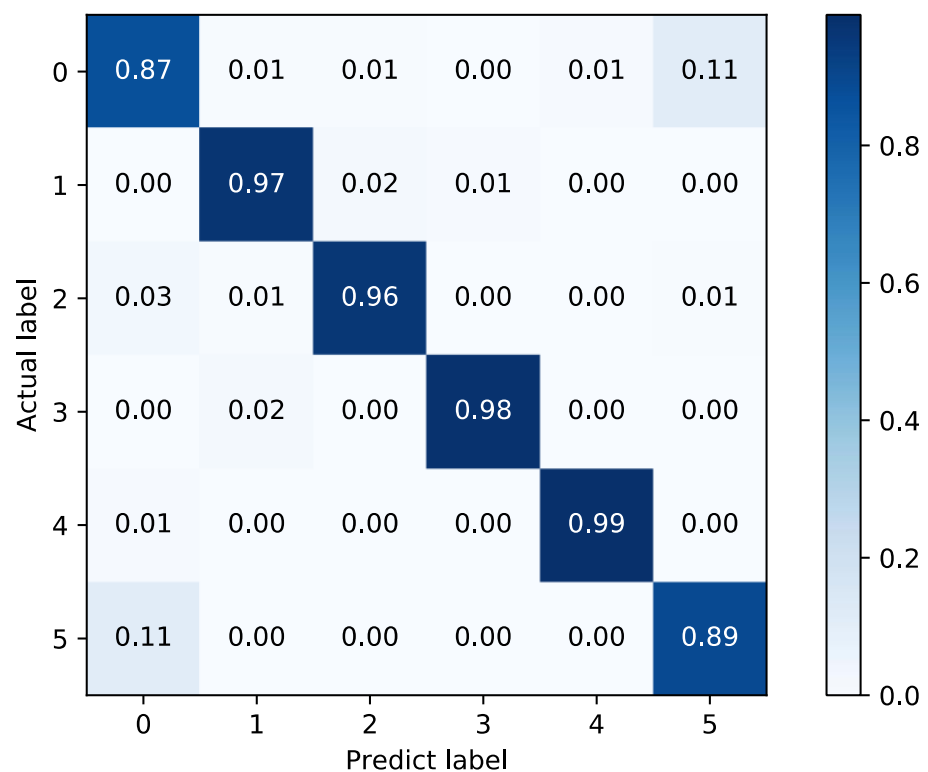


Figure 16. Confusion matrix of FFT + SAE.

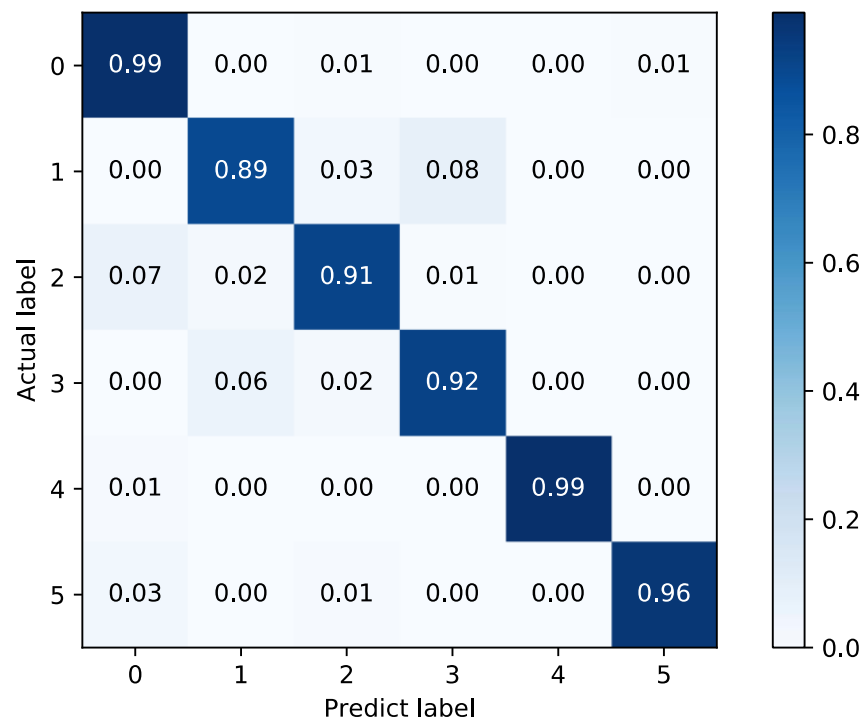


Figure 17. Confusion matrix of WPD + NN.

In Figures 16 and 17, it can be found that the FFT + SAE classifier has a larger misclassification rate in labels 0 and 5, while 1 and 3 are classified more accurately, which is exactly the opposite in the WPD + NN classifier. Therefore, the combination of these two complementary methods can be more effective for fault diagnosis, as shown in Figure 18.

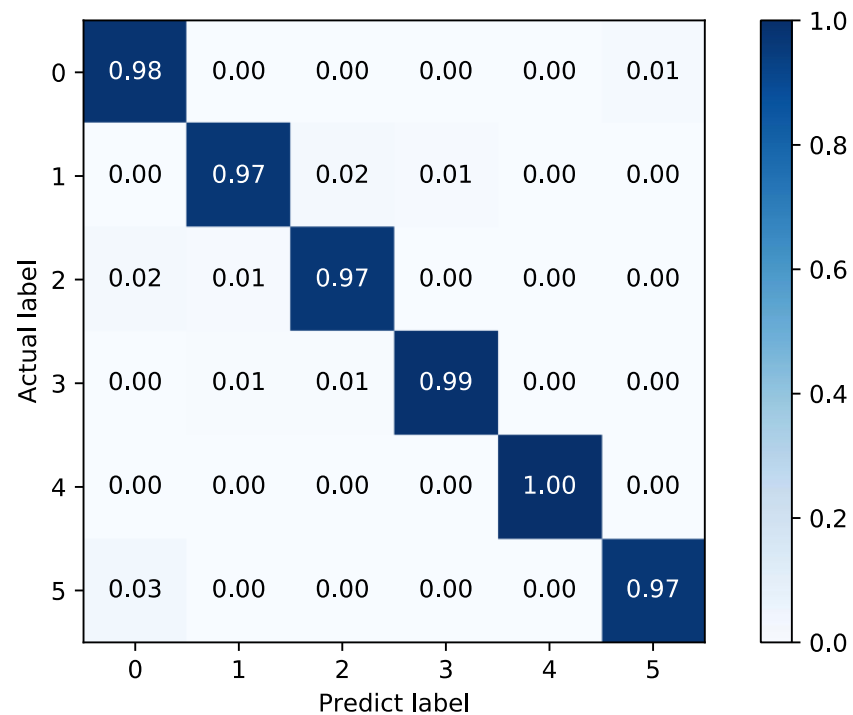


Figure 18. Confusion matrix of the fusion model.

2. Comparison with other methods.

To achieve a better comparison, we used the same test dataset (20%) for the proposed fusion model and other methods.

BP and SVM are commonly used intelligent algorithms in fault diagnosis. In this experiment, the Gaussian kernel function of SVM was used, and parameters C and gamma were set as default. BP was a three-layer network with node numbers 1000, 512, 64, and 6 used. The DFF network [24] extracted features through two layers of Conv, and the Conv 1 parameters were [4, 11, 2, 5, 2], representing the kernel number = 4, kernel size = 11, stride = 2, padding = 5, MaxPool = 2. The Conv 2 parameters were [8, 7, 1, 3, 2]. The node number of the FC layer was the same as that of BP. ISSVM [21] is an optimized SVM with an improved sparrow search algorithm. The C range was [1, 100], gamma ranges was [10⁻⁵, 0.1], optimization C = 57.75, and gamma = 0.0048. These two methods further improved the accuracy of diagnosis compared to the BP network and SVM methods, which proved the importance of feature extraction for the original signal. Further, the random forest (RF) method of integrated learning was considered, where the number of trees is 100, and each decision tree is an independent classifier. RF is a method that combines the diagnostic results of these trees. However, it only considers one model, namely the decision tree, therefore it is easy to over-fit, thereby leading to a lower accuracy. The fusion model proposed in this paper integrated these problems, adopted two different feature extraction methods, and selected the different network according to these feature vectors. The comparison results, shown in Table 5, revealed that the fusion model enjoys higher diagnosis accuracy than other methods.

Table 5. The accuracy (%) table comparison with other methods.

Method	Feature	Testing Accuracy
SVM [5]	Time Domain	68.67
BP [4]	Time Domain	74.58
DFF [24]	CNN	90.25
ISSVM [23]	WPD	92.25
RF [7]	Time Domain	93.58
Fusion model	/	97.93

4. Conclusions

In this paper, a fusion method was proposed for the fault diagnosis of FOG under vibration conditions. To solve inaccurate fault feature extraction caused by the vibration noise interference of FOG, the proposed method fuses a sparse auto encoder diagnostic model based on fast Fourier transform and a neural network diagnostic model based on wavelet packet decomposition. In the experiments, a total of 6000 sample fault datasets, obtained by simulating the output data under random vibration of FOG, were used for algorithm training and testing. The experimental results showed that the proposed method can diagnose the fault of FOG more accurately under vibration conditions compared with existing methods. Once the model is trained, it can be simply imported and reused in a short time. The time consumed for each step is shown in Appendix B Table A2. The method can be used in the shield construction environment, and significantly improve the reliability of the FOG measurement system.

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

Funding: This study was financially supported by the National Key RandD Program of China (Grant No. 2020YFB1713900).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

Appendix A

Table A1. GJB150.16A-2009 C.7 Vibration Environment Inflection.

Hz	g^2/Hz	Hz	g^2/Hz
5	0.0451	104	0.0033
6	0.0303	107	0.0044
7	0.0761	111	0.0032
13	0.0127	147	0.0029
15	0.0327	161	0.0052
16	0.0134	175	0.0022
21	0.0120	233	0.0013
23	0.0261	257	0.0027
25	0.0090	314	0.0016
26	0.0090	333	0.0053
30	0.0137	339	0.0009
34	0.0053	382	0.0017
36	0.0079	406	0.0008
46	0.0039	482	0.0019
50	0.0067	500	0.0007
55	0.0042		

1.29 grms

Appendix B

Table A2. The time required for the different sections of the fault diagnosis testing period.

Section	Time to First Reuse (ms)	Time to Normal Use (ms)
WPD	70.29	70.29
FFT	41.94	41.94
SAE	15.02	1.01
NN	6.96	0.99
Fusion model	6.98	0.99
Total	141.19	115.22
Time offset		1000.00

References

- Gao, Z.; Cecati, C.; Ding, S.X. A Survey of Fault Diagnosis and Fault-Tolerant Techniques—Part I: Fault Diagnosis with Model-Based and Signal-Based Approaches. *IEEE Trans. Ind. Electron.* **2015**, *62*, 3757–3767. [\[CrossRef\]](#)
- Shao, H.; Jiang, H.; Lin, Y.; Li, X. A novel method for intelligent fault diagnosis of rolling bearings using ensemble deep auto-encoders. *Mech. Syst. Signal Process.* **2018**, *102*, 278–297. [\[CrossRef\]](#)
- Zuo, L.; Zhang, L.; Zhang, Z.-H.; Luo, X.-L.; Liu, Y. A spiking neural network-based approach to bearing fault diagnosis. *J. Manuf. Syst.* **2021**, *61*, 714–724. [\[CrossRef\]](#)
- Li, G.; Deng, C.; Wu, J.; Chen, Z.; Xu, X. Rolling Bearing Fault Diagnosis Based on Wavelet Packet Transform and Convolutional Neural Network. *Appl. Sci.* **2020**, *10*, 770. [\[CrossRef\]](#)
- Fernández-Francos, D.; Martínez-Rego, D.; Fontenla-Romero, O.; Alonso-Betanzos, A. Automatic bearing fault diagnosis based on one-class ν -SVM. *Comput. Ind. Eng.* **2013**, *64*, 357–365. [\[CrossRef\]](#)
- Li, H.; Hu, G.; Li, J.; Zhou, M. Intelligent fault diagnosis for large-scale rotating machines using binarized deep neural networks and random forests. *IEEE Trans. Autom. Sci. Eng.* **2021**. [\[CrossRef\]](#)
- Chen, M.; Wu, J.; Wang, Y.; Jia, B.; Jiang, Y. Random forest based intelligent fault diagnosis for PV arrays using array voltage and string currents. *Sensors* **2018**, *18*, 1221. [\[CrossRef\]](#)

8. Ahmad, Z.; Nguyen, T.-K.; Ahmad, S.; Nguyen, C.D.; Kim, J.-M. Multistage Centrifugal Pump Fault Diagnosis Using Informative Ratio Principal Component Analysis. *Sensors* **2022**, *22*, 179. [[CrossRef](#)]
9. Nguyen, C.D.; Prosvirin, A.E.; Kim, C.H.; Kim, J.-M. Construction of a sensitive and speed invariant gearbox fault diagnosis model using an incorporated utilizing adaptive noise control and a stacked sparse autoencoder-based deep neural network. *Sensors* **2021**, *21*, 18. [[CrossRef](#)]
10. Li, W.; Huang, R.; Li, J.; Liao, Y.; Chen, Z.; He, G.; Yan, R.; Gryllias, K. A perspective survey on deep transfer learning for fault diagnosis in industrial scenarios: Theories, applications and challenges. *Mech. Syst. Signal Process.* **2022**, *167*, 108487. [[CrossRef](#)]
11. Wen, L.; Gao, L.; Li, X. A new deep transfer learning based on sparse auto-encoder for fault diagnosis. *IEEE Trans. Syst. Man Cybern. Syst.* **2019**, *49*, 136–144. [[CrossRef](#)]
12. Han, Y.; Tang, B.; Deng, L. Multi-level wavelet packet fusion in dynamic ensemble convolutional neural network for fault diagnosis. *Measurement* **2018**, *127*, 246–255. [[CrossRef](#)]
13. Tang, S.; Yuan, S.; Zhu, Y.; Li, G. An integrated deep learning method towards fault diagnosis of hydraulic axial piston pump. *Sensors* **2020**, *20*, 6576. [[CrossRef](#)]
14. Zhang, X.; Wan, S.; He, Y.; Wang, X.; Dou, L. Teager energy spectral kurtosis of wavelet packet transform and its application in locating the sound source of fault bearing of belt conveyor. *Measurement* **2021**, *173*, 108367. [[CrossRef](#)]
15. Sun, Y.; Cao, Y.; Li, P. Fault diagnosis for train plug door using weighted fractional wavelet packet decomposition energy entropy. *Accid. Anal. Prev.* **2022**, *166*, 106549. [[CrossRef](#)] [[PubMed](#)]
16. Chinara, S. Automatic classification methods for detecting drowsiness using wavelet packet transform extracted time-domain features from single-channel EEG signal. *J. Neurosci. Methods* **2021**, *347*, 108927.
17. Cai, J.; Wang, S.; Guo, W. Unsupervised embedded feature learning for deep clustering with stacked sparse auto-encoder. *Expert Syst. Appl.* **2021**, *186*, 115729. [[CrossRef](#)]
18. Lei, Y. Clustering algorithm-based fault diagnosis. In *Intelligent Fault Diagnosis and Remaining Useful Life Prediction of Rotating Machinery*; Butterworth-Heinemann: Oxford, UK, 2017; pp. 175–229.
19. Xiao, T. Research on Pose Measurement System of Shield Using the Combination of Inclinometer and Gyroscope. Master's Thesis, Huazhong University of Science and Technology, Wuhan, China, 2013.
20. Cai, H. Error Analysis and Compensation of Attitude Measurement System for TBM. Master's Thesis, Huazhong University of Science and Technology, Wuhan, China, 2009.
21. Yu, J.; Zhou, D.; He, P.; Huang, J. A new method for gyroscope fault diagnosis based on CGA RBFNN and multi-wavelet entropy. In Proceedings of the 2013 International Conference on Mechatronic Sciences, Electric Engineering and Computer (MEC), Shenyang, China, 20–22 December 2013; IEEE: New York, NY, USA, 2013; pp. 39–43.
22. Liu, Q.; Cheng, J.; Guo, W. Research on Gyro Fault Diagnosis Method Based on Wavelet Packet Decomposition and Multi-class Least Squares Support Vector Machine. In *Recent Trends in Intelligent Computing, Communication and Devices*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 789–797.
23. Chen, X.; Xiao, M.; Sun, Y. Fiber optic gyro fault diagnosis based on improved sparrow search algorithm and support vector machine. *J. Air Force Eng. Univ. Nat. Sci. Ed.* **2021**, *22*, 33–40.
24. Guan, J. *Research on Fault Diagnosis Technology of Fiber Optic Gyro Based on Neural Network*; Harbin Institute of Technology: Harbin, China, 2019.
25. Song, H.; Hu, S.-L.; Zhou, K.-Y. Review on Development of Fault Diagnosis for Gyroscope. In Proceedings of the ITM Web of Conferences, Birmingham, UK, 2–4 May 2017; EDP Sciences: Les Ulis, France, 2017; p. 07001.
26. Patton, R.J.; Uppal, F.J.; Simani, S.; Polle, B. Reliable fault diagnosis scheme for a spacecraft attitude control system. *Proc. Inst. Mech. Eng. Part O J. Risk Reliab.* **2008**, *222*, 139–152. [[CrossRef](#)]
27. Chamoun, J.N.; Digonnet, M. Noise and Bias Error Due to Polarization Coupling in a Fiber Optic Gyroscope. *J. Lightwave Technol.* **2015**, *33*, 2839–2847. [[CrossRef](#)]
28. Song, R.; Chen, X. Analysis of fiber optic gyroscope vibration error based on improved local mean decomposition and kernel principal component analysis. *Appl. Opt.* **2017**, *56*, 2265–2272. [[CrossRef](#)] [[PubMed](#)]
29. Vincent, P.; Larochelle, H.; Lajoie, I.; Bengio, Y.; Manzagol, P.-A.; Bottou, L. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *J. Mach. Learn. Res.* **2010**, *11*, 3371–3408.
30. Rifai, S.; Vincent, P.; Muller, X.; Glorot, X.; Bengio, Y. *Contractive Auto-Encoders: Explicit Invariance during Feature Extraction*; ICML: Broken Arrow, OK, USA, 2011.
31. Zhang, K.; Zhang, J.; Ma, X.; Yao, C.; Zhang, L.; Yang, Y.; Wang, J.; Yao, J.; Zhao, H. History matching of naturally fractured reservoirs using a deep sparse autoencoder. *SPE J.* **2021**, *26*, 1700–1721. [[CrossRef](#)]
32. Yang, Z.; Baraldi, P.; Zio, E. A method for fault detection in multi-component systems based on sparse autoencoder-based deep neural networks. *Reliab. Eng. Syst. Saf.* **2022**, *220*, 108278. [[CrossRef](#)]
33. Ranzato, M.A.; Poultney, C.; Chopra, S.; Cun, Y. Efficient learning of sparse representations with an energy-based model. In *Advances in Neural Information Processing Systems 19: Proceedings of the 2006 Conference*; MIT Press: Cambridge, MA, USA, 2006; Volume 19.
34. Mao, Z.; Xia, M.; Jiang, B.; Xu, D.; Shi, P. Incipient fault diagnosis for high-speed train traction systems via stacked generalization. *IEEE Trans. Cybern.* **2021**. [[CrossRef](#)]

35. Yuan, Z.; Song, N.; Pan, X. Fault diagnosis for space-borne fiber-optic gyroscopes using a hybrid method. *IEEE Access* **2020**, *8*, 194147. [[CrossRef](#)]
36. Wu, W. Research on Fault Diagnosis Method for Gyroscope Based on Fuzzy Support Vector Machines. Master's Thesis, Nanjing University of Aeronautics and Astronautics, Nanjing, China, 2016.
37. Heredia, G.; Ollero, A.; Mahtani, R.; Bejar, M.; Musial, M. Detection of Sensor Faults in Autonomous Helicopters. In Proceedings of the IEEE International Conference on Robotics & Automation, Barcelona, Spain, 18–22 April 2006.