

Article

Gear Fault Diagnosis Method Based on Multi-Sensor Information Fusion and VGG

Dongyue Huo ¹, Yuyun Kang ^{2,*} , Baiyang Wang ¹ , Guifang Feng ^{3,4,*}, Jiawei Zhang ⁵ and Hongrui Zhang ⁶

¹ School of Information Science and Engineering, Linyi University, Linyi 276000, China

² School of Logistics, Linyi University, Linyi 276000, China

³ School of Life Science, Linyi University, Linyi 276000, China

⁴ International College, Philippine Christian University, Manila 1004, Philippines

⁵ Linyi Trade Logistics Science and Technology Industry Research Institute, Linyi 276000, China

⁶ School of Mechanical and Vehicle Engineering, Linyi University, Linyi 276000, China

* Correspondence: kangyuyun@lyu.edu.cn (Y.K.); fengguifang@lyu.edu.cn (G.F.); Tel.: +86-155-8908-0373 (Y.K.)

Abstract: The gearbox is an important component in the mechanical transmission system and plays a key role in aerospace, wind power and other fields. Gear failure is one of the main causes of gearbox failure, and therefore it is very important to accurately diagnose the type of gear failure under different operating conditions. Aiming at the problem that it is difficult to effectively identify the fault types of gears using traditional methods under complex and changeable working conditions, a fault diagnosis method based on multi-sensor information fusion and Visual Geometry Group (VGG) is proposed. First, the power spectral density is calculated with the raw frequency domain signal collected by multiple sensors before being transformed into a power spectral density energy map after information fusion. Second, the obtained energy map is combined with VGG to obtain the fault diagnosis model of the gear. Finally, two datasets are used to verify the effectiveness and generalization ability of the method. The experimental results show that the accuracy of the method can reach 100% at most on both datasets.

Keywords: gear fault diagnosis; multi-sensor information fusion; VGG



Citation: Huo, D.; Kang, Y.; Wang, B.; Feng, G.; Zhang, J.; Zhang, H. Gear Fault Diagnosis Method Based on Multi-Sensor Information Fusion and VGG. *Entropy* **2022**, *24*, 1618. <https://doi.org/10.3390/e24111618>

Academic Editors: Claude Delpha and Demba Diallo

Received: 29 August 2022

Accepted: 1 November 2022

Published: 6 November 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

The gearbox is an important component in a mechanical transmission system and plays a key role in aerospace, wind power and other fields. Because of the complex structure of the gearbox and its continuous operation in harsh environments, it is easy to cause failures [1]. Gear failure is one of the main reasons for gearbox failure. Due to the complex working conditions of the gear and the easily affected tooth surface, its performance gradually degrades. When the gears are degraded to a certain extent, serious failures occur, resulting in huge economic losses and even major safety accidents [2,3]. Therefore, the intelligent diagnosis of gear faults is important research [4].

The traditional intelligent diagnosis method is mainly divided into three steps. First, the signal data of the sensor is collected when the machine is running. Second, feature extraction and selection are performed. Finally, the extracted features are fed into the model for training and fault classification. Many researchers have done a lot of research on gear fault diagnosis using traditional intelligent diagnosis methods [5–10]. Inturi et al. [11] extracted effective features of gear fault signals using DWT and EMD and input the extracted features into the SVM model for training. After completing several experiments, the accuracy rate obtained by combining SVM and DWT was better than that obtained by combining SVM and EMD. This is due to the problem of modal aliasing in EMD analysis compared to DWT. Therefore, the diagnostic accuracy obtained by the combination of DWT and SVM is higher. Wang et al. [12] first used the local weighted scatter smoothing method (LOWESS) to preprocess the signal spectrum, used WPT to extract signal features

second and finally used the least squares support vector machine optimized by adaptive particle swarm optimization to diagnose fault types. After simulation experiments, the fault diagnosis was accurate. In the data preprocessing stage, the combination of LOWESS and WPT reduced the interference of noise to a certain extent, while enhancing the ability of certain features in classification. In the classification and diagnosis stage, the ASPO algorithm was used to select and adjust the optimal value of the parameters, which achieved higher accuracy than the traditional SVM algorithm. The rate reached 98.33%. Ramtek et al. [13] used Flexible Analytic Wavelet Transform (FAWT) to decompose signal data into sub-band signals and used various entropies to extract features from all sub-band signals, then input these features into a LS-SVM classifier. Since the feature of the log energy entropy has properties better than other entropies, it provided the best diagnostic accuracy for signal classes. The results of simulation experiments showed that the accuracy of this method is higher than that of Continuous Wavelet Transform (CWT), Discrete Wavelet Transformation (DWT), Wavelet Packet Transform (WPT), Dual-Tree Complex Wavelet Transform (DTCWT) and Tunable Q-factor Wavelet Transform (TQWT). Vos et al. [14] combined LSTM and SVM for diagnosing gear faults. Two-step LSTM and SVM were used for bearing fault diagnosis. Simulation experiments showed that the proposed method is more reliable and effective.

With the continuous development of modern industry, the complexity of machinery and equipment continues to increase, and the operating conditions of many gearboxes are usually variable operating conditions of speed and load. At the same time, the data generated by mechanical equipment have also grown exponentially, and mechanical fault diagnosis has entered the era of big data. Although traditional machine learning methods such as SVM have good performance on small sample datasets, these methods require a lot of expertise and domain knowledge, and the feature engineering steps need to be done manually. Since deep learning does not require manual feature extraction, deep learning has been applied to gear fault diagnosis by scholars [15–20]. Nguyen et al. [21] used ANC and DNN to build a sensitive gear fault diagnosis model that can effectively and accurately diagnose the types of faults with different shaft speeds and have high classification accuracy. Cui et al. [22] proposed a semi-supervised fault diagnosis method, which first extracted features from vibration signals through unsupervised training, then used a small amount of labeled data for supervised fine-tuning for fault diagnosis. Accuracies of 99.42% and 91.97% can be achieved on the bearing dataset and gear datasets, respectively. Cao et al. [23] proposed an unsupervised domain-shared convolutional neural network for efficient fault diagnosis of machines from stable to time-varying speeds. The method analyzed the original vibration signal data of bearings and gears at different speeds. As demonstrated by their experiments, the method has higher diagnostic accuracy, faster convergence speed and better robustness than other methods. Wu et al. [24] adopted a transfer learning method based on A2ResNet-aCoral to solve the problem of small sample datasets. After experimental verification, the accuracy of this method on a small sample set is 92.71%.

At present, the traditional gear fault diagnosis mainly focuses on the signal of a single sensor [25–27]. Dong et al. [28] used adversarial denoising autoencoders as feature extractors and introduced joint ensemble and statistical alignment methods to process the deep features of samples to reduce the geometric and statistical differences between the source and target domains. This improved the accuracy of the fault diagnosis. Hasan et al. [29] proposed a fault diagnosis method based on acoustic emission signal-based acoustic spectrum imaging (ASI) combined with transfer learning. ASI converts the spectral component amplitude of the windowed time-domain acoustic emission signal into a spectrogram and uses the image to represent the characteristics of the acoustic emission signal. It then uses CNN for feature extraction for diagnosis. The accuracy of this method can reach 94.67%. Li et al. [30] combined the autoencoder and the aggregated neural network to build a fault diagnosis model and used four loss functions to extract the features of the fault information, finally achieving an accuracy of 94.89%. Although the research using only a single sensor

has the advantages of simplicity and speed, the fault information it contains is limited, and it has poor performance in fault tolerance.

In order to solve the problem of lower diagnostic accuracy under variable working conditions for the signal characteristics of a single sensor, information fusion and VGG are used to solve the problem of gear fault diagnosis. A gear fault diagnosis method based on multi-sensor information fusion and VGG is proposed. The main contributions of the paper are as follows:

1. The signal drawn directly using the plot function is no longer used as the input of the neural network. Instead, the power spectrum density energy diagrams obtained after power spectral density analysis of the original signal is used as the input of the neural network.
2. In order to obtain richer information and features, it is no longer only using the data collected by a single sensor for training and diagnosis. Instead, it uses data signals from multiple sensors, which are fused together and used for training and diagnostics.
3. In order to improve the accuracy and generalization ability of fault diagnosis, instead of using traditional machine learning methods, the obtained power spectrum density energy diagrams are combined with VGG16. Finally, its accuracy and generalization ability are verified on two datasets.

2. Methods

2.1. Convolutional Neural Network (CNN)

CNN is a deep neural network with a convolutional structure. It has powerful automatic feature extraction capabilities. The input layer, convolution layer, activation layer, pooling layer and fully connected layer constitute CNN. Convolution, activation and pooling operations can be reused. The convolution layer realizes the extraction of data features before the pooling layer reduces the dimensionality of the data and combines the extracted features through the fully connected layer to achieve classification.

The process of extracting features in the convolutional layer is the process of sliding the convolution kernel according to the specified step size and performing the convolution operation. The calculation process of the convolution operation is shown by Formula (1):

$$X_{i+1} = W_i \otimes X_i + b_i, \quad (1)$$

where X_i is the images input to the convolutional layer. X_{i+1} represents the image after the convolution operation. The symbol $\otimes$ represents the convolution operation. W_i represents the weight of the convolution kernel, and b_i is the bias.

The activation layer is used to nonlinearly transform the feature map after the convolution operation through the activation function to improve the expression ability of the model [31]. The mathematical expression after adding the activation function is shown in Formula (2):

$$y_i = f(x_{i+1}) = f(W_i \otimes X_i + b_i), \quad (2)$$

where y_i represents the feature map after the output of the activation layer. $f(*)$ represents the activation function, and the selection of different activation functions has different effects on the performance of the model.

The main purpose of the pooling layer is to reduce the dimensionality of the data, and the most used in CNN is Max Pooling. Its mathematical model is shown in Formula (3):

$$f = \text{Max}(x_{m,n}, x_{m+1,n}, x_{m,n+1}, x_{m+1,n+1}) \quad (3)$$

$$(1 \leq m \leq M, 1 \leq n \leq N).$$

The fully connected layer synthesizes multiple features to be extracted. Its mathematical expression is shown in Formula (4):

$$h_{\theta}(x_i) = \begin{bmatrix} p(y_i = 1|x_i; \theta) \\ p(y_i = 2|x_i; \theta) \\ \vdots \\ p(y_i = k|x_i; \theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^k e^{\theta_j^T x_i}} \begin{bmatrix} e^{\theta_1^T x_i} \\ e^{\theta_2^T x_i} \\ \vdots \\ e^{\theta_k^T x_i} \end{bmatrix}, \quad (4)$$

where $\theta_1, \theta_2, \dots, \theta_k$ are the learning parameters of the model, multiplied by $\frac{1}{\sum_{j=1}^k e^{\theta_j^T x_i}}$ to make the probability distribution between [0,1].

The VGG network proposed by the Visual Geometry Group of Oxford University is one of the current mainstream models of CNN [32]. To reduce the required parameters, the VGG network abandons the use of large-scale convolution kernels but uses multiple 3×3 convolution kernels. Under the premise of the same receptive field, the 5×5 convolution kernel is replaced by two 3×3 convolution kernels. The 7×7 convolution kernel is replaced by three 3×3 convolution kernels. The size of the pooling kernel is 2×2 . Through two fully connected layers with 4000 nodes and an activation function, a $1 \times 1 \times 4096$ vector is obtained. Then, through a fully connected layer of S (S is the number of categories) nodes, a $1 \times 1 \times S$ vector is obtained. Finally, the one-dimensional vector obtained through the fully connected layer is input to the softmax activation function, and the prediction result is converted into a probability distribution. Figure 1 shows the structure of VGG-16. The expression of the adopted ReLU activation function is shown in Formula (5). Therefore, VGG-16 has the advantages of simple structure, small convolution kernel, small pooling kernel, deeper layers and wider feature maps. At the same time, VGG-16 uses the ReLU function instead of Tanh and Sigmoid as the activation function because it is an unsaturated function, which can reduce the error of back propagation, and the speed of network convergence will be faster. This can greatly reduce training time.

$$f(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (5)$$

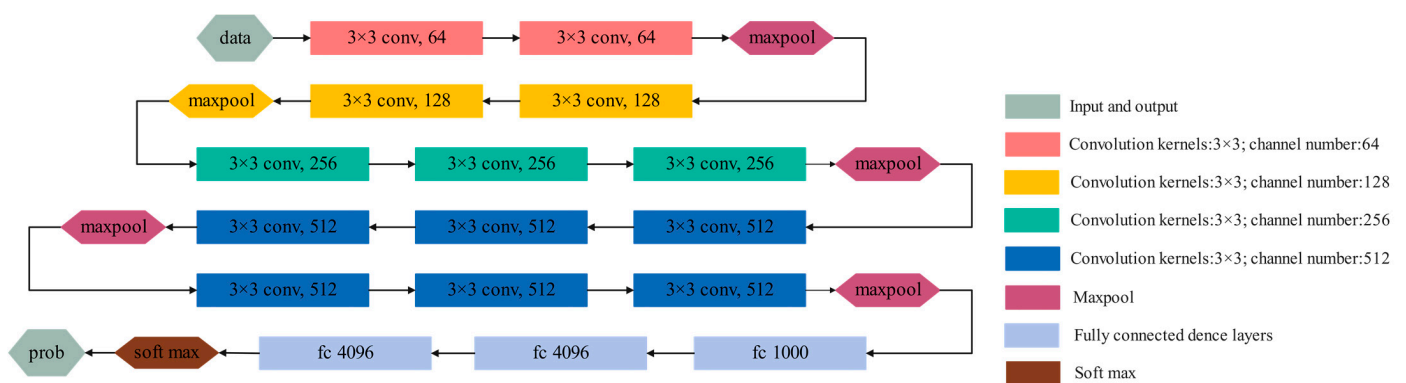


Figure 1. Network structure of VGG-16.

VGG-16 was originally used for projects with 1000 categories, that is, $N = 1000$, because the fault category of gear fault diagnosis is much less than 1000. To improve the efficiency of the model, the number of nodes in the first two fully connected layers is halved. In other words, 4096 nodes of the first two fully connected layers are modified to 2048. The number of nodes of the last fully connected layer is consistent with the number of failure categories of the dataset.

2.2. Power Spectrum Density Energy Diagrams (PSDED)

Power spectral density (PSD) is a method to calculate the power of each frequency band of a signal. The definition of the PSD of a signal $x(n)$ is shown in Formula (6). The specific process of calculating the PSD of the signal is: first perform fast Fourier transform (FFT) on the signal $u(n) = 0, 1, \dots, N - 1$, and then pass the spectral power of each frequency band through the absolute value form table out.

$$P(\omega) = \sum_{k=-\infty}^{+\infty} r(k) e^{-i\omega k} \quad (6)$$

The formula for discrete Fourier transform is shown in Formula (7). Among them, $W_N = e^{-j2\pi/N}$, $k = 0, 1, \dots, N - 1$. Both W_N and $U(k)$ have periodicity, and their period size is N .

$$U(k) = \sum_{n=0}^{N-1} u(n) W_N^{kn} \quad (7)$$

The effect diagram obtained after PSD analysis is shown in Figure 2, where (a) is the PSDED with 8 channels and (b) is the PSDED with 4 channels.

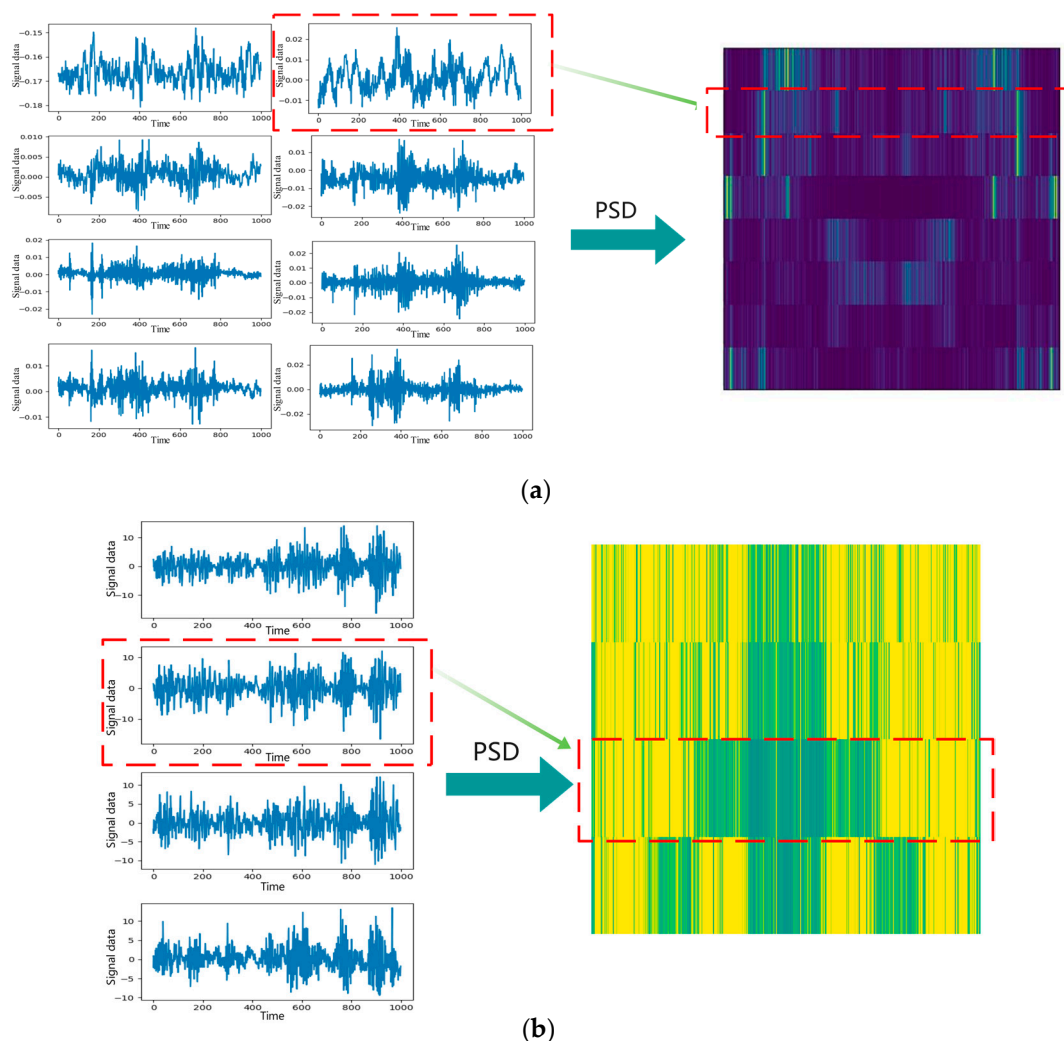


Figure 2. Power Spectrum Density Energy Diagrams: (a) 8 channels; (b) 4 channels.

2.3. Model Training and Diagnosis Process

Figure 3 shows a flowchart of the method. The process consists of three steps: data preprocessing, model training, and model diagnosis. The specific operation steps are as follows.

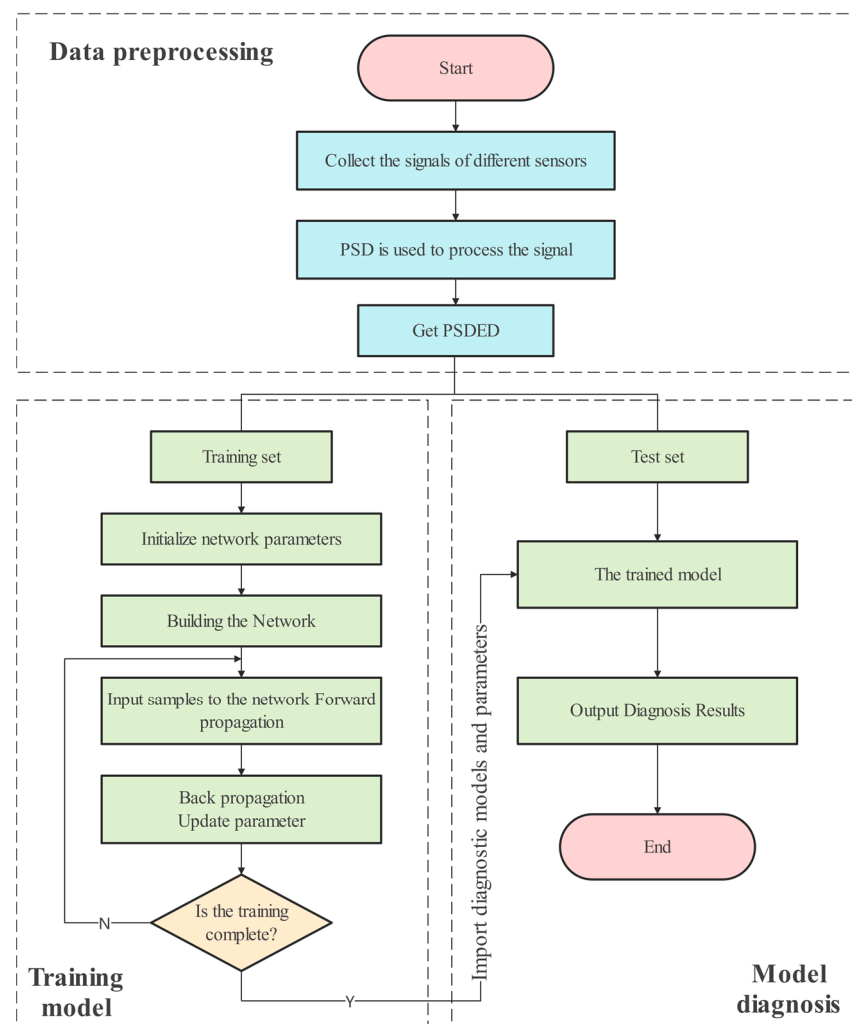


Figure 3. Flow chart of gear fault diagnosis.

Step one: Data preprocessing. The PSD is calculated for each sensor vibration signal collected, and after PSD analysis, the information of multiple sensors is fused to obtain a two-dimensional picture, namely PSEDED. A PSEDED is a sample. The samples are divided into multiple data sets according to the working conditions, and the data sets are divided into training sets and test sets according to a certain proportion.

Step two: Train the model. Build the VGG model and initialize parameters. Input the divided dataset into the built model and train it. Finally, the trained model is obtained.

Step three: Model diagnosis. Test samples are fed into the trained model. The model will perform adaptive feature extraction and pattern recognition on the gear fault data. The final output of the diagnosis results.

3. Dataset

Two public gear failure datasets are used to verify the effectiveness and generalization ability of the proposed method. Dataset 1 is a five-category dataset of 8 sensors from Southeast University. Dataset 2 is a binary public dataset with four sensor data.

3.1. Dataset 1 from Southeast University

Dataset 1 is a gearbox fault diagnosis dataset from Southeast University. These data were collected from Drivetrain Dynamic Simulator. The experimental setup of the gearbox dataset is shown in Figure 4 [33]. The dataset consists of a bearing dataset and a gear dataset. This paper selects the gear fault sub-dataset in this dataset. The data of the dataset

were collected under two working conditions with the rotating speed system loads of “20hz and 0v” and “30hz and 2v”. The data set collected data of five types: health, crack at the tooth foot, lack of one tooth in the gear, crack at the root of the tooth foot and wear on the tooth surface. Each type of data is composed of signals collected by eight sensors. One means motor vibration. Two, three and four represent the vibration of the planetary gearbox in the three directions of X, Y and Z, respectively. Five represents the motor torque. Six and seven denote parallel gearboxes, respectively. Eight represents the three directions of vibration X, Y and Z. Table 1 describes the category information of the dataset.

Table 1. Description of the type in the dataset 1.

Type	Description
Chipped	Cracks in the gear feet
Health	Healthy gear
Miss	Missing one of feet in the gear
Root	Cracks at the root of the tooth
Surface	Wear on gear surfaces

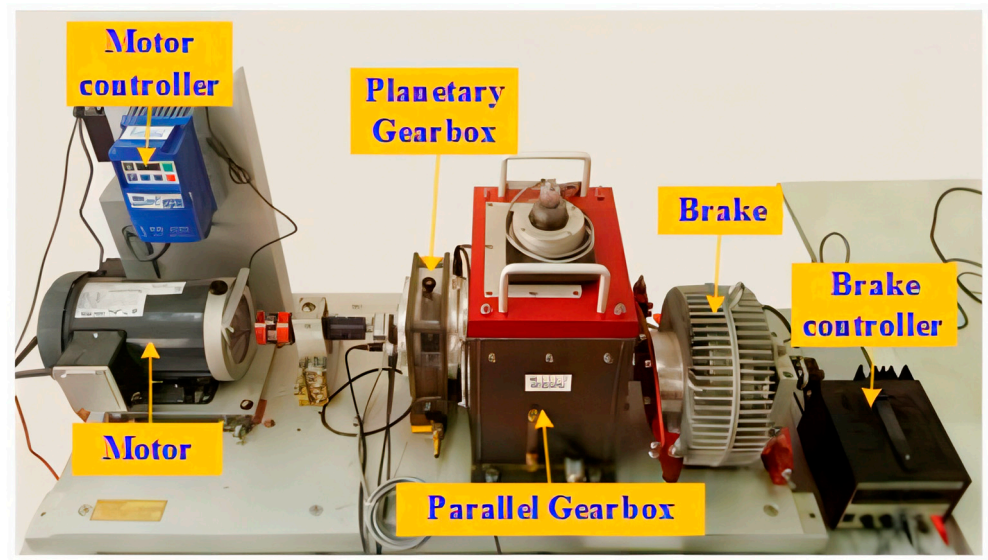


Figure 4. Experimental setup for gearbox dataset. Reprinted with permission from Ref. [33]. 2018, Siyu Shao.

The raw data collected under each condition ($C = ['20\text{hz and } 0\text{v}', '30\text{hz and } 2\text{v}']$) and under each type were divided into $D = 2000$ segments. The dataset is treated as a five-category ($T = ['\text{Health}', '\text{Chipped}', '\text{Miss}', '\text{Root}', '\text{Surface}']$) fault diagnosis under two conditions. There are 2000 sample points for each category, and the total number of sample points is shown in Formula (8).

$$D_{Total} = \text{len}(C) \times \text{len}(T) \times D \quad (8)$$

3.2. Dataset 2 from a Public Dataset

Dataset 2 is a gear fault diagnosis dataset disclosed in [34]. The dataset includes vibration measurements on healthy and damaged gearboxes under various loads (0–90%, in 10% steps) and a constant rotational speed of 30 Hz. Measurements were recorded with Spectra Quest’s Transmission Troubleshooting Simulator. The dataset is divided into two categories: health and damage. The signal data in each case were composed of arrays of signals measured separately by four sensors placed in different locations. The type description of this dataset is shown in Table 2.

Table 2. Description of the type in the dataset 2.

Type	Description
Broken	Broken gear
Health	Healthy gear

The raw data collected under each condition ($C = [\text{'Health'}, \text{'Broken'}]$) and each load ($F = [0', 10', \dots, 90']$) were divided into $D = 880$; each segment has $MD = 1000$ data points, and each segment contains $S = 4$ time series measured by sensors placed at different locations. The total number of segments is as shown in Formula (9).

$$D_{Total} = len(C) \times len(F) \times D \quad (9)$$

4. Experiment and Analysis

Several experiments were conducted on a Lenovo PC with Windows 10 64-bit operating system, Nvidia GeForce RTX 3060 GPU and 16 GB RAM. The operating environment of the experiment was: python3.9, Keras deep learning library implementation, cuda computing platform.

Each dataset was divided into a training set and a test set according to the ratio of 1:9. The batch size was set to 32. The number of iterations for training was set to 50. The learning rate was set to 0.0001.

4.1. Evaluation Indicators

Deep learning evaluation indicators were used to evaluate the performance of the deep learning model and were an important basis for model design. It is very limiting for evaluating models to count only the number of correct or incorrect predictions. Therefore, in order to verify the effectiveness of the proposed method from multiple angles and aspects, three evaluation indicators were used: loss function and accuracy, confusion matrix and t-distributed stochastic neighbor embedding (t-SNE).

4.1.1. Loss Function and Accuracy

Cross-entropy describes the distance between two probability distributions. The smaller the cross-entropy, the closer the two probability distributions are. The cross-entropy function is defined as Formula (10), where n is the number of samples, m represents the number of categories, y_{ic} is the label value and p_{ic} is the predicted value.

$$Loss = \frac{1}{n} \sum_i \sum_{c=1}^m y_{ic} \log(p_{ic}) \quad (10)$$

In order to define the evaluation index accuracy (*Accuracy*), four concepts of True Positive, False Positive, True Negative and False Negative were proposed. *TP* (True Positive): The predicted value of the sample is consistent with the true value, and both are positive. *FP* (False Positive): The predicted value of the sample is positive, and the true value is negative. *FN* (False Negative): The predicted value of the sample is negative, and the true value is positive. *TN* (True Negative): The predicted value of the sample is consistent with the true value, and both are negative. The calculation formula of Accuracy is shown in Formula (11).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (11)$$

4.1.2. Confusion Matrix

A confusion matrix is a table of data used to understand the performance of a classification model. It helps us understand how the test data are divided into different classes. Each column of the confusion matrix represents the predicted class, and the total in each column represents the amount of data predicted for that class. Each row of the confusion matrix

represents the true category of the data, and the total amount of data in each row represents the number of data instances of that category. The values in each column represent the amount of true data predicted to be in that category. The structure of the confusion matrix is shown in Figure 5.

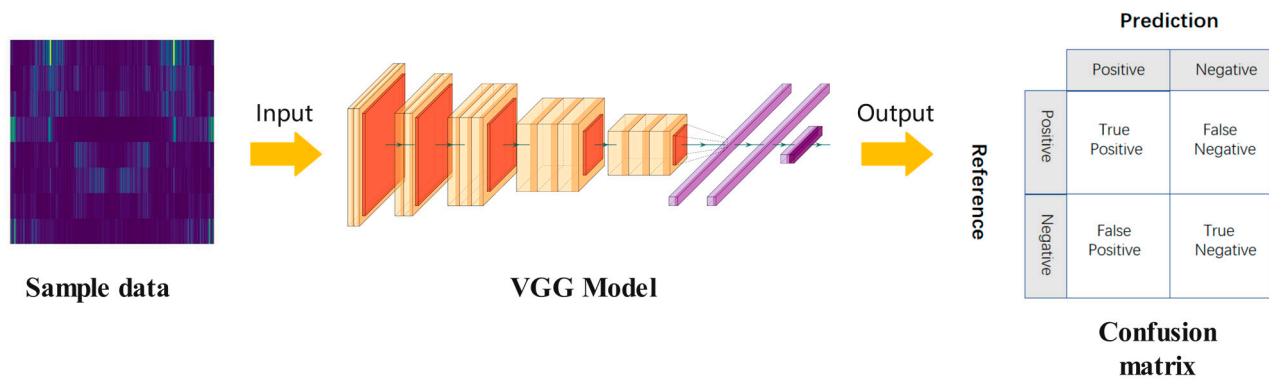


Figure 5. Structure of the confusion matrix.

4.1.3. t-SNE

t-SNE is a common clustering method used for dimensionality reduction and cluster visualization of numerical variables. The main idea is to use the conditional probability to represent the similarity of the distance of high-dimensional distribution points. A cluster generated by clustering is a collection of data objects that are similar to objects in the same cluster and different from objects in other clusters. t-SNE mainly consists of two steps:

Step 1: a probability distribution between high-dimensional objects is constructed by t-SNE, so that similar objects have a higher probability of being selected, and dissimilar objects have a lower probability of being selected. Step 2: the probability distribution of these points in the low-dimensional space is constructed by t-SNE, so that the two probability distributions are as similar as possible.

4.2. Results and Analysis of Dataset 1

Five types of data were collected under two working conditions in the dataset. The power spectral density of all collected signals were calculated and analyzed, and the power spectral density energy map (PSDED) was obtained. The dataset was divided into a training set and a test set by a 9:1 ratio. The sample data are shown in Table 3.

Table 3. Number of samples for Dataset 1.

Label	0	1	2	3	4
	Chipped	Health	Miss	Root	Surface
20hz_0v	1800	1800	1800	1800	1800
	200	200	200	200	200
30hz_2v	1800	1800	1800	1800	1800
	200	200	200	200	200

The proposed method was used to diagnose gear faults in the case of 20hz0v and 30hz2v. The loss value and accuracy obtained from the training set and test set in the two cases are shown in Figures 6 and 7. It can be seen from the figure that the proposed method achieved low loss value and high accuracy on both the training set and the test set under two variable working conditions. At the same time, it also had a good performance in the convergence speed. The loss value and accuracy obtained on the training set and test set are shown in Table 4. The proposed method achieved 100% accuracy under both variable conditions. Under the condition of 20hz0v, the loss of 1.54×10^{-7} and the accuracy of 100% were obtained on the training set, and the loss of 3.6×10^{-10} and the accuracy of 100%

were obtained on the validation set. Under the condition of 30hz2v, a loss of 1.75×10^{-7} and an accuracy of 100% were achieved on the training set, and a loss of 0.00019 and an accuracy of 100% were achieved on the validation set.

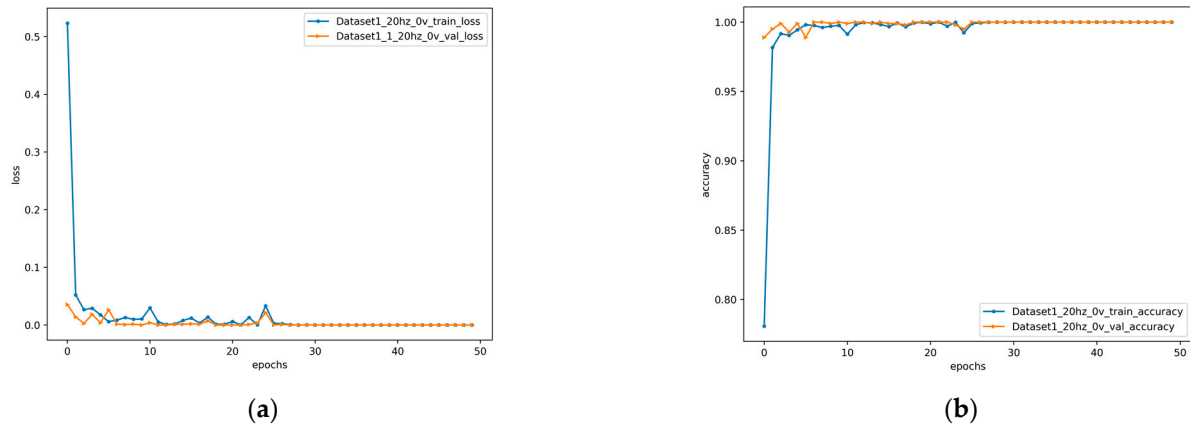


Figure 6. Loss and accuracy of 20hz0v: (a) loss; (b) accuracy.

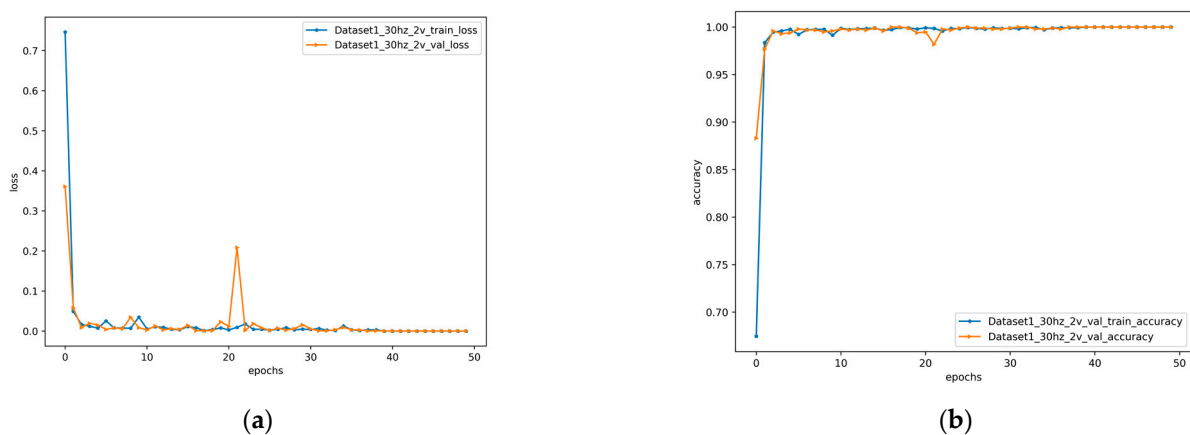


Figure 7. Loss and accuracy of 30hz2v: (a) loss; (b) accuracy.

Table 4. Loss and accuracy of 20hz0v and 30hz2v.

	Train Loss	Train Acc	Val Loss	Val Acc
20hz_0v	1.54×10^{-7}	1.0	3.6×10^{-10}	1.0
30hz_2v	1.75×10^{-7}	1.0	0.00019	1.0

In order to more intuitively show the effect of the proposed method in the fault diagnosis of variable working conditions, the confusion matrix and t-SNE of 20hz0v and 30hz2v are shown in Figures 8 and 9, respectively. It can be seen from Figure 8 that under the working condition of 20hz0v, the accuracy of identifying the health status, Crack occurrence in the root of the gear feet and Wear occurrence in the surface of the gear reached 100%. The misclassification of the proposed method was mainly to identify 6% of Crack occurrence in the gear feet as Wear occurs in the surface of the gear and 1% of Missing one foot in the gear as Wear occurs in the surface of the gear. Correspondingly, it can be seen in the cluster diagram in Figure 8b that some points belonging to the 0th and 2nd categories were identified in the 4th category. It can be seen from Figure 10 that under the working condition of 30hz2v, the accuracy of identifying the health status, missing one foot in the gear and wear occurrence in the surface of gear reached 100%. The misclassification of the proposed method was mainly to identify 1% of Crack occurrences in the gear feet

as Wear occurs in the surface of the gear, and 1% of Crack occurrences in the root of the gear feet as Wear occurs in the surface of the gear. Correspondingly, it can be seen in the cluster diagram in Figure 10b that some points belonging to the 0th and 3rd categories were identified in the 4th category.

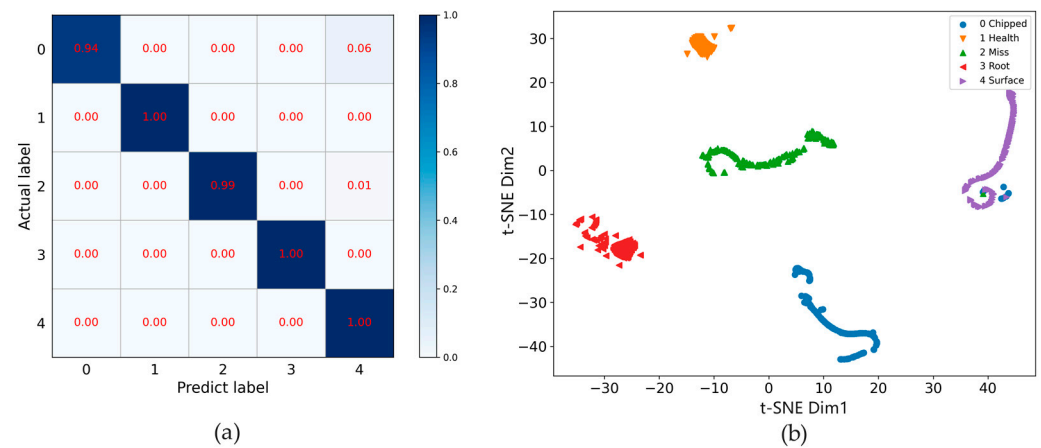


Figure 8. Confusion matrix and t-SNE of 20hz0v: (a) Confusion matrix; (b) t-SNE.

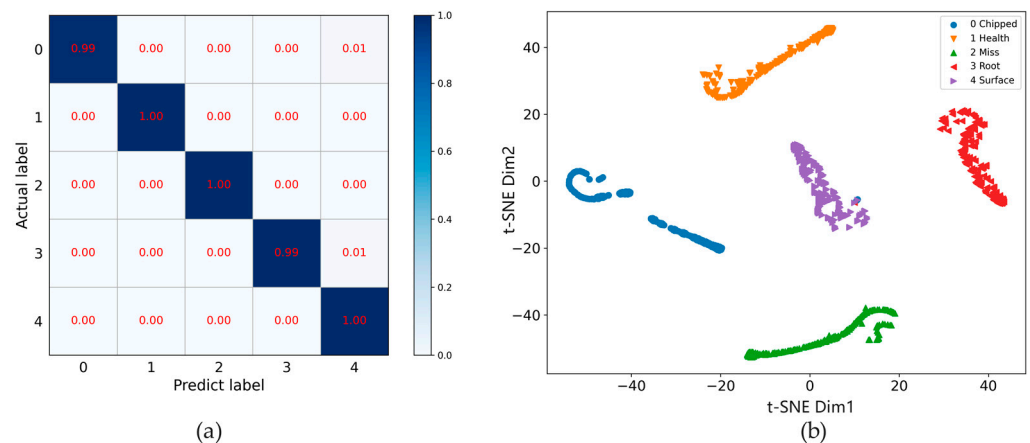


Figure 9. Confusion matrix and t-SNE of 30hz2v: (a) Confusion matrix; (b) t-SNE.

In order to verify that multi-sensor information fusion can reflect the signal characteristics of gear health and damage more comprehensively and completely, each sensor used the same method to convert the signal into PSD and input it as a sample to VGG for training. The accuracies of each sensor and multiple sensors are shown in Table 5. In the case of 20hz0v, the accuracy rate of using multi-sensor information fusion was as high as 100%, while the accuracy rate of using a single sensor varied from 95.6% to 99.6%, with an average accuracy rate of 98%. In the case of 30hz2v, the accuracy rate of using multi-sensor information fusion was 100%, the accuracy rate of using a single sensor was between 96.7% and 92.4% and the average accuracy rate was only 94.7%. Therefore, the use of multi-sensor information fusion can provide more comprehensive and accurate information for gear fault diagnosis, thereby improving the accuracy of fault identification.

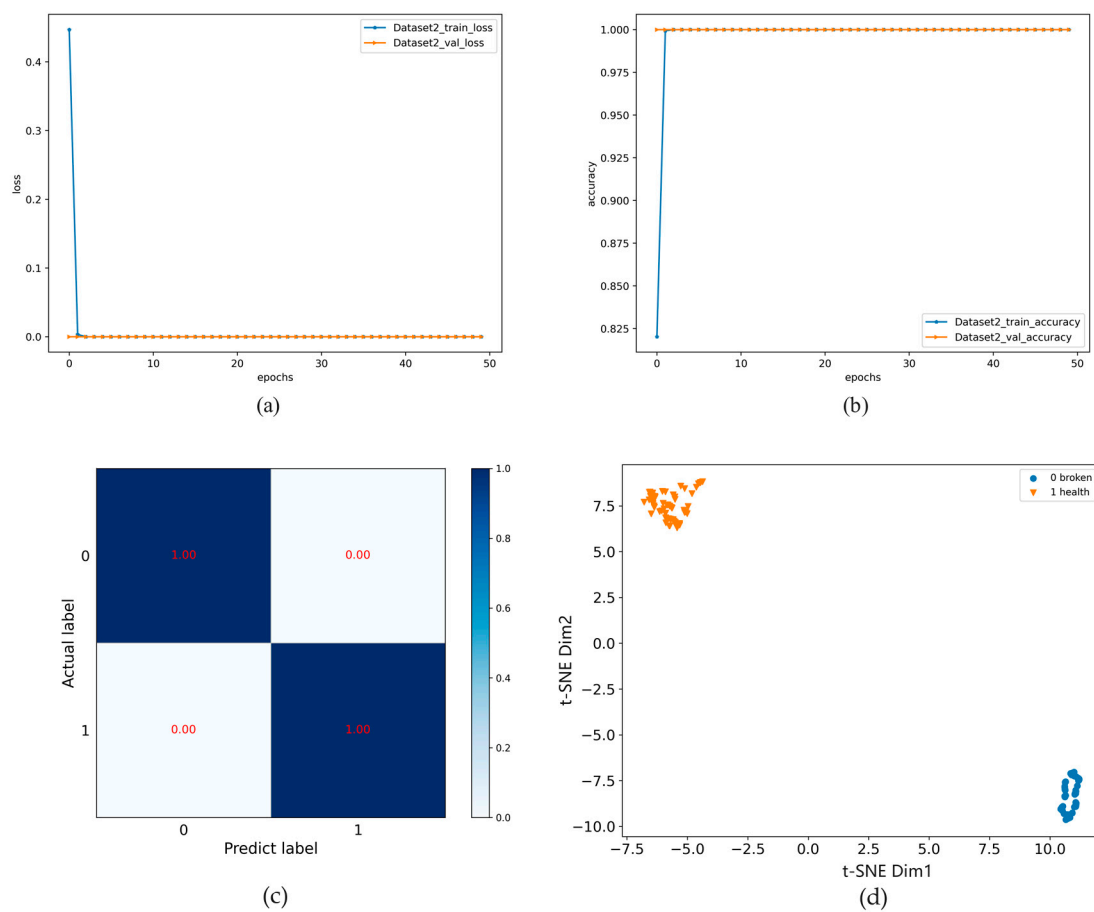


Figure 10. Results of Dataset 2: (a) loss; (b) accuracy; (c) Confusion matrix; (d) t-SNE.

Table 5. Accuracy of single-channel and multi-channel.

Channel	20hz_0v	30hz_2v
Sensor 1	99.4%	94.4%
Sensor 2	99.6%	95.9%
Sensor 3	95.9%	96.7%
Sensor 4	98.4%	96.2%
Sensor 5	98.1%	93.9%
Sensor 6	98.6%	96.6%
Sensor 7	95.6%	92.6%
Sensor 8	98.4%	92.4%
Fusion from sensor 1 to sensor 8	100%	100%

To verify the effectiveness of the proposed method, a comparison between the experimental results and other papers is shown in Table 6. As can be seen from Table 6, the accuracy obtained by using the model multi-label recurrent translation adversarial network (MCTAN) [35] based on the multimodal neural network was about 68%. The accuracy obtained by using the multimodal neural-network-based model [36] was 97.83%. The accuracy obtained by using the wavelet transform and VGG16 [33] was 99.64%. This clearly shows that the proposed method can achieve better results.

Table 6. Comparison with results from different papers using dataset 1.

Methods	Accuracy
MCTAN [35]	68%
multimodal neural-network-based model [36]	97.83%
wavelet transform and VGG16 [33]	99.64%
Proposed method	100%

4.3. Results and Analysis of Dataset 2

The dataset was divided into healthy and damaged categories under various loads and a constant rotational speed of 30 Hz. The power spectral density of the collected signal was calculated and analyzed to obtain a power spectral density energy map (PSDED). The dataset was divided into a training set and a test set in a ratio of 9:1. The sample data is shown in Table 7.

Table 7. Number of samples for Dataset 2.

Label	0 Broken	1 Health
Train	792	792
Test	88	88

VGG neural network was used to train 50 epochs, and the training results are shown in Figure 10. As shown in Figure 10a,b, the loss value and accuracy of the model on the test set were 0 and 100%, respectively. The confusion matrix and t-SNE are shown in Figure 10c,d. The results show that the proposed method performed well on this dataset, reaching almost 100% accuracy. Therefore, the proposed method still has good generalization ability in the fault diagnosis of gears with different damage categories in different datasets.

To verify the effectiveness of the proposed method, the experimental results compared to other methods using dataset 2 are shown in Table 8. As seen in Table 8, the accuracies of the method “Long Short Term Memory (LSTM) and Particle Swarm Optimization (PSO)” [37] and the method “LSTM and Cuckoo” [37] were both 87.5%, and the accuracy of the method “Empirical Mode Decomposition (EMD) and Artificial Neural Network (ANN)” [34] was about 73%. It is clearly shown that the proposed method can achieve better results.

Table 8. Results comparison of different methods about dataset 2.

Methods	Accuracy
EMD and ANN [34]	73%
“LSTM + PSO” and “LSTM + Cuckoo” [37]	87.5%
Proposed method	100%

5. Conclusions

Aiming at the problem that it is difficult to effectively identify the fault types of gears with traditional methods under complex and changeable conditions, a fault diagnosis method based on information fusion and a convolutional neural network is proposed. First, the data power spectral density was calculated on the raw signals collected by multiple sensors to obtain the frequency domain signal, which was converted to a power spectral density energy map after information fusion and PSD analysis. Second, the obtained energy map was input to VGG16 as a sample for training and diagnosis. Then two datasets were used to verify the effectiveness and generalization ability of the proposed method. For dataset 1, the diagnostic accuracy under the two working conditions was obtained, which reached 100%. Then the accuracies of single-sensor and multiple-sensors were compared. Under the two working conditions, the accuracy of multi-sensor information

fusion was 2% and 5.3% higher than that of a single-sensor, respectively. For dataset 2, its fault diagnosis rate was as high as 100%. Therefore, the fusion of the data of multiple sensors can provide more comprehensive and accurate information for gear fault diagnosis, and the combination with VGG16 can greatly improve the accuracy of fault identification. In future work, more advanced classifiers will be considered for gear fault diagnosis to improve its accuracy and robustness.

Author Contributions: D.H. contributed to the methodology; B.W., H.Z. and J.Z. contributed to the experiment; Y.K. and G.F. contributed to the conceptualization. All authors have read and agreed to the published version of the manuscript.

Funding: Shandong Province Higher Educational Science and Technology Program (No. J18KB018), Shandong Province Innovation and entrepreneurship training program for college students (No. S202210452042) and Research and application of key technologies of cargo distribution and transportation based on BDS and 5G (No. CYY-TD-2022-004).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets used in the study are available from the corresponding authors upon reasonable request.

Conflicts of Interest: The authors declare no competing interest. All authors listed have approved the manuscript that is enclosed.

References

1. Babu, T.N.; Ali, P.S.N.; Prabha, D.R.; Mohammed, V.N.; Wahab, R.S.; Vijayalakshmi, S. Fault Diagnosis in Bevel Gearbox Using Coiflet Wavelet and Fault Classification Based on ANN Including DNN. *Arab. J. Sci. Eng.* **2022**, 1–27. [\[CrossRef\]](#)
2. Wang, Z.; Huang, H.; Wang, Y. Fault diagnosis of planetary gearbox using multi-criteria feature selection and heterogeneous ensemble learning classification. *Measurement* **2020**, *173*, 108654. [\[CrossRef\]](#)
3. Li, C.; Sánchez, R.V.; Zurita, G.; Cerrada, M.; Cabrera, D.; Vásquez, R.E. Gearbox fault diagnosis based on deep random forest fusion of acoustic and vibratory signals. *Mech. Syst. Signal Process.* **2016**, *76–77*, 283–293. [\[CrossRef\]](#)
4. Surendran, R.; Khalaf, O.I.; Romero, C.A.T. Deep Learning Based Intelligent Industrial Fault Diagnosis Model. *Comput. Mater. Contin.* **2022**, *70*, 6323–6338. [\[CrossRef\]](#)
5. Meng, Z.; Huo, H.; Pan, Z.; Cao, L.; Li, J.; Fan, F. A gear fault diagnosis method based on improved accommodative random weighting algorithm and BB-1D-TP. *Measurement* **2022**, *195*, 111169. [\[CrossRef\]](#)
6. Xu, P.; Huang, L.; Song, Y. An optimal method based on HOG-SVM for fault detection. *Multimed. Tools Appl.* **2022**, *81*, 6995–7010. [\[CrossRef\]](#)
7. Grzeszkowski, M.; Nowoisky, S.; Scholzen, P.; Kappmeyer, G.; Gühmann, C.; Brimmers, J.; Brecher, C. Classification of gear pitting damage using vibration measurements. *Tm-Tech. Mess.* **2021**, *88*, 282–293. [\[CrossRef\]](#)
8. Xiao, Y.; Xue, J.; Li, M.; Yang, W. Low-Pass Filtering Empirical Wavelet Transform Machine Learning Based Fault Diagnosis for Combined Fault of Wind Turbines. *Entropy* **2021**, *23*, 975. [\[CrossRef\]](#)
9. He, Y.; Ye, L.; Zhu, X.; Wang, Z. Feature extraction based on PSO-FC optimizing KPCA and wear fault identification of planetary gear. *J. Mech. Sci. Technol.* **2021**, *35*, 2347–2357. [\[CrossRef\]](#)
10. Choudhary, A.; Goyal, D.; Letha, S.S. Infrared Thermography-Based Fault Diagnosis of Induction Motor Bearings Using Machine Learning. *IEEE Sens. J.* **2020**, *21*, 1727–1734. [\[CrossRef\]](#)
11. Inturi, V.; Pratyush, A.S.; Sabareesh, G.R. Detection of Local Gear Tooth Defects on a Multistage Gearbox Operating Under Fluctuating Speeds Using DWT and EMD Analysis. *Arab. J. Sci. Eng.* **2021**, *46*, 11999–12008. [\[CrossRef\]](#)
12. Wang, H.; Chan, L. Wavelet Packet Transform-Assisted Least Squares Support Vector Machine for Gear Wear Degree Diagnosis. *Math. Probl. Eng.* **2021**, *2021*, 9889933. [\[CrossRef\]](#)
13. Ramteke, D.S.; Pachori, R.B.; Parey, A. Automated Gearbox Fault Diagnosis Using Entropy-Based Features in Flexible Analytic Wavelet Transform (FAWT) Domain. *J. Vib. Eng. Technol.* **2021**, *9*, 1703–1713. [\[CrossRef\]](#)
14. Vos, K.; Peng, Z.; Jenkins, C.; Shahriar, R.; Borghesani, P.; Wang, W. Vibration-based anomaly detection using LSTM/SVM approaches. *Mech. Syst. Signal Process.* **2021**, *169*, 108752. [\[CrossRef\]](#)
15. Li, D.; Zhao, Y.; Zhao, Y. A dynamic-model-based fault diagnosis method for a wind turbine planetary gearbox using a deep learning network. *Prot. Control Mod. Power Syst.* **2022**, *7*, 1–14. [\[CrossRef\]](#)
16. Fu, Y.; Chen, X.; Liu, Y.; Son, C.; Yang, Y. Multi-Source Information Fusion Fault Diagnosis for Gearboxes Based on SDP and VGG. *Appl. Sci.* **2022**, *12*, 6323. [\[CrossRef\]](#)
17. Han, T.; Ma, R.; Zheng, J. Combination bidirectional long short-term memory and capsule network for rotating machinery fault diagnosis. *Measurement* **2021**, *176*, 109208. [\[CrossRef\]](#)

18. Kumar, A.; Vashishtha, G.; Gandhi, C.; Tang, H.; Xiang, J. Sparse transfer learning for identifying rotor and gear defects in the mechanical machinery. *Measurement* **2021**, *179*, 109494. [\[CrossRef\]](#)
19. Ha, J.M.; Youn, B.D. A Health Data Map-Based Ensemble of Deep Domain Adaptation Under Inhomogeneous Operating Conditions for Fault Diagnosis of a Planetary Gearbox. *IEEE Access* **2021**, *9*, 79118–79127. [\[CrossRef\]](#)
20. Zhu, X.; Ye, X.; Wang, R.; Zhao, W.; Luo, X.; Zhao, J.; Han, Z.; Gao, X. Investigation and Experimental Study on Gearbox Vibration Fault Diagnosis Method Based on Fusion Feature Convolutional Learning Network. *Exp. Tech.* **2022**, 1–12. [\[CrossRef\]](#)
21. 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* **2020**, *21*, 18. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Cui, L.; Tian, X.; Shi, X.; Wang, X.; Cui, Y. A Semi-Supervised Fault Diagnosis Method Based on Improved Bidirectional Generative Adversarial Network. *Appl. Sci.* **2021**, *11*, 9401. [\[CrossRef\]](#)
23. Cao, H.; Shao, H.; Zhong, X.; Deng, Q.; Yang, X.; Xuan, J. Unsupervised domain-share CNN for machine fault transfer diagnosis from steady speeds to time-varying speeds. *J. Manuf. Syst.* **2022**, *62*, 186–198. [\[CrossRef\]](#)
24. Wu, D.; Yan, P.; Pei, J.; Su, Y.; Zhou, H.; Yi, R.; Hu, G. Quality estimation method for gear hobbing based on attention and adversarial transfer learning. *Measurement* **2021**, *188*, 110383. [\[CrossRef\]](#)
25. Wang, J.; Yang, J.; Wang, Y.; Bai, Y.; Zhang, T.; Yao, D. Ensemble decision approach with dislocated time–frequency representation and pre-trained CNN for fault diagnosis of railway vehicle gearboxes under variable conditions. *Int. J. Rail Transp.* **2021**, *10*, 186–198.
26. Li, Y.; Zou, W.; Jiang, L. Fault diagnosis of rotating machinery based on combination of Wasserstein generative adversarial networks and long short term memory fully convolutional network. *Measurement* **2022**, *191*, 110826. [\[CrossRef\]](#)
27. Zhao, X.; Yao, J.; Deng, W.; Ding, P.; Ding, Y.; Jia, M.; Liu, Z. Intelligent Fault Diagnosis of Gearbox Under Variable Working Conditions with Adaptive Intra-class and Inter-class Convolutional Neural Network. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**. [\[CrossRef\]](#)
28. Dong, S.; He, K.; Tang, B. The fault diagnosis method of rolling bearing under variable working conditions based on deep transfer learning. *J. Braz. Soc. Mech. Sci. Eng.* **2020**, *42*, 1–13. [\[CrossRef\]](#)
29. Hasan, M.J.; Islam, M.M.M.; Kim, J.-M. Acoustic spectral imaging and transfer learning for reliable bearing fault diagnosis under variable speed conditions. *Measurement* **2019**, *138*, 620–631. [\[CrossRef\]](#)
30. Li, S.; An, Z.; Lu, J. A Novel Data-Driven Fault Feature Separation Method and Its Application on Intelligent Fault Diagnosis Under Variable Working Conditions. *IEEE Access* **2020**, *8*, 113702–113712. [\[CrossRef\]](#)
31. Xia, M.; Li, T.; Xu, L.; Liu, L.; de Silva, C.W. Fault Diagnosis for Rotating Machinery Using Multiple Sensors and Convolutional Neural Networks. *IEEE/ASME Trans. Mechatron.* **2018**, *23*, 101–110. [\[CrossRef\]](#)
32. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. In Proceedings of the International Conference on Learning Representations, San Diego, CA, USA, 7–9 May 2015.
33. Shao, S.; McAleer, S.; Yan, R.; Baldi, P. Highly Accurate Machine Fault Diagnosis Using Deep Transfer Learning. *IEEE Trans. Ind. Inform.* **2018**, *15*, 2446–2455. [\[CrossRef\]](#)
34. Malik, H.; Pandya, Y.; Parashar, A.; Sharma, R. *Feature Extraction Using EMD and Classifier through Artificial Neural Networks for Gearbox Fault Diagnosis*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 309–317.
35. Zhao, B.; Niu, Z.; Liang, Q.; Xin, Y.; Qian, T.; Tang, W.; Wu, Q. Signal-to-Signal Translation for Fault Diagnosis of Bearings and Gears with Few Fault Samples. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 3526910. [\[CrossRef\]](#)
36. Ma, Y.; Wen, G.; Cheng, S.; He, X.; Mei, S. Multimodal convolutional neural network model with information fusion for intelligent fault diagnosis in rotating machinery. *Meas. Sci. Technol.* **2022**, *33*, 125109. [\[CrossRef\]](#)
37. Durbhaka, G.K.; Selvaraj, B.; Mittal, M.; Saba, T.; Rehman, A.; Goyal, L.M. Swarm-LSTM: Condition Monitoring of Gearbox Fault Diagnosis Based on Hybrid LSTM Deep Neural Network Optimized by Swarm Intelligence Algorithms. *Comput. Mater. Contin.* **2021**, *66*, 2041–2059. [\[CrossRef\]](#)