

## Article

# Rolling Bearing Fault Diagnosis Using Hybrid Neural Network with Principal Component Analysis

Keshun You, Guangqi Qiu and Yingkui Gu \* 

School of Mechanical and Electrical Engineering, Jiangxi University of Science and Technology, Ganzhou 341000, China

\* Correspondence: guyingkui@163.com

**Abstract:** With the rapid development of fault prognostics and health management (PHM) technology, more and more deep learning algorithms have been applied to the intelligent fault diagnosis of rolling bearings, and although all of them can achieve over 90% diagnostic accuracy, the generality and robustness of the models cannot be truly verified under complex extreme variable loading conditions. In this study, an end-to-end rolling bearing fault diagnosis model of a hybrid deep neural network with principal component analysis is proposed. Firstly, in order to reduce the complexity of deep learning computation, data pre-processing is performed by principal component analysis (PCA) with feature dimensionality reduction. The preprocessed data is imported into the hybrid deep learning model. The first layer of the model uses a CNN algorithm for denoising and simple feature extraction, the second layer makes use of bi-directional long and short memory (BiLSTM) for greater in-depth extraction of the data with time series features, and the last layer uses an attention mechanism for optimal weight assignment, which can further improve the diagnostic precision. The test accuracy of this model is fully comparable to existing deep learning fault diagnosis models, especially under low load; the test accuracy is 100% at constant load and nearly 90% for variable load, and the test accuracy is 72.8% at extreme variable load (2.205 N·m/s–0.735 N·m/s and 0.735 N·m/s–2.205 N·m/s), which are the worst possible load conditions. The experimental results fully prove that the model has reliable robustness and generality.

**Keywords:** PHM; intelligent fault diagnosis; complex extreme variable loading; hybrid deep neural network; robustness and generality



**Citation:** You, K.; Qiu, G.; Gu, Y. Rolling Bearing Fault Diagnosis Using Hybrid Neural Network with Principal Component Analysis. *Sensors* **2022**, *22*, 8906. <https://doi.org/10.3390/s22228906>

Academic Editors: Yangquan Chen, Subhas Mukhopadhyay, Nunzio Cennamo, M. Jamal Deen, Junseop Lee, Simone Morais and Yolanda Vidal

Received: 8 October 2022

Accepted: 15 November 2022

Published: 17 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

Due to complex working conditions and frequently changing loads in actual production, a large number of mechanical system failures are caused by faults in bearings [1]. The mechanism of bearing damage is very complex; the machine operating environment [2], frequent fluctuations in load [3–5], and improper installation, etc., can all cause different types of bearing faults, mainly including abrasion failure, fatigue failure, corrosion failure, and cavitation failure [6]. It is very difficult and unrealistic to analyze and diagnose faults by only studying the mechanism [7], but some studies have modeled bearing dynamics in terms of the radial internal clearance of rolling bearings as a way of analyzing bearing failure and life [8,9], which provide good references. Therefore, we can combine mechanism analyses to research a better intelligent fault diagnosis method. Rolling bearings, as important rotating parts in machinery and equipment, are also one of the important sources of faults in machinery and equipment [10]. Rolling bearings are one of the most common and widely used kinds of bearing; therefore, the fault diagnosis method of rolling bearings has been one of the key technologies in the development of machinery fault diagnosis [11].

Fault prognostic and health management (PHM) systems need to have a complete, practical, intelligent, reliable, and systematic solution for rolling bearing health management [12,13], which includes raw data pre-processing, feature value selection and

extraction, failure mode identification, performance degradation assessment, performance trend prediction, and maintenance decision making [14]. Raw data pre-processing mainly includes outlier processing, zero-meaning processing, trend term elimination, and digital filtering [15]. Eigenvalue selection and extraction mainly includes time domain analysis, frequency domain analysis, time-frequency domain analysis, data constraint methods, and sensor information fusion [16,17]. Fault diagnosis mainly includes knowledge-based expert systems, analytic model-based state estimation, parameter estimation, and equivalence space [18], as well as data-driven neural networks, support vector machines (SVM) [19], information fusion, and multivariate statistical analysis [20]. Fault prediction mainly includes traditional reliability prediction methods based on fault tree analysis and lifetime distribution models, failure mechanism models based on physical methods of failure (FMM), Bayesian models based on statistics, Hidden Markov Models (HMM), and data-driven autoregressive models, neural network regression models, and support vector regression (SVR) models [21]. Performance trends are mainly predicted and evaluated in terms of model and algorithm errors.

In this paper, we propose a data-driven hybrid neural-network-based intelligent multi-classification algorithm that can automatically extract and process features, learn, reason, and decide whether a rolling bearing is normal or faulty, as well as what kind of fault it is, directly from the raw vibration data. The PCA method is used as data pre-processing to reduce computational complexity and improve feature extraction, and a convolutional neural network (CNN) algorithm with strong feature extraction capability is designed in the first layer of the network, which can autonomously design the size of the convolutional layers and enable initial learning and inference of the data. Then, a bi-directional Long Short Time Memory (BiLSTM) structure is designed in the second layer, which maximizes the learning performance of the model in learning time series data features. In order to reduce the number of parameters computed by the model, and to improve the speed and diagnostic performance of the model, an attention-based soft classification algorithm is designed in the last layer. In order to test the generality of the proposed model, we trained the model using operating data at 1 horsepower (1 hp = 0.735 N·m/s) load and operating conditions, conducted experimental tests using operating data at 0–2.205 N·m/s, and finally, we used bearing vibration data at six different variable load conditions (1-2 hp, 1-3 hp, 2-1 hp, 2-3 hp, 3-1 hp, 3-2 hp) for prognostic diagnosis.

The remainder of the paper is organized as follows: Section 2 analyzes related work involved in the rolling bearing fault diagnosis. Section 3 describes the theory and methods of rolling bearing fault diagnosis through the proposed hybrid deep neural network classification algorithm. Section 4 then analyses the experimental data and results, including a performance comparison with alternative methodologies. Finally, Section 5 concludes key results of the research.

## 2. Related Works

About 30–40% of equipment failures are caused by bearing failures. Since the 1990s [22], the research of algorithms for fault diagnosis of bearings has been a key topic. From traditional statistical methods to machine learning methods, and recently the application of deep learning algorithms, algorithms have been and continue to be areas of continued improvement and research.

Before the deep learning boom, classical machine learning (ML)-based algorithms were sensibly used in industrial bearing fault diagnosis [21]. For example: artificial neural networks (ANN) have been applied for more than 30 years [23]; principal component analysis (PCA) also has a great advantage in the extraction of sensitive features of bearing fault data [24,25]; K-Nearest Neighbors (K-NN) has been applied in bearing fault diagnosis for more than a decade [26]; and SVM has been famous for its powerful nonlinear processing capability with good generalization performance since its introduction. SVM is renowned for its powerful nonlinear processing ability, good generalization performance, being suitable for learning small sample data, and performance in bearing fault diagnosis,

which is far better than ANN, KNN, etc. [27], and it is often used in combination with other algorithms in practical research to achieve better diagnosis results [28]. Furthermore, neuro-fuzzy networks [29], Bayesian networks [16], self-organizing maps [30], extreme learning machines (ELM) [31], transfer learning [32], linear discriminant analysis [33], quadratic discriminant analysis [34], random forests [35], independent component analysis [36], empirical pattern decomposition [37], correlation analysis [38], affinity propagation [39], and dictionary learning [40] have all been widely used in bearing fault diagnosis, among others.

With the advent of big data and the development of high performance processors, deep learning algorithms have emerged in bearing fault diagnosis. The CNN's excellent denoising and classification performance, which has been widely used in image processing, has also been gradually applied to bearing fault diagnosis [41]. Examples include Adaptive CNN (ADCNN) [42], LeNet-5 based CNN [43], Deep Fully Convolutional Neural Network (DFCNN) [44], Multiscale CNN (MS-DCNN) [45], Pythagorean Spatial Pyramid Pooling (PSPP) CNN [46], and Adaptive Overlapping CNN (AOCNN) [47]. With the demand for unsupervised learning, a combination of auto-encoders (AE) were developed for application to bearing fault diagnosis [48]. Various deep learning convolutional neural networks have been developed and applied to fault diagnosis, such as the Deep Belief Network (DBN) [49], the Recurrent Neural Network (RNN) [50], the Generative Adversarial Network (GAN) [51], Long Short Time Memory (LSTM) [52], transfer learning based deep learning [53], the Graph Neural Network (GNN) [54], and their variants, etc.

The accuracy of bearing fault diagnosis based on advanced deep learning algorithms has reached excellent accuracy of over 90%, but more efficient, versatile, and more precisely accurate hybrid deep learning neural network models are needed to cope with fault diagnosis problems under various variable load conditions.

### 3. Introduction to Theory and Methodology

#### 3.1. Level I: Fault Generation of Rolling Bearings

Most bearings cannot reach the designed life during operation, mainly because of poor lubrication, unreasonable assembly, and manufacturing defects. In order to diagnose the failure of the bearings during operation, we generally use more advanced sensors to obtain the vibration signal of the corresponding position, and then combine the signal processing method.

As shown in Figure 1, first assume that the number of bearing balls is  $Z$ , the diameter of balls is  $d$ , the bearing raceway pitch is  $D$ , the bearing contact angle is  $\alpha$ , the inner raceway radius is  $r_1$ , the outer raceway radius is  $r_2$ , and the bearing inner ring speed is  $n$ . Theoretically, the characteristic frequency equation of rolling bearing has the following:

Inner ring rotation frequency  $f_i$

$$f_i = \frac{n}{60} \quad (1)$$

Relative rotation frequency  $f_r$  of inner and outer ring, because the outer ring of the rolling bearing does not rotate, the outer ring rotation frequency  $f_0$  is 0,

$$f_r = f_i - f_0 = f_i \quad (2)$$

Frequency  $f_{ic}$  of rolling body passing a point of inner ring,

$$f_{ic} = \frac{1}{2}Z \left( 1 + \frac{d}{D} \cos \alpha \right) f_r \quad (3)$$

Frequency  $f_{oc}$  of rolling body passing a point on the outer ring,

$$f_b = \frac{1}{2}Z \left( 1 - \frac{d}{D} \cos \alpha \right) f_r \quad (4)$$

Rotational frequency  $f_b$  of the rolling body. The calculation formula is equivalent to the calculation formula of cage rotation frequency  $f_c$ .

$$f_b = \frac{1}{2} \left( 1 - \frac{d}{D} \cos \alpha \right) f_r \quad (5)$$

When a fault occurs, the fault frequency of the bearing can be empirically calculated, at which time the fault frequency of the inner ring becomes:

$$f_i = 0.6 \times Z \times f_r \quad (6)$$

Fault frequency of outer ring:

$$f_0 = 0.4 \times Z \times f_r \quad (7)$$

Frequency of cage faults:

$$f_c = 0.381 - 0.4 \times f_r \quad (8)$$

Frequency of rolling body faults:

$$f_c = \begin{cases} 0.23 \times Z \times f_r, & Z < 0 \\ 0.18 \times Z \times f_r, & Z \geq 10 \end{cases} \quad (9)$$

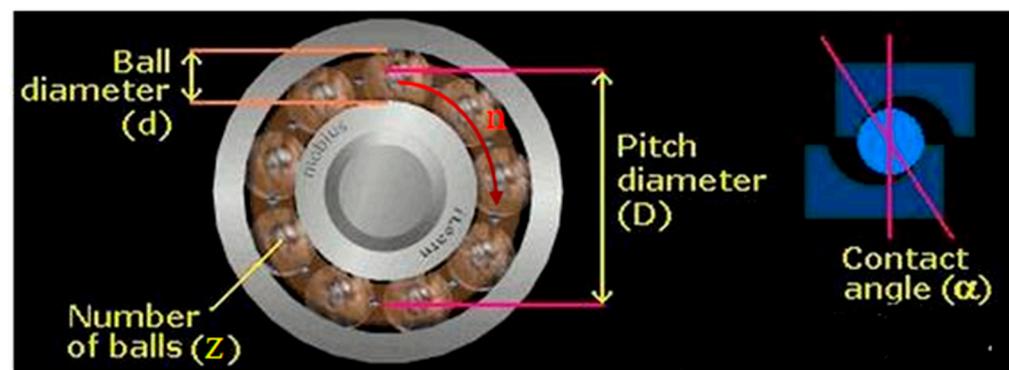
Frequency relationship between outer ring and cage:

$$f_0 = Z \times f_c \quad (10)$$

Frequency relationship between outer ring and inner ring:

$$f_0 + f_i = Z \times f_r \quad (11)$$

Based on the analysis of the failure mechanism of rolling bearings, combined with the actual production experience derived from the failure formula, to a certain extent, we can indeed produce a practical guidance, but the role produced is limited; the biggest drawback is it is unreliable and precision diagnosis is too low.

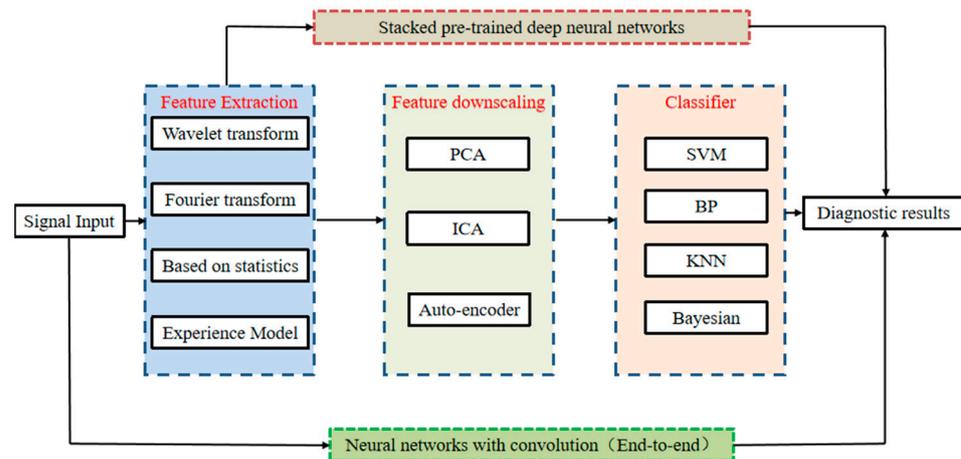


**Figure 1.** Schematic diagram of the geometric parameters of rolling bearings.

### 3.2. Level II: Fault Diagnosis Methods

Bearing fault diagnosis has been a popular area of research, and an algorithm usually includes two parts: signal feature extraction and classification. Common feature extraction algorithms include fast Fourier variation, wavelet transform, empirical pattern decomposition, and statistical features of the signal, etc. As shown in Figure 2, the intelligent diagnosis method based on machine signal processing through feature extraction algorithm combined

with classifier requires expert experience, a time-consuming design, and cannot guarantee generality, and thus, it is difficult to meet the requirements of large data and accuracy.



**Figure 2.** Intelligent fault diagnosis method.

### 3.3. Level III: The Proposed Hybrid Neural Network Fault Diagnosis Method with PCA

#### 3.3.1. Data Pre-Processing with PCA

PCA is the most commonly used linear dimensionality reduction method. The goal is to map high-dimensional data into a low-dimensional space by linear projection and expect the maximum amount of information (maximum variance) in the projected dimension, so as to use fewer data dimensions while retaining the characteristics of more original data points. The purpose is to reduce the noise or computational effort of the data while trying to ensure that the amount of information is not distorted.

First, assume that the data set  $X = [x_1, x_2, \dots, x_n]$  has  $n$  sets of data and each set has  $m$  features.

- (1) Normalization of the data, i.e.,

$$Z_{ij} = \frac{x_{ij} - \bar{x}_j}{\sqrt{\sigma(x_j)}} \quad (12)$$

- (2) Calculating the covariance matrix  $C$  of the normalized data.

$$C(X_i, X_j) = \frac{\sum_{k=1}^n (x_i^k - \bar{x}_i^k)(x_j^k - \bar{x}_j^k)}{m - 1} \quad (13)$$

- (3) Computing the eigenvalues  $(\lambda_1, \lambda_2, \dots, \lambda_m)$  of the covariance matrix  $C$ .

$$Cu = \lambda u \quad (14)$$

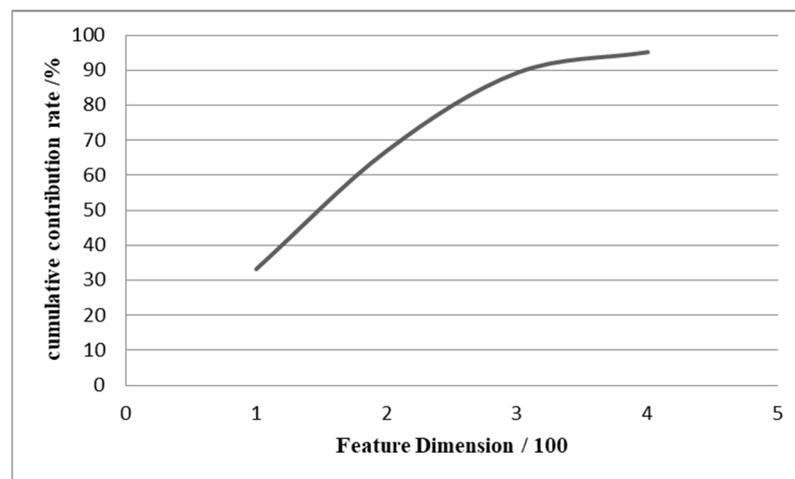
- (4) Calculating the cumulative contribution of the first  $k$  principal components. When the cumulative contribution rate  $(p) \geq 90\%$ , only the first  $k$  feature vectors can be extracted as sample features, and the larger the cumulative contribution rate is, the more original information is included.

$$(p) = \sum_{i=1}^p \lambda_i / \sum_{i=1}^m \lambda_i \quad (15)$$

- (5) Deriving the principal component feature vector  $Y$  with reduced dimensional.

$$Y = \begin{bmatrix} y_i^i \\ \bullet \\ \bullet \\ \bullet \\ y_k^i \end{bmatrix} = \begin{bmatrix} u_1^T \cdot (x_1^i, x_2^i, \dots, x_n^i)^T \\ u_2^T \cdot (x_1^i, x_2^i, \dots, x_n^i)^T \\ \bullet \\ \bullet \\ u_k^T \cdot (x_1^i, x_2^i, \dots, x_n^i)^T \end{bmatrix} \quad (16)$$

As shown in Figure 3, the original feature data is 470 dimensions, and the 330 most sensitive features can be obtained by calculating the cumulative contribution rate of 90%. By using this data pre-processing method with PCA, we can achieve a successful feature dimension reduction of 30%, which will effectively reduce the workload of the later deep learning computation and will also produce some denoising.



**Figure 3.** The feature dimension of 90% cumulative contribution rate.

### 3.3.2. End-to-End Hybrid Neural Network Fault Diagnosis Method

The proposed end-to-end hybrid neural-network-based intelligent diagnosis algorithm is used to simultaneously complete feature extraction and fault detection, which integrates all the advantages of CNN, LSTM, and attention mechanisms. Being a hybrid deep learning convolutional neural network model, it has already achieved good prospects in medical biology and other applications [55], and still continues to prosper in engineering applications.

As shown in Figure 4, the first layer of the model has a 1-dimensional CNN convolutional layer, which is mainly responsible for the denoising and feature extraction of the original vibration data; its calculation is shown in Equation (17).

$$(i, j) = I * K(i, j) = \sum_m \sum_n I(i + m, j + n) K(m, n) \quad (17)$$

where  $S$  represents the result of the operation;  $I$  is the original image;  $K$  is the convolution kernel;  $m, n$  are the height and width of the convolution kernel;  $i, j$  represent the position of the convolved.

The second layer is designed as a bi-directional LSTM (Bi-LSTM), which is a type of recurrent neural network (RNN). In practice, RNNs have been found to have problems such as gradient disappearance, gradient explosion, and a poor ability to rely on information over long distances; thus, the LSTM was introduced. The LSTM is similar to a RNN in terms of its main structure, but the main improvement is the addition of three gates in the hidden layer  $h$ , which are a forgetting gate, input gate, and output gate, as well as the addition of a new cell state. The principle is shown in Figure 4.  $f(t)$ ,  $i(t)$ , and  $o(t)$  represent the values of the forgetting gate, input gate, and output gate at time  $t$ , respectively.  $\alpha(t)$

denotes the initial feature extraction of  $h(t - 1)$  and  $x(t)$  at time  $t$ . The specific calculation process is shown in Equations (18)–(21).

$$f(t) = \sigma(W_f h_{t-1} + U_f x_t + b_f) \tag{18}$$

$$i(t) = \sigma(W_i h_{t-1} + U_i x_t + b_i) \tag{19}$$

$$\alpha(t) = \tanh(W_a h_{t-1} + U_a x_t + b_a) \tag{20}$$

$$o(t) = \sigma(W_o h_{t-1} + U_o x_t + b_o) \tag{21}$$

where  $x_t$  denotes the input at time  $t$ ;  $h_{t - 1}$  denotes the hidden state value at time  $t - 1$ ;  $W_f$ ,  $W_i$ ,  $W_o$ , and  $W_a$  denote the weight coefficients of  $h_{t-1}$  in the forgetting gate, input gate, output gate, and feature extraction process, respectively;  $U_f$ ,  $U_i$ ,  $U_o$ , and  $U_a$  denote the weight coefficients of  $x_t$  in the forgetting gate, input gate, output gate, and feature extraction process, respectively;  $b_f$ ,  $b_i$ ,  $b_o$ , and  $b_a$  denote the bias values of  $x_t$  in the forgetting gate, input gate, output gate, and feature extraction process, respectively;  $U_f$ ,  $U_i$ ,  $U_o$ , and  $U_a$  denote the weight coefficients of the forgetting gate, input gate, output gate, and feature extraction process  $x_t$ , respectively;  $b_f$ ,  $b_i$ ,  $b_o$ , and  $b_a$  denote the bias values of the forgetting gate, input gate, output gate, and feature extraction process, respectively;  $\tanh$  denotes the tangent hyperbolic function and  $\sigma$  denotes the activation function Sigmoid.

$$\tanh(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}} \tag{22}$$

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{23}$$

The result of the forgetting gate and input gate calculation acting on  $c(t - 1)$  constitutes the cell state  $c(t)$  at time  $t$ , which is expressed in Equation (19) as:

$$c(t) = c(t - 1) \odot f(t) + i(t) \odot \alpha(t) \tag{24}$$

where  $\odot$  is the Hadamard product. Eventually, the hidden layer state  $h(t)$  at time  $t$  is solved by the output gate  $o(t)$  and the cell state  $c(t)$  at the current moment.

$$h(t) = o(t) \odot \tanh(c(t)) \tag{25}$$

As shown in Figure 5, the BiLSTM neural network structure model consists of two independent LSTM, as shown in Figure 6. The input sequences are input into the two LSTM in positive and negative order, respectively, for feature extraction, and the two output vectors (i.e., the extracted feature vectors) are stitched together to form the final feature expression of the output.

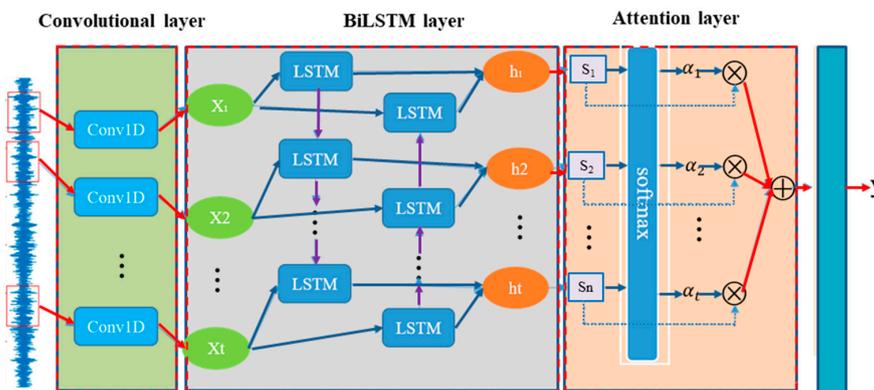


Figure 4. The proposed hybrid neural network fault detection method.

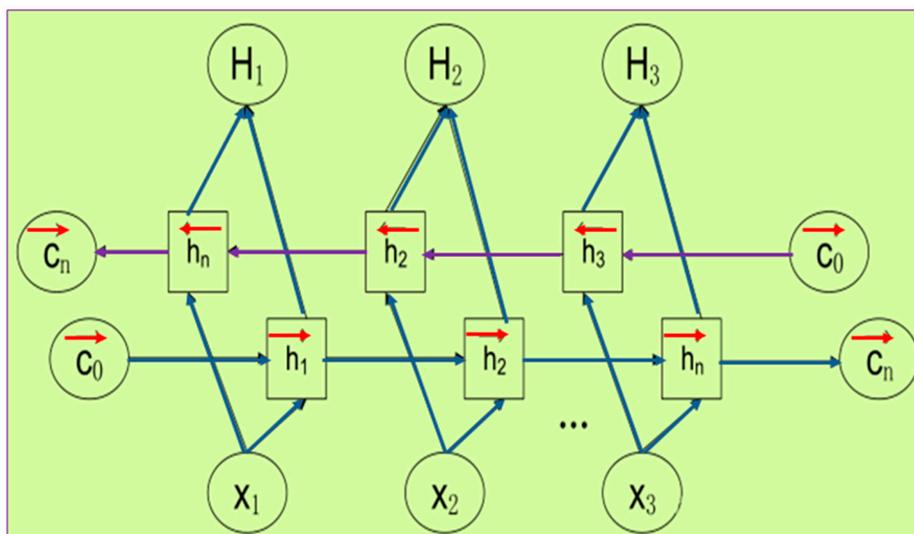


Figure 5. BiLSTM principle.

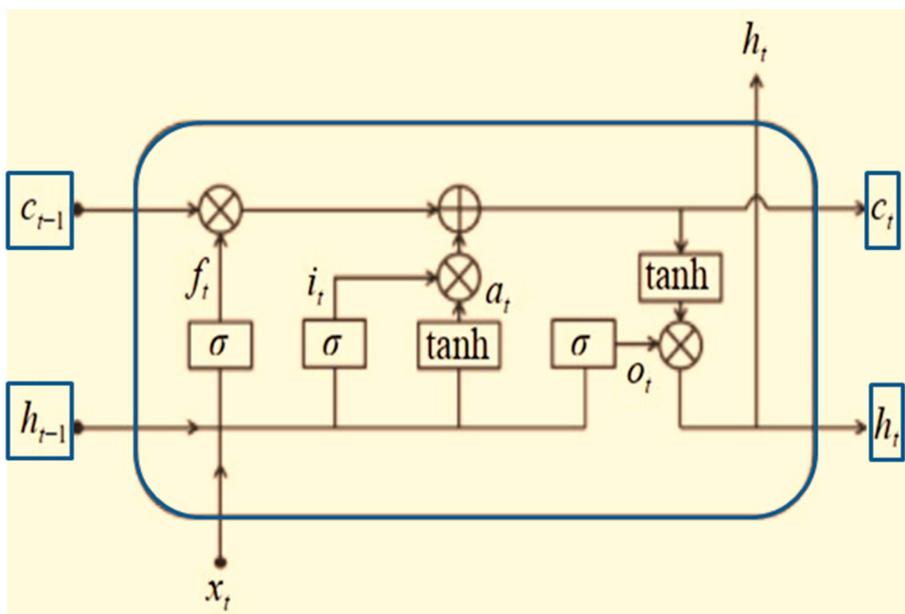


Figure 6. LSTM principle.

The BiLSTM model is designed so that the feature data obtained at time  $t$  has information between the past and the future at the same time. Experimentally, this neural network structure has proven to be more efficient and perform better than a single LSTM structure for feature extraction of time series data. It is worth mentioning that the parameters of the 2 LSTM neural networks in BiLSTM are independent of each other.

As shown in Figure 7, when the feature data extracted by BiLSTM is fed to the Attention mechanism layer, the Attention technique causes the data to be classified as more feature-specific by weighting the data with different features and reassigning the weights through the learning and scoring results. Here, the Score is first defined as Equation (26).

$$Score(h_t, \bar{h}_s) = h_t^T \cdot W \cdot \bar{h}_s \tag{26}$$

where  $h_t$  is the hidden state of the decoder at time  $t$ , and  $\bar{h}_s$  denotes the hidden states of the encoder,  $W$  is a matrix to be learned, which is used throughout the process. After the score is obtained, we can find the weight of attention  $\alpha_{ts}$ .

$$\alpha_{ts} = \frac{\exp(\text{Score}(h_t, \bar{h}_s))}{\sum_{s'=1}^S \exp(\text{Score}(h_t, \bar{h}_{s'}))} \quad (27)$$

Then, the weights are multiplied with the hidden states in the encoder to obtain the feature vector  $c_t$ .

$$c_t = \sum_s \alpha_{ts} \bar{h}_s \quad (28)$$

After that, we can calculate the Attention vector  $\alpha_t$ , combined with the weights of attention  $\alpha_{ts}$ , and the final value of attention can be derived.

$$\alpha_t = f(c_t, h_t) = \tan(W_c[c_t, h_t]) \quad (29)$$

The Attention Mechanism is an information filtering method that further alleviates the problem of long-term dependency in LSTM and GRU [56]. In general, this can be achieved in three steps: first, a task-relevant representation vector is introduced as a benchmark for feature selection, a manually specified hyperparameter, which can be either a dynamically generated vector or a learnable parameter vector; then, a scoring function is chosen to calculate the correlation between the input features and this vector to obtain the probability distribution of the features being selected, which is called the attention distribution; finally, a weighted average of the input features by the attention distribution filters out the task-relevant feature information.

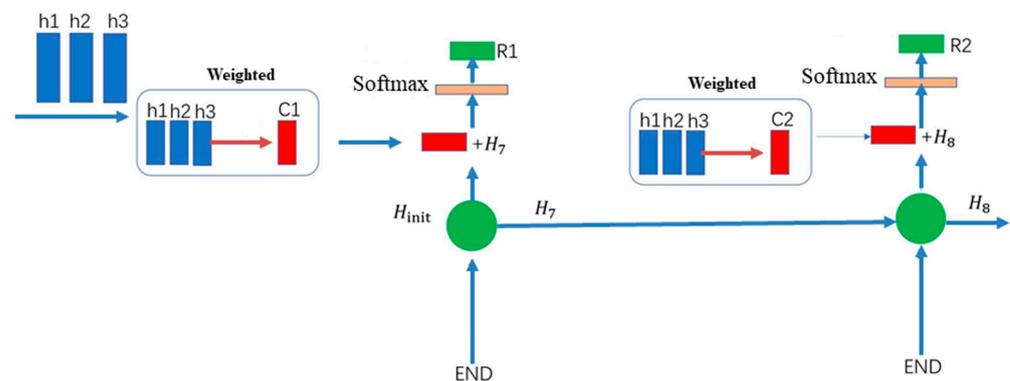


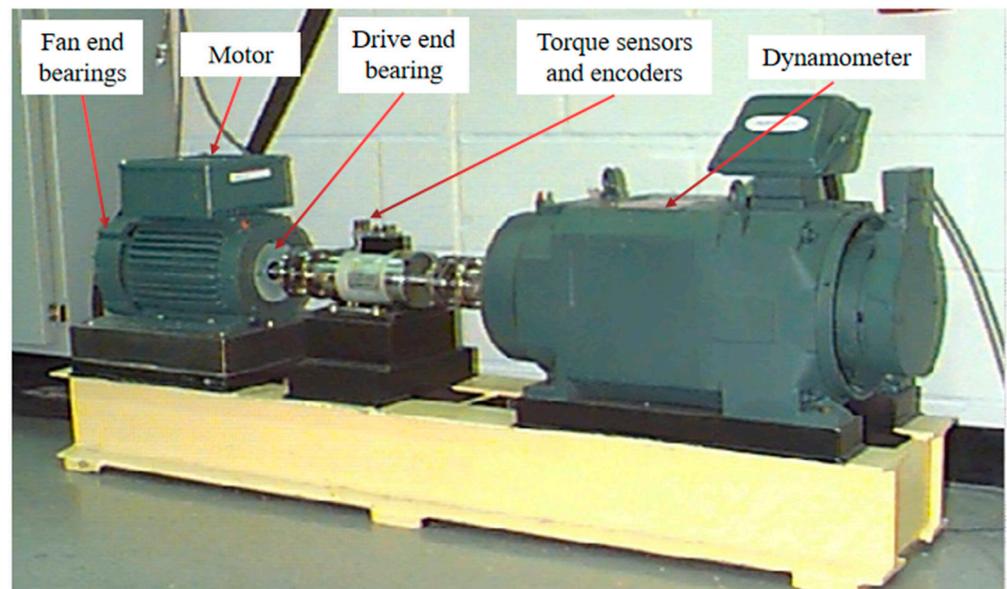
Figure 7. Attention technique.

#### 4. Bearing Diagnostic Performance Verification of the Proposed Model

As shown in Figure 8, the experimental platform consists of a drive motor, a torque transducer, and a power tester (right side of the figure) [57].

##### 4.1. Level I: Introduction to the Conditions and Data Set of the Experiment

Rolling bearing fault diagnosis is generally performed using the CWRU dataset to standardize the strengths and weaknesses of detection algorithms. As shown in Table 1, the data in this experiment uses DE (drive end) accelerometer data and a bearing with SKF6205 type load for 0–3 horsepower (0–2.205 N·m/s) corresponding to the approximate motor speed of 1797 r/min, 1772 r/min, 1750 r/min, and 1730 r/min; the sampling frequency is 48 kHz, the experimental single point damage diameter of the selected bearing is 0.007 mm, 0.014 mm, and 0.021 mm, and each fault diameter contains a rolling body fault, inner ring fault, and outer ring fault [58].



**Figure 8.** Equipment for experimental data acquisition (from Case Western Reserve University (CWRU)).

**Table 1.** CWRU bearing open data set sample processing.

Sampling Position	Bearing Type	Sampling Frequency	Load	Single-Point Loss Diameter
Drive end(DE) 6 o'clock	SKF6205	48 KHZ	0–2.205 N·m/s	0.007 mm, 0.014 mm, 0.021 mm

As shown in Table 2, the experimental dataset consists of nine fault datasets and one normal dataset, and the datasets used for training are generated by combining the ten classes of datasets corresponding to 0–3 hp, respectively. The sample size of each class is 256 and the sample size of the combined dataset is 1024; 70% of the combined dataset is used as the training set and 30% as the test set.

**Table 2.** Open CWRU bearing data set sample.

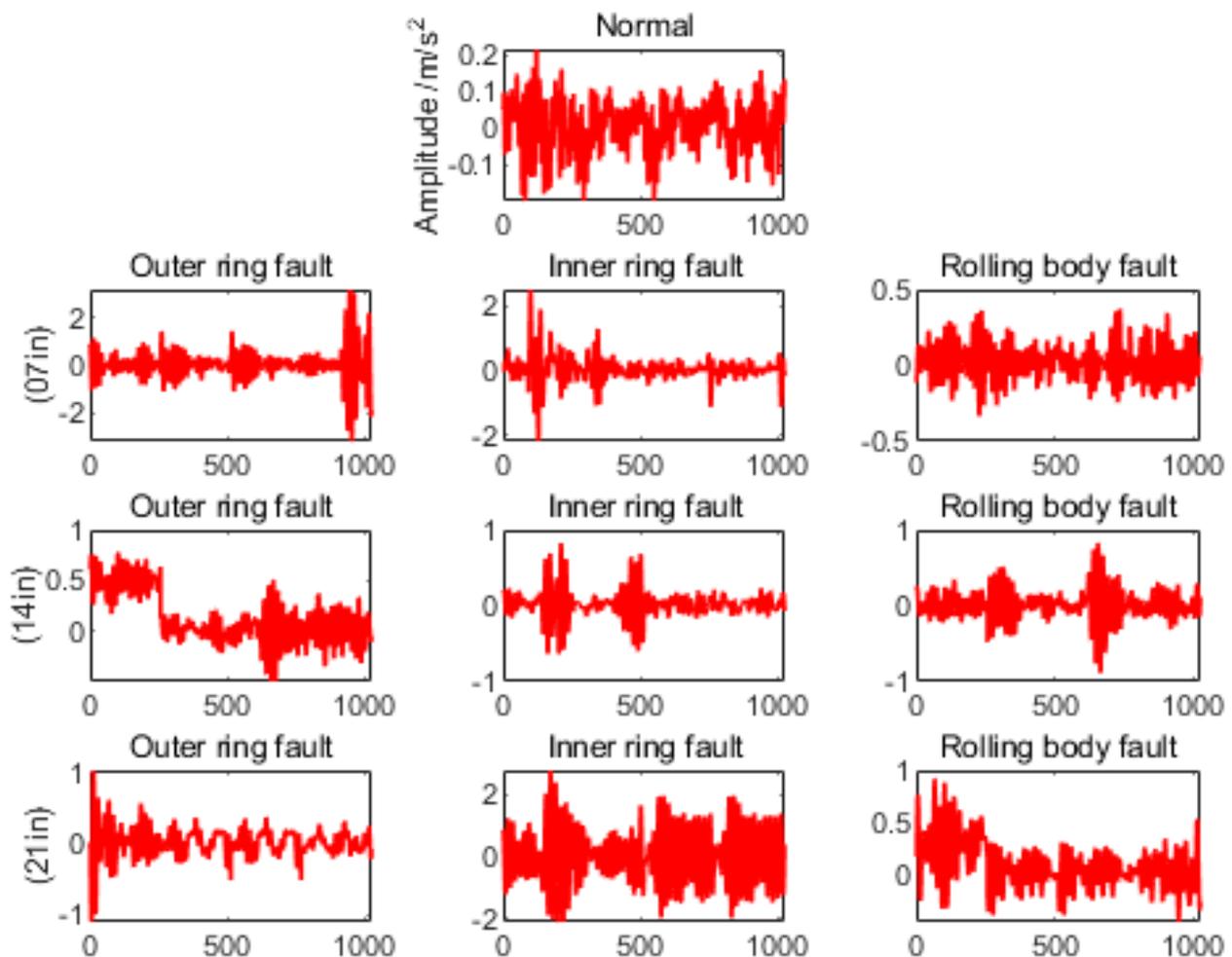
Sample Type	Sample Number	Sample Length	Training Sets	Test Set	Category Marker
Normal	1000	470	700	300	0
Outer ring fault (07 in)	1000	470	700	300	1
Outer ring fault (14 in)	1000	470	700	300	2
Outer ring fault (21 in)	1000	470	700	300	3
Inner ring fault (07 in)	1000	470	700	300	4
Inner ring fault (14 in)	1000	470	700	300	5
Inner ring fault (21 in)	1000	470	700	300	6
Rolling body fault (07 in)	1000	470	700	300	7
Rolling body fault (14 in)	1000	470	700	300	8
Rolling body fault (21 in)	1000	470	700	300	9

As shown in Figure 9, by observing the number of features, magnitude, fluctuation period, and phase difference in the bearing vibration signal, it can be found that the normal bearing vibration is more regular and the period is more stable, but after carefully observing the rolling bearing vibration signal with a fault, it is found that it is difficult to classify the bearing with a fault by manual observation of these data due to the influence of noise, different working conditions, and limited human perception ability. As shown in Figure 10, the linear FFT analysis allows for a rough determination of the frequency that produces the

maximum vibration signal; i.e., the most likely frequency of the fault. Although the FFT analysis method gives us a simple and reliable way to diagnose faults directly to the senses, it is limited by problems such as noise in the vibration signal and imbalance in the data, so it is not really a very objective diagnostic method.

#### 4.2. Level II: Training Results of the Model and Testing under Different Load Cases

The model is trained using the created dataset, and the trained model is also saved. The accuracy curve and loss rate curve of the model in the training process are shown in Figure 11. We can see that the model is well trained and there is no overfitting phenomenon. This is because, in the training process, we use a 10-fold cross-validation method, which groups the raw dataset into a training set and a validation set or test set. Firstly, divide the dataset into ten parts. Then, take turns to allocate nine of them for training and one for validation, and finally, use the mean of the ten results as an estimate of the accuracy of the algorithm.



**Figure 9.** Visual analysis of rolling bearing data (occurring on outer ring, inner ring, and rolling body with single point diameter damage of 0.007 mm (07 in), 0.014 mm (14 in), and 0.021 mm (21 in), respectively).

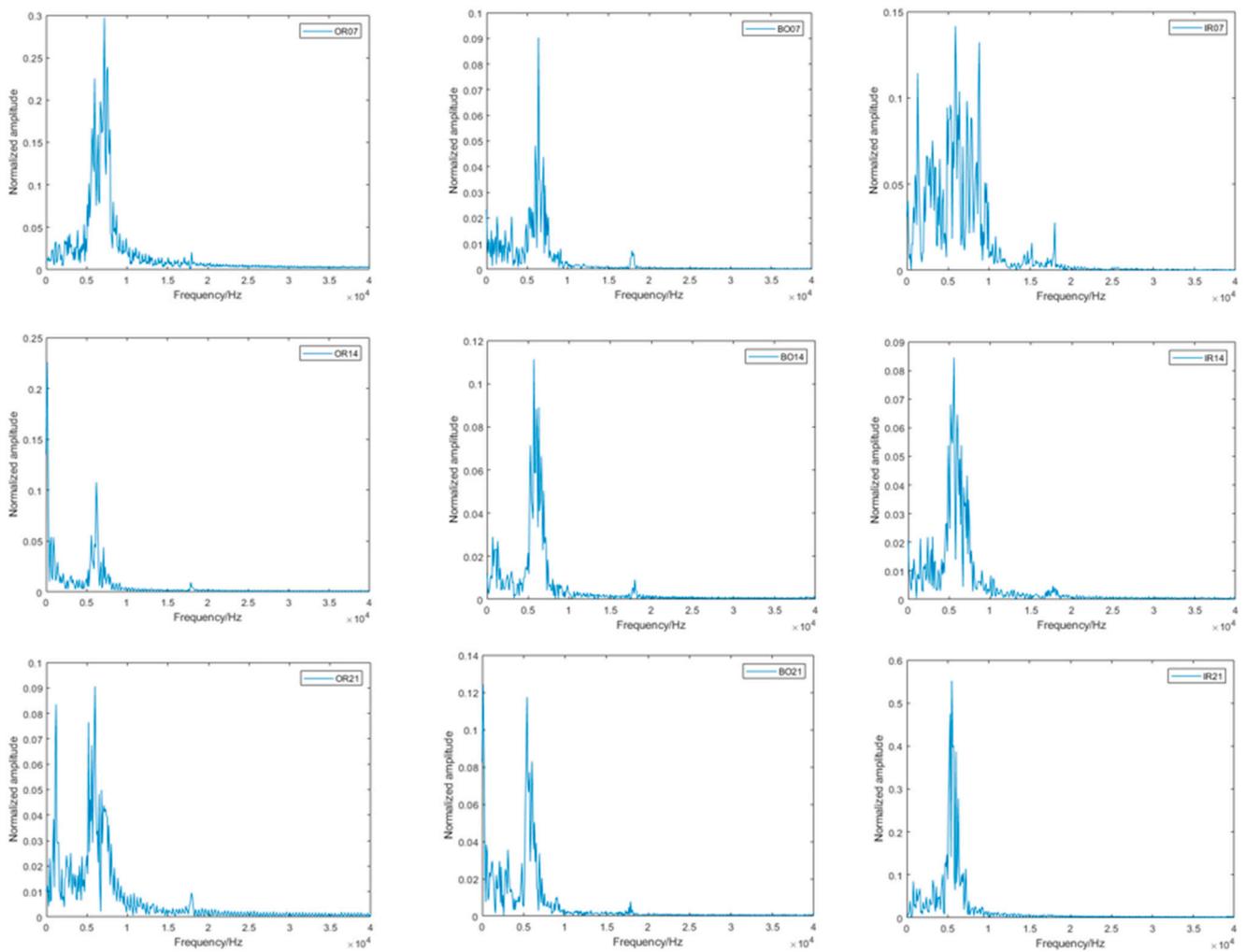


Figure 10. FFT analysis of rolling bearing faults data.

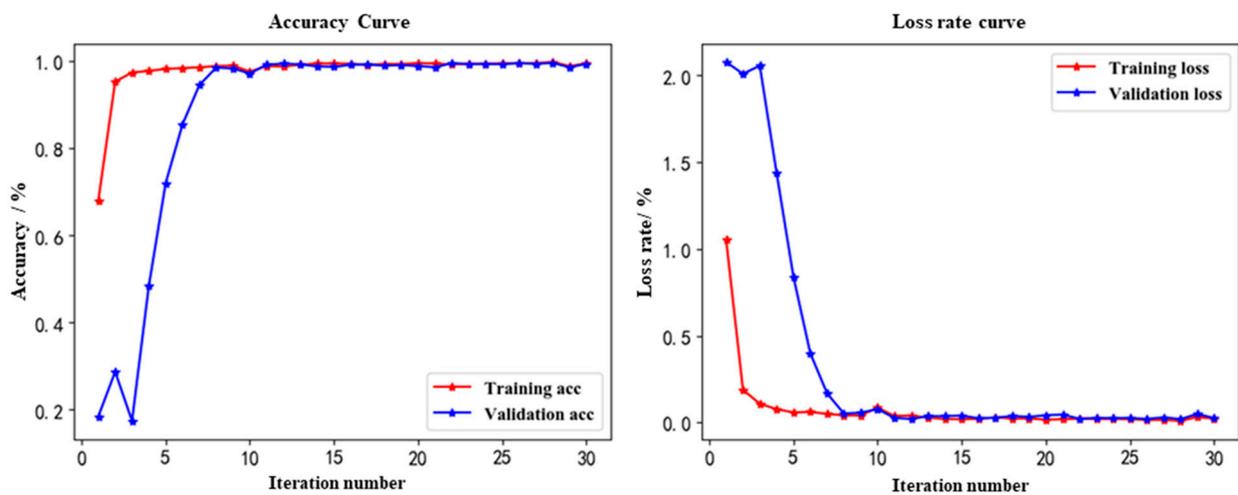


Figure 11. Training result.

To refine the fault diagnosis of the model, the model was tested using data under 0 hp, 1 hp, 2 hp, and 3 hp loads, and a confusion matrix was used to represent the fault diagnosis results. As can be seen from the confusion matrix in Figure 12, only some of the samples

are incorrectly identified, and most of them are 100% identified. Combining the four plots, the diagnostic results are a little better under loads of 0 hp, 1 hp, and 2 hp conditions. To more clearly represent the feature extraction ability of the model, t-SNE is introduced to downscale and visualize the features of each network layer of the model; only the feature extraction results of each network layer under the load of 1 hp condition are shown in this paper.

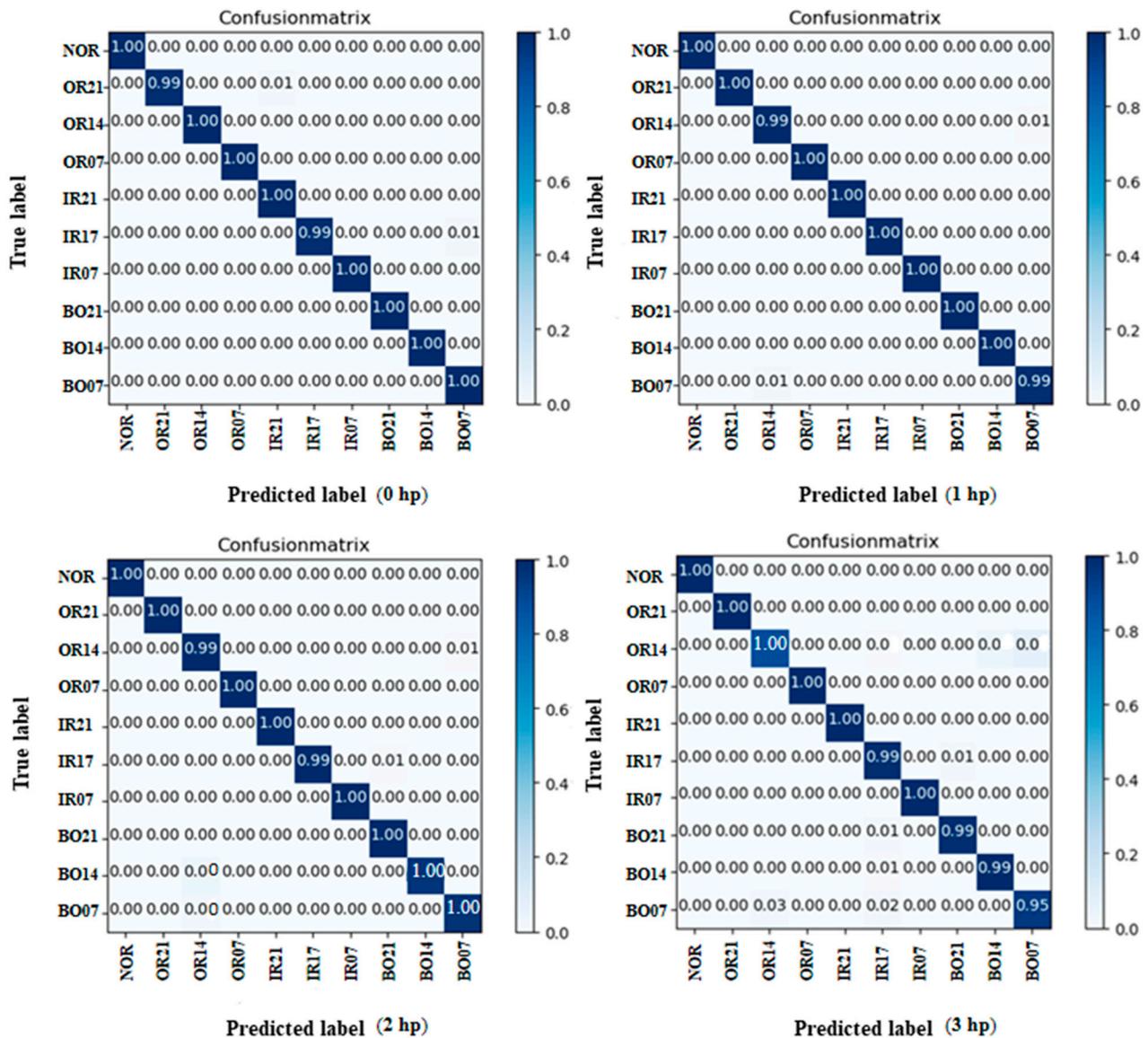


Figure 12. Test results of model under a load of 0.735 N·m/s~2.205 N·m/s.

As shown in Figure 13, which provides the t-SNE visualization results of each layer of the model when the input layer is a time-domain signal, the data of the bearings in different operation states are mixed with each other, and their clustering effect is extremely poor. From the t-SNE visualization results of the convolutional layer, we can see that some samples with the same type have already started to aggregate, and, as the network layer goes deeper, the t-SNE visualization results of the BiLSTM of second layer have basically completed the accurate classification of the vast majority of samples, and only a small number of samples are misclassified. Finally, the clustering effect is more obvious in the attention layer of the third layer, which is consistent with the results of the confusion matrix above, and also proves that the model has superior diagnostic ability. As shown in Table 3,

the end-to-end rolling bearing fault diagnosis model, based on the hybrid deep neural network proposed in this paper, is not only efficient but also has high accuracy and some advantages compared with other deep learning neural network fault diagnosis methods.

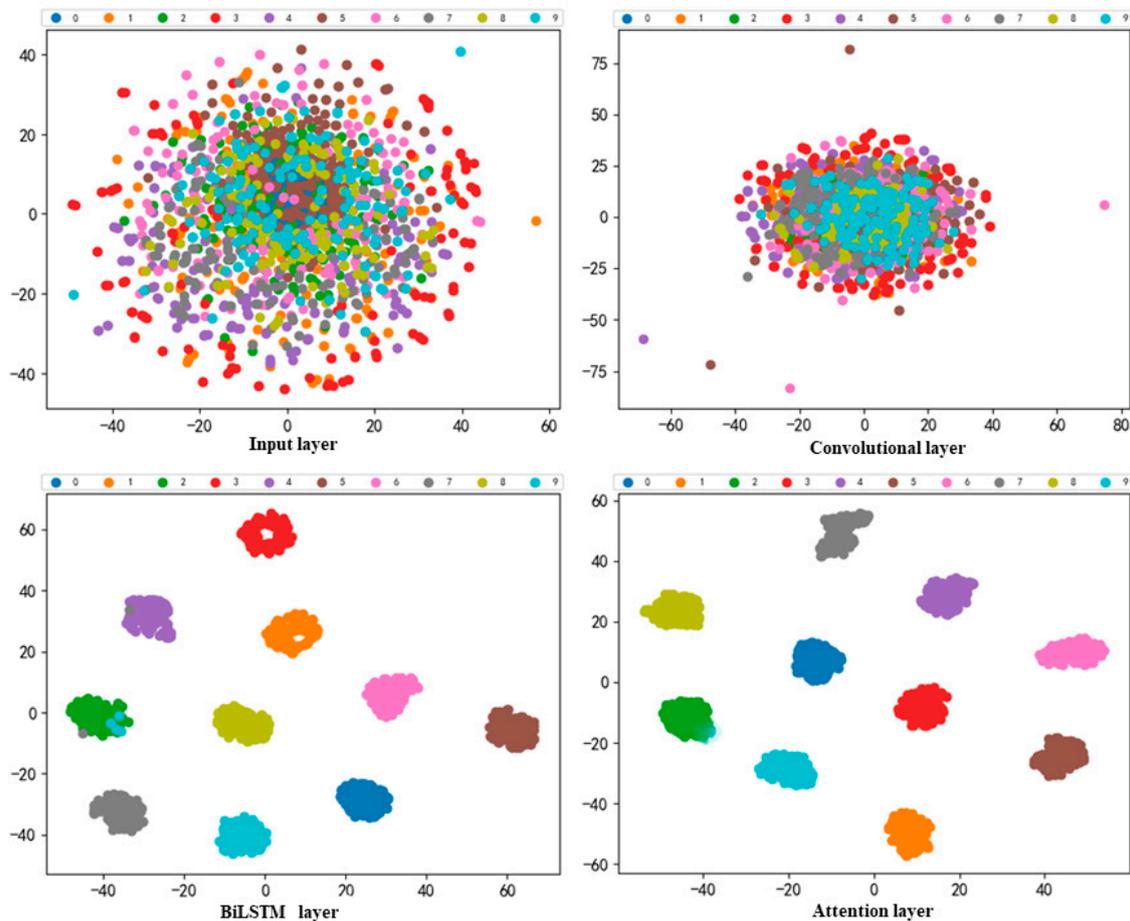


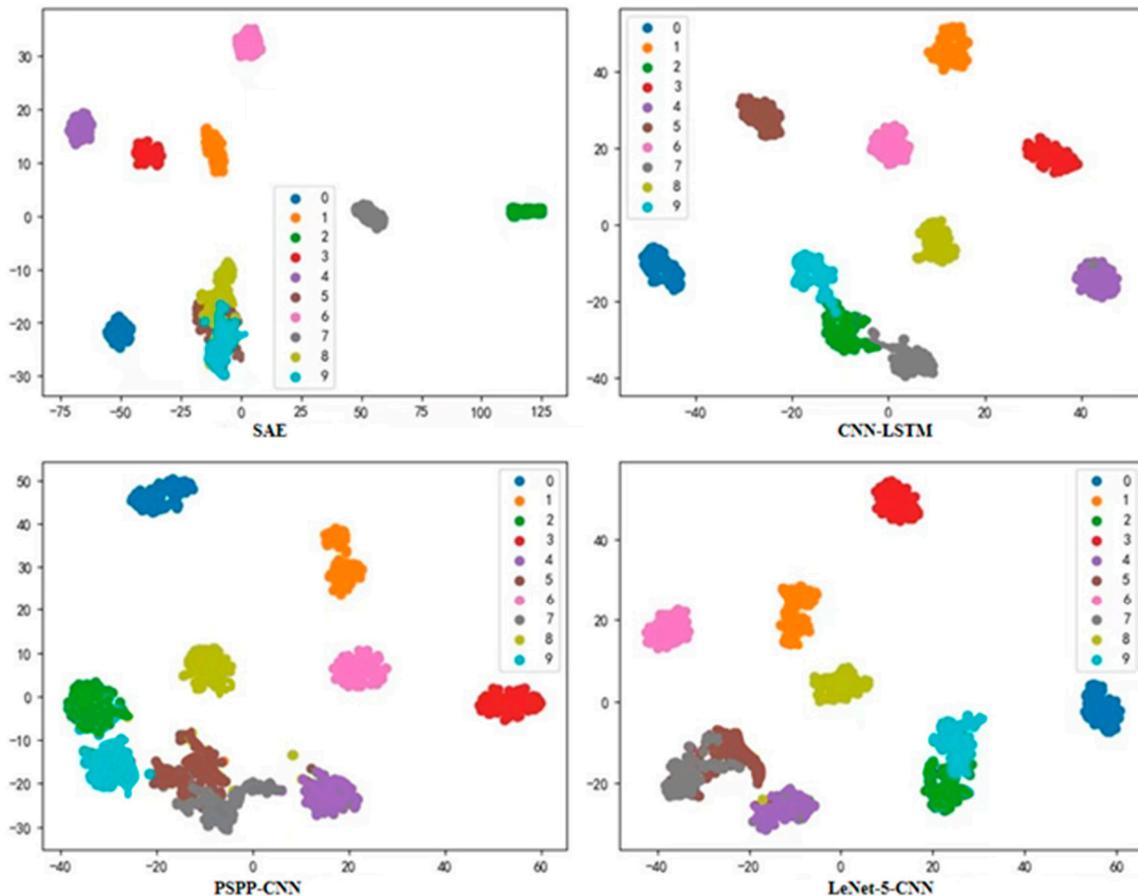
Figure 13. Visualization results of t-SNE for model feature extraction.

Table 3. Comparison of the accuracy of the proposed model with other models in rolling bearing fault diagnosis for CWRU.

Fault Diagnosis Model	Classifier	Percentage of Training Samples	Average Accuracy
ADCNN	Softmax	50%	97.9%
CNN	Softmax	90%	92.6%
LeNet-5-CNN	FC layer	83%	99.79%
IDS-CNN	Softmax	80%	98.92%
PSPP-CNN	Softmax	67%	99.19%
AOCNN	Softmax	50%	99.19%
SAE	ELM	50%	99.61
DBN	Softmax	N/A	98.8%
CNN-LSTM	Softmax	83%	99.6%
DC-GAN	SVM	96%	86.33%
GAAN-SDAE	Softmax	78%	99.2%
<b>The proposed</b>	<b>Softmax</b>	<b>70%</b>	<b>99.98%</b>

As shown in Figure 14, in order to compare the classification effectiveness of the proposed model with other intelligent fault diagnosis models, four existing, more advanced deep learning fault diagnosis models (SAE, CNN-LSTM, PSPP-CNN, and LeNet-5-CNN) were statistically analyzed. The CWRU dataset was still used for testing, and the objective

of the test was to de-classify ten types of rolling bearing faults from 0–9 categories at a motor load of 2.205 N·m/s. The classification results from the t-SNE visualization statistics of the four models show that although they are good enough for fault diagnosis, there is still some gap in classification effectiveness with the hybrid deep learning model proposed in this study.

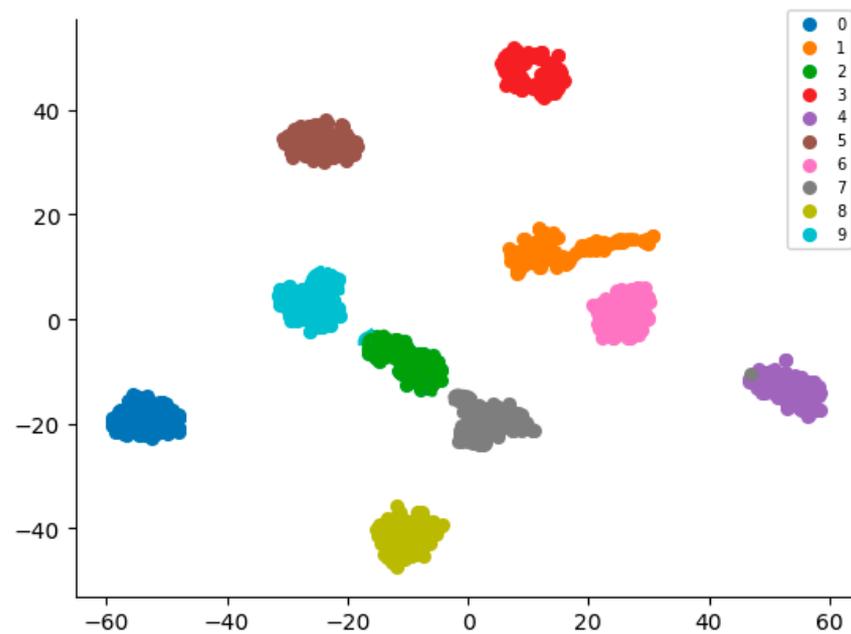


**Figure 14.** Visualization results of t-SNE in four model tests (for comparison with the proposed model).

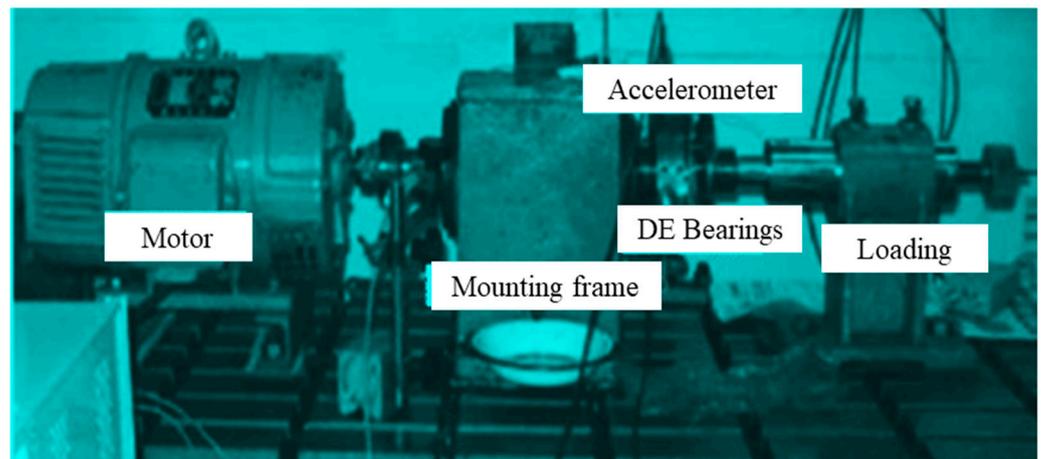
As shown in Figure 15, to test the generalization performance of the model proposed in this study, this paper conducted a cross-dataset test, using the same amount of data from the real dataset of an official industrial big data competition to test the model. It was also found that the deep learning intelligent fault diagnosis model proposed in this study performs very well in other datasets, basically classifying all types of faults.

#### 4.3. Level III: Diagnostic Performance Verification with Load Variation by Practical Testing

Changes in work load are common for a mechanical system, and when the load changes, the signal measured by the sensor will also change. Under different loads, the number of features in the vibration signal is not the same, and so, the amplitude size is not the same, and the fluctuation period and phase difference are also large. The above situation will cause the classifier to be unable to accurately classify the extracted features, thus reducing the generalization performance of the intelligent fault diagnosis system. In order to verify the diagnostic performance of the model under the changed actual working environment, we built a bearing diagnostic signal acquisition platform, as shown in Figure 16, and put the vibration signal sensor to the DE side for data acquisition. Adjusting the motor speed to 1–3 hp, we recorded the signal with and without the faulty bearing, respectively.



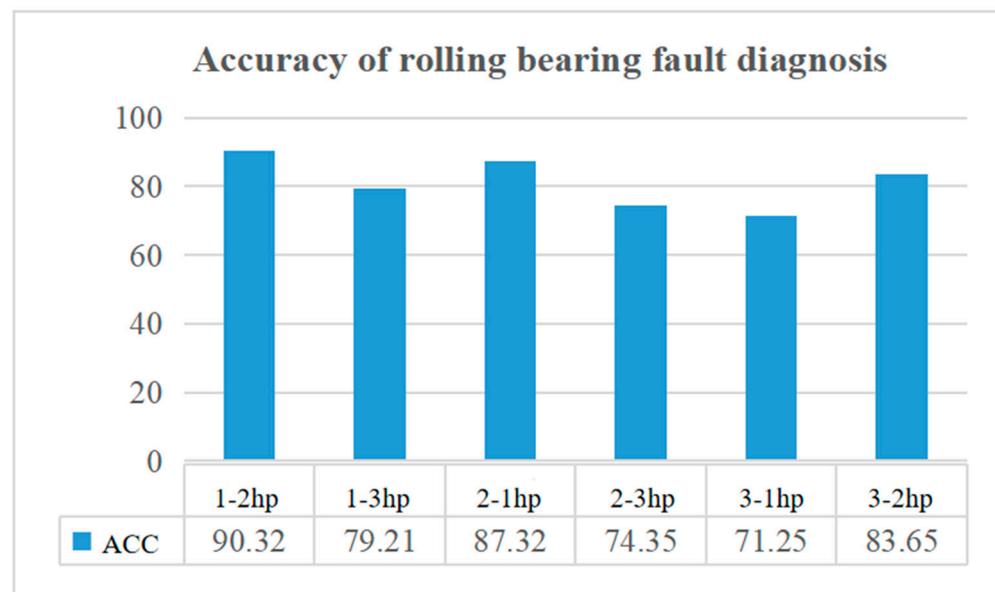
**Figure 15.** Visualization results of t-SNE in other data sets (from an industrial big data competition).



**Figure 16.** Practical rolling bearing fault data acquisition platform.

Although the model achieved good diagnostic results under the condition of constant load, in practice, the load of the bearing was variable. To further verify the generalization ability of the model proposed in this paper, the diagnostic performance under load variation conditions were tested in the practical platform. Training samples with loads of 1 hp, 2 hp, and 3 hp were used to train the model, and the remaining test samples were used to test the generalization ability of the model.

The test results are shown in Figure 17. The model has the highest fault diagnosis accuracy under variable loads of 1-2 hp and 2-1 hp, with an average accuracy close to 90%, and the worst fault diagnosis accuracy under variable loads of 3-1 hp, 2-3 hp, with an average accuracy of about 72.8%. On the whole, although the diagnostic accuracy of the model is poor under the variable operating conditions of high load, the diagnostic accuracy is high under the variable operating conditions of low load, and the overall average accuracy is more than 80%. Therefore, it can be seen that the model can be applied to fault diagnosis under conventional load variation conditions.



**Figure 17.** Fault diagnosis performance of rolling bearings under variable loads.

In this study, to test the nonlinear robustness and generalization capability of the fault classification model, we built a practical test platform and conducted six sets of destructive experiments with variable loads. In actual machine operation, the motor load does not generally drastically fluctuate, so that a generally acceptable bearing will not produce the faults that we want in experiments over a short period of time, even under normal conditions of variable workloads. Since it is difficult to attain a clear indication in the experiment whether a fault has been generated, or what the specific type of fault is, etc., we do not know when to use the sensor to obtain data on the vibration signal. Therefore, the use of good diagnostic model measurements in practice has a certain level of randomness, which makes the actual experimental testing very challenging. Moreover, extreme variable load conditions are difficult to generate in general, and, even under extreme variable load conditions, machine failure does not necessarily immediately occur because the machine will have a certain load carrying capacity.

Through practical experimental tests, we clearly understand that fault diagnosis research under extreme variable load conditions in the study is both significant and a challenge. Therefore, we should consider some randomness issues of the classification model and the practical tests, so as to improve the real performance of the model.

## 5. Conclusions

With the continuous development of fault prognostics and health management (PHM) technology, intelligent fault diagnosis models have emerged. Most of the rolling bearing intelligent diagnosis models based on deep learning have considerable testing accuracy, but there are still problems in terms of insufficient generality and robustness, especially under the working conditions of rapidly changing loads.

In this study, we first explored the generation of rolling bearing faults and some empirical frequency formulas for bearing fault vibration, and concluded the inefficiency and unreliability of traditional fault diagnosis methods. Therefore, an end-to-end hybrid deep neural-network-based model with PCA was constructed and applied to the fault diagnosis of rolling bearings. When the original unclustered data were input to the model, the data were first pre-processed by feature dimensionality reduction, and then features were extracted and denoised by a CNN algorithm in the first layer, followed by a bi-directional LSTM algorithm for feature extraction and memory of time series data. Finally, the attention mechanism was used to improve the weight of different categories of feature data for Softmax classification to improve the accuracy of diagnosis.

The CWRU dataset was used to train and test the model, and the experimental results showed that the highest test accuracy of the model under low load conditions was close to 100%; the overall test accuracy was 99.98%, which surpassed most existing deep learning fault diagnosis models. In order to test the diagnostic accuracy of the model under variable load conditions, six sets of achievable experiments were designed. The experimental results showed that this model still had 72.8% diagnostic accuracy under extreme variable load conditions, and more than 80% diagnostic accuracy under overall variable load conditions, which indicated that the diagnostic model has considerable robustness and versatility.

In terms of next steps, we will design more practical test experiments and sufficiently overcome the randomness caused by model testing to conduct repeated cross-platform tests. Moreover, we will do our best to deploy and improve the intelligent fault diagnosis model proposed in this study.

**Author Contributions:** Software, G.Q.; Writing—review & editing, K.Y.; Supervision, Y.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by National Natural Science Foundation of China (Grant Nos. 61,963,018 and 52165012), Natural Science Foundation of Jiangxi Province in China (Grant Nos. 20212ACB202004).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** Available Based on Request. The datasets generated and/or analyzed during the current study are not publicly available due to extension of the submitted research work, but are available from the corresponding author on reasonable request.

**Conflicts of Interest:** The authors declare that they have no conflict of interest to report regarding the present study.

## References

1. Lanham, C. *Understanding the Tests That Are Recommended for Electric Motor Predictive Maintenance*; Baker Instrument Company: New York, NY, USA, 2002; Energy Publication.
2. Hoang, D.-T.; Kang, H.-J. A survey on Deep Learning based bearing fault diagnosis. *Neurocomputing* **2019**, *335*, 327–335. [[CrossRef](#)]
3. Wu, H.; Li, J.; Zhang, Q.; Tao, J.; Meng, Z. Intelligent fault diagnosis of rolling bearings under varying operating conditions based on domain-adversarial neural network and attention mechanism. *ISA Trans.* **2022**, *130*, 477–489. [[CrossRef](#)] [[PubMed](#)]
4. Zhang, R.; Tao, H.; Wu, L.; Guan, Y. Transfer Learning with Neural Networks for Bearing Fault Diagnosis in Changing Working Conditions. *IEEE Access* **2017**, *5*, 14347–14357. [[CrossRef](#)]
5. Zhang, Z.; Zhou, F.; Li, S. A Cross Working Condition Multiscale Recursive Feature Fusion Method for Fault Diagnosis of Rolling Bearing in Multiple Working Conditions. *IEEE Access* **2022**, *10*, 78502–78518. [[CrossRef](#)]
6. *ISO 15243; Rolling Bearings: Damage and Failures—Terms, Characteristics and Causes*. International Organization for Standardization: Geneva, Switzerland, 2004.
7. Liu, R.; Yang, B.; Zio, E.; Chen, X. Artificial intelligence for fault diagnosis of rotating machinery: A review. *Mech. Syst. Signal Process.* **2018**, *108*, 33–47. [[CrossRef](#)]
8. Ambrozkiwicz, B.; Litak, G.; Georgiadis, A.; Syta, A.; Meier, N.; Gassner, A. Effect of Radial Clearance on Ball Bearing's Dynamics Using a 2-DOF Model. *Int. J. Simul. Model.* **2021**, *20*, 513–524. [[CrossRef](#)]
9. Wang, P.; Yang, Y.; Ma, H.; Hongyang, X.; Xu, L.; Zhong, L.; Bangchun, W. Vibration characteristics of rotor-bearing system with angular misalignment and cage fracture: Simulation and experiment. *Mech. Syst. Signal Process.* **2023**, *182*, 109545. [[CrossRef](#)]
10. Khan, S.; Yairi, T. A review on the application of deep learning in system health management. *Mech. Syst. Signal Process.* **2018**, *107*, 241–265. [[CrossRef](#)]
11. Cerrada, M.; Sánchez, R.-V.; Li, C.; Pacheco, F.; Cabrera, D.; de Oliveira, J.V.; Vásquez, R.E. A review on data-driven fault severity assessment in rolling bearings. *Mech. Syst. Signal Process.* **2018**, *99*, 169–196. [[CrossRef](#)]
12. Zhao, R.; Yan, R.; Chen, Z.; Mao, K.; Wang, P.; Gao, R.X. Deep learning and its applications to machine health monitoring. *Mech. Syst. Signal Process.* **2019**, *115*, 213–237. [[CrossRef](#)]
13. Huo, Z.; Martinez-Garcia, M.; Zhang, Y.; Shu, L. A Multisensor Information Fusion Method for High-Reliability Fault Diagnosis of Rotating Machinery. *IEEE Trans. Instrum. Meas.* **2021**, *71*, 1–12. [[CrossRef](#)]
14. Zhang, S.; Zhang, S.; Wang, B.; Habetler, T.G. Deep Learning Algorithms for Bearing Fault Diagnostics—A Comprehensive Review. In Proceedings of the 2019 IEEE 12th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED), Toulouse, France, 27–30 August 2019; IEEE: Piscataway, NJ, USA, 2019.

15. Soualhi, A.; Taleb, S. Data Fusion for Fault Severity Estimation of Ball Bearings. In Proceedings of the 2018 IEEE International Conference on Industrial Technology (ICIT), Lyon, France, 20–22 February 2018; pp. 2105–2110. [\[CrossRef\]](#)
16. Wu, J.; Wu, C.; Cao, S.; Or, S.W.; Deng, C.; Shao, X. Degradation Data-Driven Time-To-Failure Prognostics Approach for Rolling Element Bearings in Electrical Machines. *IEEE Trans. Ind. Electron.* **2019**, *66*, 529–539. [\[CrossRef\]](#)
17. Hu, Q.; Qin, A.; Zhang, Q.; He, J.; Sun, G. Fault Diagnosis Based on Weighted Extreme Learning Machine With Wavelet Packet Decomposition and KPCA. *IEEE Sens. J.* **2018**, *18*, 8472–8483. [\[CrossRef\]](#)
18. Kang, S.; Cui, L.; Wang, Y.; Li, F.; Mikulovich, V.I. Method of assessing the multi-state of a rolling bearing based on CFOA-HSVM two measures combination. In Proceedings of the 2017 Prognostics and System Health Management Conference (PHM-Harbin), Harbin, China, 9–12 July 2017; pp. 1–5. [\[CrossRef\]](#)
19. Jallepalli, D.; Kakhki, F.D. Data-Driven Fault Classification Using Support Vector Machines. *Intell. Hum. Syst. Integr.* **2021**, *1322*, 316–322. [\[CrossRef\]](#)
20. Nayana, B.R.; Geethanjali, P. Analysis of Statistical Time-Domain Features Effectiveness in Identification of Bearing Faults From Vibration Signal. *IEEE Sens. J.* **2017**, *17*, 5618–5625. [\[CrossRef\]](#)
21. Awadallah, M.; Morcos, M. Application of AI tools in fault diagnosis of electrical machines and drives—an overview. *IEEE Trans. Energy Convers.* **2003**, *18*, 245–251. [\[CrossRef\]](#)
22. Nandi, S.; Toliyat, H.; Li, X. Condition monitoring and fault diagnosis of electrical motors—A review. *IEEE Trans. Energy Convers.* **2005**, *20*, 719–729. [\[CrossRef\]](#)
23. Chow, M.-Y.; Mangum, P.; Yee, S. A neural network approach to real-time condition monitoring of induction motors. *IEEE Trans. Ind. Electron.* **1991**, *38*, 448–453. [\[CrossRef\]](#)
24. Malhi, A.; Gao, R. PCA-Based Feature Selection Scheme for Machine Defect Classification. *IEEE Trans. Instrum. Meas.* **2004**, *53*, 1517–1525. [\[CrossRef\]](#)
25. Xue, X.; Zhou, J. A hybrid fault diagnosis approach based on mixed-domain state features for rotating machinery. *ISA Trans.* **2017**, *66*, 284–295. [\[CrossRef\]](#)
26. Ettetfagh, M.M.; Ghaemi, M.; Asr, M.Y. Bearing fault diagnosis using hybrid genetic algorithm K-means clustering. In Proceedings of the 2014 IEEE International Symposium on Innovations in Intelligent Systems and Applications (INISTA) Proceedings, Alberobello, Italy, 23–25 June 2014; pp. 84–89. [\[CrossRef\]](#)
27. Teotrakool, K.; Devaney, M.J.; Eren, L. Bearing Fault Detection in Adjustable Speed Drives via a Support Vector Machine with Feature Selection using a Genetic Algorithm. In Proceedings of the 2008 IEEE Instrumentation and Measurement Technology Conference, Victoria, BC, Canada, 12–15 May 2008; pp. 1129–1133. [\[CrossRef\]](#)
28. Liu, H.; You, K. Optimization of dewatering process of concentrate pressure filtering by support vector regression. *Sci. Reports* **2022**, *12*, 7135. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Chen, C.; Zhang, B.; Vachtsevanos, G. Prediction of Machine Health Condition Using Neuro-Fuzzy and Bayesian Algorithms. *IEEE Trans. Instrum. Meas.* **2011**, *61*, 297–306. [\[CrossRef\]](#)
30. Li, W.; Zhang, S.; He, G. Semisupervised distance-preserving self-organizing map for machine-defect detection and classification. *IEEE Trans. Instrum. Meas.* **2013**, *62*, 869–879. [\[CrossRef\]](#)
31. Tong, Q.; Cao, J.; Han, B.; Zhang, X.; Nie, Z.; Wang, J.; Lin, Y.; Zhang, W. A Fault Diagnosis Approach for Rolling Element Bearings Based on RSGWPT-LCD Bilayer Screening and Extreme Learning Machine. *IEEE Access* **2017**, *5*, 5515–5530. [\[CrossRef\]](#)
32. Shen, F.; Chen, C.; Yan, R.; Gao, R.X. Bearing fault diagnosis based on SVD feature extraction and transfer learning classification. In Proceedings of the 2015 Prognostics and System Health Management Conference (PHM), Beijing, China, 21–23 October 2015.
33. Zhao, X.; Jia, M. Fault diagnosis of rolling bearing based on feature reduction with global-local margin Fisher analysis. *Neurocomputing* **2018**, *315*, 447–464. [\[CrossRef\]](#)
34. Cococcioni, M.; Lazzerini, B.; Volpi, S.L. Robust Diagnosis of Rolling Element Bearings Based on Classification Techniques. *IEEE Trans. Ind. Inform.* **2012**, *9*, 2256–2263. [\[CrossRef\]](#)
35. Wang, Z.; Zhang, Q.; Xiong, J.; Xiao, M.; Sun, G.; He, J. Fault Diagnosis of a Rolling Bearing Using Wavelet Packet Denoising and Random Forests. *IEEE Sens. J.* **2017**, *17*, 5581–5588. [\[CrossRef\]](#)
36. Yang, T.; Pen, H.; Wang, Z.; Chang, C.S. Feature Knowledge Based Fault Detection of Induction Motors Through the Analysis of Stator Current Data. *IEEE Trans. Instrum. Meas.* **2016**, *65*, 549–558. [\[CrossRef\]](#)
37. He, D.; Li, R.; Zhu, J. Plastic Bearing Fault Diagnosis Based on a Two-Step Data Mining Approach. *IEEE Trans. Ind. Electron.* **2012**, *60*, 1. [\[CrossRef\]](#)
38. Chen, C.; Shen, F.; Yan, R. Topic Correlation Analysis for Bearing Fault Diagnosis Under Variable Operating Conditions. *J. Phys. Conf. Ser.* **2017**, *842*, 12045. [\[CrossRef\]](#)
39. Wei, Z.; Wang, Y.; He, S.; Bao, J. A novel intelligent method for bearing fault diagnosis based on affinity propagation clustering and adaptive feature selection. *Knowl. Based Syst.* **2017**, *116*, 1–12. [\[CrossRef\]](#)
40. Lu, Y.; Xie, R.; Liang, S.Y. Adaptive online dictionary learning for bearing fault diagnosis. *Int. J. Adv. Manuf. Technol.* **2018**, *101*, 195–202. [\[CrossRef\]](#) [\[PubMed\]](#)
41. Janssens, O.; Slavkovikj, V.; Vervisch, B.; Stockman, K.; Loccupier, M.; Verstockt, S.; Van de Walle, R.; Van Hoecke, S. Convolutional Neural Network Based Fault Detection for Rotating Machinery. *J. Sound Vib.* **2016**, *377*, 331–345. [\[CrossRef\]](#)
42. Guo, X.; Chen, L.; Shen, C. Hierarchical adaptive deep convolutionneural network and its application to bearing fault diagnosis. *Measurement* **2016**, *93*, 490–502. [\[CrossRef\]](#)

43. Wen, L.; Li, X.; Gao, L.; Zhang, Y. A New Convolutional Neural Network-Based Data-Driven Fault Diagnosis Method. *IEEE Trans. Ind. Electron.* **2017**, *65*, 5990–5998. [[CrossRef](#)]
44. Zhang, W.; Zhang, F.; Chen, W.; Jiang, Y.; Song, D. Fault State Recognition of Rolling Bearing Based Fully Convolutional Network. *Comput. Sci. Eng.* **2019**, *21*, 55–63. [[CrossRef](#)]
45. Zhuang, Z.; Wei, Q. Intelligent fault diagnosis of rolling bearing using one-dimensional multi-scale deep convolutional neural network based health state classification. In Proceedings of the 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), Zhuhai, China, 27–29 March 2018; pp. 1–6.
46. Guo, S.; Yang, T.; Gao, W.; Zhang, C.; Zhang, Y. An Intelligent Fault Diagnosis Method for Bearings with Variable Rotating Speed Based on Pythagorean Spatial Pyramid Pooling CNN. *Sensors* **2018**, *18*, 3857. [[CrossRef](#)]
47. Qian, W.; Li, S.; Wang, J.; An, Z.; Jiang, X. An intelligent fault diagnosis framework for raw vibration signals: Adaptive overlapping convolutional neural network. *Meas. Sci. Technol.* **2018**, *29*, 095009. [[CrossRef](#)]
48. Shao, H.; Jiang, H.; Zhao, H.; Wang, F. A novel deep autoencoder feature learning method for rotating machinery fault diagnosis. *Mech. Syst. Signal Process.* **2017**, *95*, 187–204. [[CrossRef](#)]
49. Chen, Z.; Li, W. Multisensor Feature Fusion for Bearing Fault Diagnosis Using Sparse Autoencoder and Deep Belief Network. *IEEE Trans. Instrum. Meas.* **2017**, *66*, 1693–1702. [[CrossRef](#)]
50. Abed, W.; Sharma, S.; Sutton, R.; Motwani, A. A Robust Bearing Fault Detection and Diagnosis Technique for Brushless DC Motors Under Non-stationary Operating Conditions. *J. Control. Autom. Electr. Syst.* **2015**, *26*, 241–254. [[CrossRef](#)]
51. Xie, Y.; Zhang, T. Imbalanced Learning for Fault Diagnosis Problem of Rotating Machinery Based on Generative Adversarial Networks. In Proceedings of the 2018 37th Chinese Control Conference (CCC), Wuhan, China, 25–27 July 2018. [[CrossRef](#)]
52. Pan, H.; He, X.; Tang, S.; Meng, F. An improved bearing fault diagnosis method using one-dimensional CNN and LSTM. *J. Mech. Eng.* **2018**, *64*, 443–452.
53. Zhang, W.; Li, C.; Peng, G.; Chen, Y.; Zhang, Z. A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load. *Mech. Syst. Signal Process.* **2018**, *100*, 439–453. [[CrossRef](#)]
54. Li, T.; Zhou, Z.; Li, S.; Sun, C.; Yan, R.; Chen, X. The emerging graph neural networks for intelligent fault diagnostics and prognostics: A guideline and a benchmark study. *Mech. Syst. Signal Process.* **2021**, *168*, 108653. [[CrossRef](#)]
55. Chen, Z.; Yang, J.; Chen, L.; Jiao, H. Garbage classification system based on improved shufflenet v2. *Resour. Conserv. Recycling* **2022**, *178*, 106090. [[CrossRef](#)]
56. Lv, H.; Chen, J.; Pan, T.; Zhang, T.; Feng, Y.; Liu, S. Attention mechanism in intelligent fault diagnosis of machinery: A review of technique and application. *Measurement* **2022**, *199*, 111594. [[CrossRef](#)]
57. Case Western Reserve University (CWRU) Bearing Data Center, December 2018. Available online: <https://engineering.case.edu/bearingdatacenter> (accessed on 11 September 2022).
58. Smith, W.A.; Randall, R.B. Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study. *Mech. Syst. Signal Process.* **2015**, *64*, 100–131. [[CrossRef](#)]