Article

Bearing Fault Diagnosis Method Based on Deep Learning and Health State Division

Lin Shi, Shaohui Su, Wanqiang Wang *, Shang Gao and Changyong Chu

School of Mechanical Engineering, Hangzhou Dianzi University, Hangzhou 310018, China; 222010097@hdu.edu.cn (L.S.); ssbhui@hdu.edu.cn (S.S.); gaoshang@hdu.edu.cn (S.G.); kevin@hdu.edu.cn (C.C.)

* Correspondence: wwq@hdu.edu.cn

Abstract: As a key component of motion support, the rolling bearing is currently a popular research topic for accurate diagnosis of bearing faults and prediction of remaining bearing life. However, most existing methods still have difficulties in learning representative features from the raw data. In this paper, the Xi’an Jiaotong University (XJTU-SY) rolling bearing dataset is taken as the research object, and a deep learning technique is applied to carry out the bearing fault diagnosis research. The root mean square (RMS), kurtosis, and sum of frequency energy per unit acquisition period of the short-time Fourier transform are used as health factor indicators to divide the whole life cycle of bearings into two phases: the health phase and the fault phase. This division not only expands the bearing dataset but also improves the fault diagnosis efficiency. The Deep Convolutional Neural Networks with Wide First-layer Kernels (WDCNN) network model is improved by introducing multi-scale large convolutional kernels and Gate Recurrent Unit (GRU) networks. The bearing signals with classified health states are trained and tested, and the training and testing process is visualized, then finally the experimental validation is performed for four failure locations in the dataset. The experimental results show that the proposed network model has excellent fault diagnosis and noise immunity, and can achieve the diagnosis of bearing faults under complex working conditions, with greater diagnostic accuracy and efficiency.

Keywords: rolling bearing; fault diagnosis; health status division; deep learning; convolutional neural network

1. Introduction

The bearing is one of the most critical parts of rotating machinery. If a fault occurs in use, it will seriously affect the normal operation of mechanical equipment, causing huge losses and disasters. Therefore, bearing fault diagnosis is an essential step for the normal operation of modern rotating machinery. Research on bearing diagnosis methods and fault mechanisms to ensure the normal operation of bearings has always been a key issue of concern for domestic and foreign experts. Nowadays, machinery system health monitoring has stepped into the era of big data, which is manifested in the use of sensors to obtain monitoring sample data, deep learning to accumulate training experience as the main technical means, and intelligent judgment of machinery health status as the ultimate goal, to ensure the reliability of equipment operation and promote efficient production.

At present, bearing fault diagnosis technologies mainly include vibration analysis, acoustic analysis, oil sample analysis, temperature analysis, and voltage current detection [1].

The vibration analysis method is a widely used diagnosis method, in which the vibration signal can be measured online and combined with deep learning technology feature extraction to determine the early failure type of the bearing. However, due to the high cost of data acquisition, the need for storage and transmission technology to be developed, and other reasons, the typical vibration signal fault diagnosis dataset is extremely scarce,
which seriously restricts the theoretical research and engineering application of mechanical equipment health management technology and fault diagnosis.

In the past 20 years, the research on deep learning fault diagnosis of bearing vibration signals has mainly included the following three aspects:

(1) Selection of datasets

The Western Reserve University dataset [2] is one of the most studied datasets for bearing fault diagnosis by many scholars. However, single dataset research also hinders the research of bearing fault diagnosis algorithms. More and more scholars are also working on newer fault diagnosis datasets. The Prognostics and Health Management (PHM) 2012 bearing full-cycle life dataset (FEMTO-ST) [3] is the most used dataset in full-cycle life prediction studies, but the disadvantages of this dataset are that the failure location is not given and the sampling duration is only 0.1 s. The frequency resolution is low and it is not possible to perform fault diagnosis classification studies. The University of Cincinnati bearing dataset [4], which contains the full cycle life and failure location of bearings, is generally used for bearing remaining life prediction studies. It is also used by more and more scholars in the field of fault classification. However, this dataset was only obtained under a single operating condition with constant values of both rotational speed and radial load, and the sample size is small. Moreover, many scholars are discouraged by the long time of disclosure and the large amount of data.

Professor Lei Yaguo’s team in the School of Mechanical Engineering, at Xi’an Jiaotong University, publicly released the XJTU-SY dataset [5], which contains the full life-cycle vibration signals of 15 rolling bearings under three working conditions. The motor speed for condition 1 is 2100 r/min and the radial force is 12 KN; the speed for condition 2 is 2250 r/min and the radial force is 11 KN; the speed for condition 3 is 2400 r/min and the radial force is 10 KN. The dataset is clearly labeled with the failure location of each bearing, which provides data support for the research in the field of PHM and promotes the algorithm research in the field of bearing remaining useful life prediction [6–9]. However, for the bearing failure locations given in this dataset, no scholars have conducted reasonable fault diagnosis classification studies. The experimental platform is shown in Figure 1.

![Figure 1](image-url)

*Figure 1.* Experimental platform (provided by the XJTU-SY dataset [5]).

The two PCB 352C33 unidirectional acceleration sensors in Figure 1 are fixed to the test bearing horizontally and vertically via magnetic bases. A DT9837 portable dynamic
signal collector was used to collect the horizontal and vertical vibration signals from the sensors. The experimental sampling frequency is 25.6 kHz, the sampling interval is 1 min, the sampling duration is 1.28 s, and the data sampled each time are 32,768 time-series vibration signals.

(2) Data preprocessing

Data preprocessing includes signal processing and dimension transformation. For example, Dong Wook Kim et al. [10] studied the effect of data preprocessing methods and super parameters on rolling bearing fault detection accuracy in deep learning. The higher diagnostic accuracy of the 2D image data format of the convolutional neural network was confirmed by one-dimensional and two-dimensional conversion of the data. Hongyu Zhong et al. [11] proposed a combined transfer learning method, which uses continuous wavelet transform to construct the original vibration signal into time–frequency images, and constructed a self-attention light convolution neural network model. The experimental results verify the effectiveness of the transfer learning method. Compared with other regular CNN models, the classification accuracy of this method reaches 99.5% when there are fewer training samples. More importantly, this shows that the transfer learning method has high accuracy while staying lightweight. Although the accuracy of bearing fault diagnosis can be improved by converting the input data of two-dimensional images, the input of image data will greatly increase the network training time.

These continuously improved time–frequency domain-based fault diagnosis methods have been able to extract the fault features in vibration signals well, but these methods require specialized background knowledge and complex signal processing to achieve better diagnostic results, and applying them in complex environments and with large amounts of data would take considerable effort.

(3) Deep learning network model

Traditional machine learning methods such as artificial neural networks (ANNs) [12] and support vector machines (SVMs) [13] have been better applied in bearing fault diagnosis. ANN learns using training mechanical fault information and diagnosis experience, and then expresses the learned fault diagnosis knowledge using connection weights distributed inside the network. SVM is a generalized linear classifier for binary classification of data using supervised learning, which transforms non-separable low-dimensional data into separable high-dimensional data and establishes the optimal separation hyperplane based on kernel functions to satisfy the classification. Compared with the methods based on the application of signal processing alone, the methods based on machine learning have better adaptability and performance. However, early machine learning methods are highly dependent on expert knowledge and manual feature selection, and these shallow machine learning methods have limited representation capability and cannot make full use of massive data, and the bearing fault diagnosis under complex operating conditions still needs to be improved.

In 2006, Ge. Hinton et al. [14] put forward deep learning for the first time, opening the door to the scientific research field of deep learning. The convolution neural network (CNN) is the most representative algorithm of deep learning technology, and the deep learning network model built using the CNN is also the direction that many scholars have continued to explore [15–19]. In 2012 at the ImageNet competition, Krizhevsky et al. [20] proposed the AlexNet large convolutional neural network. This network reduced the top-5 error rate to 15.3% and began the boom in deep learning techniques. In recent years, inspired by the design idea of AlexNet, many researchers have added activation functions, Dropout [21], Batch Normalization (BN), and other techniques based on the CNN to enhance the strong nonlinear feature extraction ability of the neural network, and many representative convolutional neural networks such as VggNet and GoogleNet have appeared [22].

In 2016, Kaiming He et al. [23] proposed the residual convolutional neural network (ResNet). This network won first place in several tracks in ILSVRC and COCO 2015 competitions: ImageNet detection, ImageNet localization, COCO detection, and COCO
segmentation. This network can add jump connections to deeper network layers at any network layer, which avoids the loss of data feature information and overfitting phenomenon. This greatly increases the depth of the network model, which leads to more superior network models.

A recurrent neural network (RNN) is also one of the representative algorithms of deep learning. The network structure is connected in a ring with a node orientation, and the internal state can display dynamic timing behavior. It is often applied to the processing of timing information such as audio and text. However, since RNN encounters the problem of gradient disappearance when processing long time series information, the network has only short-term memory. The Long Short-Term Memory (LSTM) network improved by Alex Graves [24] effectively solves the problem of short memory, but the LSTM network is more complicated. Cho et al. [25] proposed the Gate Recurrent Unit (GRU) network structure. The network simplifies the LSTM network structure and has long-term memory.

Currently, in the field of rolling bearing fault diagnosis research, deep learning is widely used for research related to mining time series of vibration signal data.

Wenglang Xie et al. [26] proposed a hybrid model based on CNN and individual classifiers to diagnose bearing faults. Experiments have verified that random forest (RF) and support vector machine (SVM) can make full use of the feature extraction ability of CNN. The average diagnosis accuracy of the CNN-RF model and CNN-SVM model on the large-scale dataset is 98.9% and 99%, respectively. YuXia et al. [27] proposed a new multi-source TL model, which uses feature learners to generate features of each source domain and target domain data, so that the joint weight classifier can predict target tags. A distance metric based on moment matching is also introduced to reduce the distance between all source domains and target domains. Experimental results, such as high diagnostic accuracies of 99.96%, support the reliability and universality of the proposed model. Ruixin Wang et al. [28] proposed a depth feature reinforcement learning method for rolling bearing fault diagnosis. Using the Elu activation function and attention mechanism model, they established a depth Q network to accurately diagnose the fault mode. The test accuracy of the proposed method was the highest, and the average test accuracy reached 98.71%, which shows that this method is superior to other intelligent diagnosis methods. Jun Li et al. [29] proposed a rolling bearing fault diagnosis model that combines a recursive neural network based on two-stage attention and a convolutional block attention module. In the experimental test of CWRU, results indicate that the accuracy of the proposed fault diagnosis method DARNN-CBAM-CNN for rolling bearings is 97.69% and the proposed fault diagnosis method has broad application prospects under the condition of unbalanced data. Jiangquan Zhang et al. [30] proposed an intelligent diagnosis algorithm based on CNN, which can automatically accomplish the process of the feature extraction and fault diagnosis. Zhibo Li et al. [31] proposed a fault diagnosis method based on the fusion of deep learning with a knowledge graph. Compared with the deep learning models such as Resnet and Inception in the noise environment of multiple working conditions, the model proposed in this paper not only shows a faster convergence speed and stable performance, but also a higher accuracy in evaluation indicators. Gaowei Xu et al. [32] proposed a novel bearing fault diagnosis method based on deep CNN and random forest ensemble learning. Jiqiang Zhang et al. [33] proposed a novel bearing fault diagnosis method based on deep separable convolution and spatial dropout regularization.

Based on the previously proposed problem of dataset selection, this paper conducts a health state classification study based on the XJTU-SY bearing dataset as the data basis for bearing fault diagnosis. For the health state classification method of bearings, domestic and foreign scholars have also conducted relevant research.

The most commonly used method to classify the health state is the observation method, and Wei Xipeng [34] classified the bearing operating health state at different stages by observing the vibration signal. Although this is the simplest and most intuitive method, the error of the division is large. Lin Feiting [35] used HHT to build the power spectral density of the bearing signal, used quartic polynomial fitting, and derived the
fitting curve. The two inflection points of the curve are taken as the multi-stage dividing points of bearing health. Yin Aijun et al. [36,37] constructed different types of health factors for deepstep feature extraction of vibration signals, and used the 3o principle to classify the bearing health state by observing the sudden change points in the smooth phase and the smooth phase signals within the threshold value. Some scholars also used the HMM model to divide the bearing health status into two stages and multi-stage processes [38–40].

The feature processing capability, adaptability, computational efficiency, and interpretability of existing deep learning methods in bearing fault diagnosis still need to be improved, and the corresponding research still needs to be perfected.

Based on the problems summarized above and the current status of the research, the research idea of this paper is proposed. To better expand the dataset, this paper performs health state segmentation on the XJTU-SY bearing dataset as the database for bearing fault diagnosis. In this paper, the one-dimensional signal characteristics of the data are retained in data preprocessing, and two network models, CNN and RNN, are used as the core, and Dropout, BN, and other techniques are introduced to jointly solve the bearing fault diagnosis problem. Under the condition of ensuring the diagnostic accuracy of the model, the noise resistance performance of the model is studied by adding noises with different SNRs to simulate different noise environments of actual industrial scenes. By comparing with other algorithmic models, it is verified that the method proposed in this paper has strong fault diagnosis and anti-noise capability.

The rest of this paper is organized as follows: Section 2 introduces the proposed bearing health status division method. Section 3 introduces the proposed deep learning network model. Section 4 is the experimental verification part of the second and third sections. Finally, conclusions are given in Section 5.

2. Signal Division Method of Bearing Health State

Bearing operation usually goes through a stable period, which means it is in the healthy stage, and after fault occurrence time (FOT) the degradation level becomes severe and the component is in the fault stage. Health state division divides the bearing’s whole life-cycle signal into different phases, which can usually be divided into the health phase and the fault phase. The result of the division expands the sample dataset for fault diagnosis, and the corresponding classification study of bearing fault diagnosis can be carried out.

2.1. Health Indicator Selection

The health indicator (HI) can characterize the degradation status information of the bearing and is an important indicator to evaluate the health status of the bearing. For health status classification, it is essential to construct HI curves that can accurately characterize the health status of bearings from the monitoring data of mechanical equipment.

To further reveal the degradation characteristics of the bearing, the time-domain characteristics are extracted from the bearing monitoring signal as HI. Some time-domain characteristics of the bearing vibration signal are shown in Table 1. Kurtosis is particularly sensitive to shock signals, and early bearing failures mainly originate from the action of alternating shock loads, so it is particularly suitable for early fault diagnosis. Furthermore, the advantage of the root mean square (RMS) is better stability. Thus, the kurtosis factor and RMS were initially selected as HI.
Table 1. Time-domain characteristics of bearing vibration signal.

<table>
<thead>
<tr>
<th>Dimensional Feature</th>
<th>Equation</th>
<th>Dimensionless Feature</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_{i}(t)$</td>
<td>pulse factor</td>
<td>$C_{pf} = \frac{F}{</td>
</tr>
<tr>
<td>Peak values</td>
<td>$F = \max</td>
<td>x_{i}(t)</td>
<td>$</td>
</tr>
<tr>
<td>root mean square</td>
<td>$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^{N} x_{i}^2(t)}$</td>
<td>Peak factor</td>
<td>$C_{pf} = \frac{F}{RMS}$</td>
</tr>
<tr>
<td>Population variance</td>
<td>$s^2 = \frac{1}{N} \sum_{i=1}^{N} (x_{i} - \bar{x})^2$</td>
<td>kurtosis factor</td>
<td>$C_{kd} = \frac{\frac{1}{N} \sum_{i=1}^{N} x_{i}^4(t)}{RMS^4}$</td>
</tr>
</tbody>
</table>

In the table, N is the number of samples and $x_{i}(t)$ is the value of the data at moment $t$.

To take into account both the time-domain and frequency-domain characteristics of the bearing signal, the sum of the frequency energy per unit acquisition period of the short-time Fourier transform (STFT SUM) is selected as the health factor in the frequency domain.

The window function is chosen as the Hemming window, the window length is set to 32,768, and the overlap is set to 0. From the XJTU-SY dataset, it is known that the data volume of the signal collected every 1.28 s is 32,768, and the STFT SUM index of the signal collected every minute can be obtained by summing its amplitude.

Based on the above-selected health factors, signal analysis was performed on the XJTU-SY dataset. The signal characteristics reflecting the health factors are shown in Figure 2. The RMS value starts to rise in the degradation stage and reaches the threshold value quickly in the damage stage. The kurtosis factor is suitable for the diagnosis of early failure, and the kurtosis indicator has a significant fluctuation of the trend in the health phase, which can easily misjudge the normal health phase as the fault phase. Compared with the other two health indicators, the kurtosis factor is more influenced by external noise and cannot reflect the degradation trend of bearing performance well.

![Figure 2. Signal analysis of bearing HI.](image-url)
2.2. Evaluation of Bearing HI

To select a reasonable health indicator, the trendability, monotonicity, and robustness are introduced to quantitatively evaluate the health indicator.

(1) Trendability

Over time, the evolution process of bearing degradation becomes more and more serious, so the HI curve representing the development process of bearing degradation should show a certain time correlation. The characteristics of this curve are defined as a trend, and the equation is as follows:

$$T_{re}(H, T) = \frac{\left| \sum_{k=1}^{K} (h_k - \bar{H})(t_k - \bar{T}) \right|}{\sqrt{\sum_{k=1}^{K} (h_k - \bar{H})^2 \sum_{k=1}^{K} (t_k - \bar{T})^2}}$$  \hspace{1cm} (1)

In Equation (1), $h_k$ is the sampling period HI value at $t_k$ time; $\bar{H} = (1/K) \sum_{k=1}^{K} h_k$ is the average value of health indicators in the whole life cycle; $\bar{T} = (1/K) \sum_{k=1}^{K} t_k$ is the mean value of each sampling period. $0 \leq T_{re}(H, T) \leq 1$; the closer the value to 1, the better the trend.

(2) Monotonicity

The monotonicity of the HI curve reflects the degree of degeneration of the HI curve. Although the bearing operation is in a stable stage, the bearing has also degraded over time, but the degree of degradation is weak. The monotonicity is calculated as follows:

$$Mon = \left| \frac{\text{num of } dF > 0}{T - 1} - \frac{\text{num of } dF < 0}{T - 1} \right|$$  \hspace{1cm} (2)

$T$ is the number of sampling points in the bearing cycle; $dF$ is the differential between the front and rear values in the HI curve.

(3) Robustness [41]

For the characteristic signal sequence $F = [f(1), f(2), \cdots f(K)]$, the time sequence $T = [t(1), t(2), \cdots t(K)]$, $f(t_k)$ represents the characteristic value obtained at the time $t_k$, where $k = 1, 2, \cdots, K$, $K$ represents the length of time. First, the Exponential Weighted Moving Average (EWMA) is used to divide the feature sequence into two parts, namely, the stationary trend term $f_T(t_k)$ and the random complementary vector $f_R(t_k)$.

$$f(t_k) = f_T(t_k) + f_R(t_k)$$  \hspace{1cm} (3)

EWMA is calculated as follows:

$$f_T(t_k) = \beta f_T(t_{k-1}) + (1 - \beta)f(t_k)$$  \hspace{1cm} (4)

Equation (4) generally takes $\beta \geq 0.9$ and $f_T(t_k)$ can be calculated by averaging the previous values.

Robustness is the tolerance to outliers and measures the effect of possible random fluctuations in the bearing degradation process due to random changes in sensor noise. The robustness evaluation index of $F$ is denoted $\text{Rob}(F)$:

$$\text{Rob}(F) = \frac{1}{K} \sum_{k} \exp \left( - \frac{|f(k) - f_T(k)|}{f(k)} \right)$$  \hspace{1cm} (5)

Based on the XJTU-SY dataset, the health indicators of 15 bearing vibration signals under 3 working conditions were evaluated respectively, and the results are shown in Table 2. It can be seen that the calculation results of RMS indicators are good, and most parameters are higher than the kurtosis factor and STFT SUM. Therefore, the paper selected RMS as the HI of bearing health status division.
Table 2. HI evaluation results.

<table>
<thead>
<tr>
<th></th>
<th>RMS</th>
<th>Kurtosis Factor</th>
<th>STFT SUM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tre</td>
<td>Mon</td>
<td>Rob</td>
</tr>
<tr>
<td>Bearing1_1</td>
<td>0.8632</td>
<td>0.6393</td>
<td>0.7908</td>
</tr>
<tr>
<td>Bearing1_2</td>
<td>0.9014</td>
<td>0.2625</td>
<td>0.7888</td>
</tr>
<tr>
<td>Bearing1_3</td>
<td>0.7560</td>
<td>0.6306</td>
<td>0.7775</td>
</tr>
<tr>
<td>Bearing1_4</td>
<td>0.3625</td>
<td>0.0083</td>
<td>0.7053</td>
</tr>
<tr>
<td>Bearing1_5</td>
<td>0.7584</td>
<td>0.4118</td>
<td>0.6760</td>
</tr>
<tr>
<td>Bearing2_1</td>
<td>0.3731</td>
<td>0.0449</td>
<td>0.7953</td>
</tr>
<tr>
<td>Bearing2_2</td>
<td>0.8838</td>
<td>0.2625</td>
<td>0.7798</td>
</tr>
<tr>
<td>Bearing2_3</td>
<td>0.8821</td>
<td>0.4812</td>
<td>0.8637</td>
</tr>
<tr>
<td>Bearing2_4</td>
<td>0.7909</td>
<td>0.7561</td>
<td>0.6504</td>
</tr>
<tr>
<td>Bearing2_5</td>
<td>0.9027</td>
<td>0.4911</td>
<td>0.8138</td>
</tr>
<tr>
<td>Bearing3_1</td>
<td>0.3207</td>
<td>0.0445</td>
<td>0.8944</td>
</tr>
<tr>
<td>Bearing3_2</td>
<td>0.7047</td>
<td>0.0790</td>
<td>0.8258</td>
</tr>
<tr>
<td>Bearing3_3</td>
<td>0.4141</td>
<td>0.2703</td>
<td>0.8358</td>
</tr>
<tr>
<td>Bearing3_4</td>
<td>0.3575</td>
<td>0.0502</td>
<td>0.8875</td>
</tr>
<tr>
<td>Bearing3_5</td>
<td>0.8719</td>
<td>0.4336</td>
<td>0.8320</td>
</tr>
</tbody>
</table>

2.3. Health State Division Method

The key to the classification of bearing health status signals is to identify the fault occurrence time (FOT). The bearing signal typically goes through a stable phase, when the bearing is in the healthy stage. After FOT, the bearing degradation becomes severe and is in the fault stage.

At the time of failure, the health factor will produce an elbow-point mutation due to the transition from the smooth phase to the failure phase. Therefore, in this paper, within the threshold condition of health factor 0.1, the range of abrupt change in health factor under the threshold condition is first observed to be determined. After that, the first-order differentiation of the RMS health factor within the range is performed to obtain the abrupt change condition of the RMS first-order differentiation point, to determine the health state of the bearing. The principle of bearing health state division is shown in Figure 3.
2.4. Health Status Division Results

Based on the bearing vibration signal in the XJTU-SY dataset, the health status is divided, and the division results are shown in Figure 4. The long red line in the figure shows the boundary line of the healthy stage. It can be seen that the vibration signal tends to be stable in the healthy stage on the left side of the bearing vibration signal. After the failure point, the characteristic amplitude of the vibration signal shows an increasing trend with the running time, which accurately reflects the bearing degradation process.

![Figure 4. Result of bearing health status division.](image-url)
Taking Working Condition 1 as an example, the bearing under Working Condition 1 is divided into four types of bearing states: normal, cage fault, outer-race fault, and mixed fault of inner and outer race.

Bearing normal data is the health status data before the FOT point of bearing1_1 to bearing1_5; outer ring failure data is the degradation data after the FOT point of bearing1_1 to bearing1_3; inner ring failure data is the degradation data after the FOT point of bearing1_4, and mixed inner and outer ring failure data is the degradation data after the FOT point of bearing1_5.

Since the vibration signal at the fault occurrence point is at the critical value of the fault and normal signals, and the signal points collected within 1.28 s have not been further divided, to use more reliable data, this paper discards the vibration signal at the bearing health fault occurrence point. The specific bearing health status division results of the three working conditions are shown in Table 3.

<table>
<thead>
<tr>
<th>Fault Location</th>
<th>Bearing1_1</th>
<th>Bearing1_2</th>
<th>Bearing1_3</th>
<th>Bearing1_4</th>
<th>Bearing1_5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>1:76</td>
<td>1:43</td>
<td>1:56</td>
<td>1:77</td>
<td>1:32</td>
</tr>
<tr>
<td>Cage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inner-Outer-race</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outer-race</td>
<td>78:123</td>
<td>45:161</td>
<td>58:158</td>
<td></td>
<td>34:52</td>
</tr>
<tr>
<td>Bearing2_1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>1:449</td>
<td>1:44</td>
<td>1:322</td>
<td>1:28</td>
<td>1:118</td>
</tr>
<tr>
<td>Cage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outer-race</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bearing3_1</td>
<td>1:2376</td>
<td>1:912</td>
<td>1:338</td>
<td>1:1414</td>
<td>1:4</td>
</tr>
<tr>
<td>Normal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cage-Outer-race</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>914:2496</td>
</tr>
<tr>
<td>Inner-race</td>
<td></td>
<td></td>
<td></td>
<td>340:371</td>
<td>1416:1515</td>
</tr>
<tr>
<td>2378:2538</td>
<td>6:114</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. Bearing Fault Diagnosis Based on Deep Learning

In this section, according to the signal processing technology, the original signal is processed by fast Fourier transform (FFT) to extract the characteristics of the bearing vibration signal.

By building a deep learning network model, we study the bearing fault diagnosis method based on the bearing health state classification data.

3.1. FFT Feature Extraction of Time-Domain Signal

Firstly, the FFT feature of the original time-domain vibration signal is extracted, and the extracted FFT feature information is input into the neural network. The Fourier transform equation is as follows:

\[
F(w) = \int_{-\infty}^{\infty} f(t) \cdot e^{iwt} \, dt 
\]

In Equation (6), \( F(w) \) is the Fourier transform of \( f(t) \) and \( w \) is the frequency.

Because the frequency-domain features extracted by the Fourier transform are symmetrical, extracting the time-domain vibration signal into the frequency domain changes can reduce the amount of data by half. Because the maximum amplitude values in the frequency domain are different, the paper uniformly processes the frequency-domain
signal divided by the sample length. FFT transformation of the time-domain vibration signal is shown in Figure 5.

![Vibration signal and FFT transform](image)

**Figure 5.** FFT of bearing vibration signal: (a) Vibration signal; (b) FFT transform.

### 3.2. Network Model of Bearing Fault Diagnosis Based on Deep Learning

Zhang Wei [42] of Harbin University of Technology designed a model known as “Deep Convolutional Neural Networks with Wide First layer kernel” (WDCNN) based on the characteristics of one-dimensional signal vibration.

Its structural feature is that the first layer is a wide convolution kernel. In the WDCNN network model, 64 × 1 feature extraction is used for the first layer of the large convolution kernel, and a 3 × 1 small convolution kernel and a pooling layer are used for the rest for further feature extraction.

In the WDCNN network model, the first layer of a large convolutional kernel using a single 64 × 1 convolutional kernel inevitably loses information in downsampling, so the first layer of the large convolutional kernel of different sizes is used in the original paper to verify the reliability of the model.

To address this feature, this paper introduces multi-scale large convolutional kernels for the first layer of large convolutional kernels based on the WDCNN model. In this paper, four different sizes of convolutional kernels, 16 × 1, 32 × 1, 64 × 1, and 96 × 1, are introduced in the first layer of a 64 × 1 large convolutional kernel to perform further feature extraction for sample information of different lengths. The structure of the first layer of the main network is shown in Figure 6.

![Multi-scale first-layer wide convolution kernel](image)
3.3. Network Model and Detailed Parameters

The bearing vibration signal sequence has the time correlation property, and the recycle neural network has good time correlation sequence processing ability. Therefore, this paper introduces a GRU recurrent neural network to process the sequence features extracted from the convolutional layer. By combining the convolutional neural network and GRU network, the time-dependent sequences can be processed more efficiently. The utilization of the features can be improved by automatically extracting the intrinsic features of the signal using the convolutional neural network. The GRU network can then enable further processing of the features to improve the network’s ability to process time-correlated sequences. The approach combining the advantages of both neural networks can increase the ability of the network to cope with bearing fault signals in complex situations, especially strong noise situations [43].

The network model proposed in this paper is shown in Figure 7. After the time-domain signal is extracted by FFT features, the first layer of the multi-scale large convolutional kernel, multi-layer 3 × 1 convolutional layer and 2 × 1 pooling layer, and GRU recurrent neural network are then used for feature extraction to further classify the rolling bearing state.

Figure 7. Structure of the network model.

The detailed network parameters are shown in Table 4, with a Relu activation function layer following each convolutional layer. The purpose of introducing the activation function is to make the otherwise linear model nonlinear, allowing the model to handle linearly indistinguishable problems.

To suppress neural network overfitting, Dropout is introduced in this paper after the first large convolutional kernel and GRU layer of the multi-scale. Usually, Dropout is artificially set to 0.5 or 0.3, so the probability of lost neurons $p$ is set to 0.3 for all Dropout layers, and the $l_2$ regularization factor is introduced to $10^{-4}$ in each convolutional and GRU layer.
3.4. Experimental Platform and Technology

The realization of the model training in this paper adopted the Tensorflow2.2_GPU version deep learning framework based on Python3.7, and Pycharm was used for the code editing. The experimental environment was a computer with AMD R5 3500X CPU, GTX 1660s GPU, 256 G system memory, and 16 GB running memory under the Win10 system.

To improve the efficiency of the experiment and preserve the optimal model parameters of the network, Early Stopping and Save Best Only techniques are used. The Early Stopping technique ensures that the training process is terminated early when the validation accuracy no longer increases, shortening the model training time and improving the training efficiency, while the Save Best Only technique saves the model with the best performance throughout the training cycle and avoids saving the degraded model. In the experiment, the training period is set to 300 epochs, and the Early Stopping technique takes the loss function as the monitor. If the validation set loss function does not degrade for 100 consecutive epochs, the Save Best Only technique is used to save the current network model parameters. The batch_size is 256 of the original WDCNN network setting.

3.5. Experimental Process

The experimental and algorithmic flow chart is shown in Figure 8. Firstly, the divided training, validation, and test datasets are pre-processed by FFT as the input of the neural network. The training set is disrupted and divided into several consecutive small batches of data (Mini Batches), each containing 256 sample data. The FFT processed training samples are then input to the neural network model, and feature extraction is performed by the first layer of the multi-scale large convolutional kernel followed by multiple small convolutional layers and pooling layers. Finally, the feature sequence information after the convolutional layer is processed by the GRU recurrent neural network, and the corresponding fault classification labels are output by the four neurons of the fully connected layer.

The model training uses a cross-entropy loss function and Adam optimizer for gradient update of output samples. The loss function of the validation set is used as the monitor, and if the loss function of the validation set does not decrease for 100 consecutive
epochs, the model with the smallest loss function is used as the optimal model. After that, the test set data are input to the optimal model and the fault diagnosis classification accuracy is output.

4. Experimental Verification

4.1. Experimental Data Processing and Experimental Process Design

4.1.1. Data Enhancement

The data enhancement method proposed in this section uses the overlapping sampling method, i.e., for the training samples, each segment of the signal is acquired from the original signal with an overlap between its subsequent segments, as shown in Figure 9. For the test samples, there is no overlap in the acquisition, and the offset is set to 28 in this paper.
4.1.2. Division of Experimental Dataset

During the experiment, the first 70% of the bearing vibration signals with divided health status are set as the training set and the last 30% as the test and validation set, as shown in Figure 10.

Table 5. Description of the experimental dataset.

<table>
<thead>
<tr>
<th>Working Condition I</th>
<th>Fault Location</th>
<th>Cage</th>
<th>Inner-Outer</th>
<th>Normal</th>
<th>Outer-Race</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type Label</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Bearing</td>
<td>1_4</td>
<td>1_5</td>
<td>1_1</td>
<td>1_2</td>
<td>1_3</td>
</tr>
<tr>
<td>Training Set</td>
<td>7000</td>
<td>7000</td>
<td>1400</td>
<td>1400</td>
<td>1400</td>
</tr>
<tr>
<td>Validation Set</td>
<td>1000</td>
<td>1000</td>
<td>200</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>Test Set</td>
<td>2000</td>
<td>2000</td>
<td>400</td>
<td>400</td>
<td>400</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Working Condition II</th>
<th>Fault Location</th>
<th>Inner-race</th>
<th>Outer-race</th>
<th>Normal</th>
<th>Cage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type Label</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Bearing</td>
<td>2_4</td>
<td>2_5</td>
<td>2_1</td>
<td>2_2</td>
<td>2_3</td>
</tr>
<tr>
<td>Training Set</td>
<td>7000</td>
<td>7000</td>
<td>1400</td>
<td>1400</td>
<td>1400</td>
</tr>
<tr>
<td>Validation Set</td>
<td>1000</td>
<td>1000</td>
<td>200</td>
<td>200</td>
<td>200</td>
</tr>
</tbody>
</table>

In the whole-life degradation process of bearings, the strong degradation characteristics of the later stage of bearings are diagnosed through the early weak degradation characteristics. The different distribution of data can better reflect the generalization ability of the bearing diagnosis model. Furthermore, in the actual bearing operation process, fault diagnosis of different late health statuses is more in line with the actual operation of the bearing under health management because it can reduce unnecessary replacement maintenance costs.

In the experiments, 10,000 samples were taken for each state type in the three divided bearing datasets. The Case 1 dataset includes a total of 27,999 training data samples, 4000 validation data samples, and 8000 test data samples. Case 2 is the same as Case 1. The training set of Case 3 includes 28,000 training samples, and the data distribution of the specific division is shown in Table 5.
4.2. Analysis of Experimental Results and Visualization of Training Classification Process

In this section, five sets of comparison networks are introduced to verify the feasibility of the method proposed in this paper through experimental results. The five sets of comparison network models are as follows: WDCNN + original vibration data, WDCNN + FFT signal processing, Propose + original vibration data, SVM + FFT (where SVM uses Gaussian kernel function and the error penalty term coefficient is taken as 1), and ANN + FFT (the number of neurons in each layer is 1000, 500, 300, 100, 50, and each layer uses a RELU activation function and l2 regularization). The first three sets of comparison networks are used as self-comparison experiments of the proposed network model. This paper presents a comparison with the original WDCNN network, which was the leading deep convolutional network-based bearing fault diagnosis method at that time, containing five convolutional layers and one fully connected layer. ANN and SVM are two traditional machine learning methods used for bearing fault diagnosis. ANN learns by training mechanical fault information and diagnosis experience, and then expresses the learned fault diagnosis knowledge using connection weights distributed inside the network. SVM has the advantage of solving small samples, nonlinear data, and strong generalization ability.

To avoid the randomness of the experiments and to ensure the credibility of the experimental comparisons, 10 repetitions of the experiments were undertaken for each model. Through the data enhancement technique, each experiment randomly grabs the divided bearing dataset under three working conditions to constitute different training sets, validation sets, and test sets to verify the reliability of the model and eliminate the influence of experimental randomness. The mean and standard deviation of the 10 experimental results were taken as the error range of the experimental results. The results are shown in Figure 11.

![Image of diagnostic results comparison](image)

**Figure 11.** Comparison of diagnostic results of different models.
It can be seen from Figure 11 that the network accuracy reflected by the same network model is very different under different working condition data. FFT signal processing combined with the network model proposed in this paper (Propose + FFT) under the three working conditions obtained an average diagnostic accuracy of 96.969%, 97.846%, and 97.904% higher than other models. The standard deviation was also significantly smaller than other models, which reflects the stability of the proposed model.

During the experiments, both the correct and loss rates of the model almost stabilized in the training set and the validation set, and the fit was good. The accuracy of the validation set was maintained at about 90%, as shown in Figure 12.

![Figure 12](image)

**Figure 12.** The trend of accuracy and the loss function of the model training process under three working conditions: (a) Working condition I; (b) Working condition II; (c) Working condition III.

After the model training is completed, the test set is used to verify the model. The confusion matrix results are shown in Figure 13, and its values of 0, 1, 2, and 3 correspond to the fault type labels in Table 5. It can be seen that the prediction results are almost all correct. The real category 3 label of Working Condition 3 is the mixed fault of the inner and outer race rolling element, and 6% of the samples are predicted to be the inner-race fault of label 0, which is tolerable in the actual bearing fault diagnosis.

![Figure 13](image)

**Figure 13.** Test set results of three working conditions’ confusion matrix: (a) Working condition I; (b) Working condition II; (c) Working condition III.

4.3. **Analysis of Model Noise Resistance Results**

This section is designed to validate the noise immunity of the model, especially for additive Gaussian white noise, since this noise is one of the most representative noises and is easy to quantify. In this paper, the strength of the bearings subjected to industrial environmental noise is simulated by adding different signal-to-noise ratios (SNRs) [44].
The SNR is defined in Equation 7, and the unit is usually decibels (dB), where a smaller SNR indicates a more contaminated signal.

\[ SNR_{dB} = 10 \log_{10} \left( \frac{P_{\text{signal}}}{P_{\text{noise}}} \right) \] (7)

In Equation (7), \( P_{\text{signal}} \) is the original signal power and \( P_{\text{noise}} \) is the added noise power.

As shown in Figure 14, the original vibration signal has been completely distorted compared with the original signal after adding Gaussian white noise with \( SNR = 0 \) dB.

![Figure 14. Schematic diagram of adding noise signal.](image)

White Gaussian noise with a signal-to-noise ratio of −10 dB to 10 dB was added to the test set, and then the test set data were input into the saved network model. The diagnostic accuracy of the model is shown in Table 6. The table shows the mean values of ten diagnostic accuracies and the range of standard deviations of ten diagnostic accuracies for different models under three working conditions.

<table>
<thead>
<tr>
<th>SNR Model</th>
<th>-10</th>
<th>-8</th>
<th>-6</th>
<th>-4</th>
<th>-2</th>
<th>0</th>
<th>2</th>
<th>4</th>
<th>6</th>
<th>8</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM + FFT</td>
<td>25.76</td>
<td>27.11</td>
<td>35.60</td>
<td>48.16</td>
<td>59.66</td>
<td>66.93</td>
<td>75.50</td>
<td>83.07</td>
<td>85.29</td>
<td>87.50</td>
<td>88.43</td>
</tr>
<tr>
<td>±</td>
<td>0.986</td>
<td>2.315</td>
<td>6.977</td>
<td>7.123</td>
<td>3.325</td>
<td>3.083</td>
<td>3.473</td>
<td>1.896</td>
<td>2.073</td>
<td>1.598</td>
<td>1.503</td>
</tr>
<tr>
<td>ANN + FFT</td>
<td>38.24</td>
<td>40.54</td>
<td>43.10</td>
<td>48.78</td>
<td>60.43</td>
<td>69.73</td>
<td>72.94</td>
<td>75.00</td>
<td>79.77</td>
<td>83.07</td>
<td>85.16</td>
</tr>
<tr>
<td>WD + Ori.</td>
<td>29.15</td>
<td>32.71</td>
<td>37.95</td>
<td>44.13</td>
<td>54.82</td>
<td>65.83</td>
<td>76.90</td>
<td>82.63</td>
<td>86.77</td>
<td>88.38</td>
<td>90.42</td>
</tr>
<tr>
<td>WD + FFT</td>
<td>30.98</td>
<td>32.31</td>
<td>38.55</td>
<td>40.26</td>
<td>49.33</td>
<td>60.74</td>
<td>71.00</td>
<td>80.70</td>
<td>84.82</td>
<td>87.32</td>
<td>88.84</td>
</tr>
<tr>
<td>Pro + Ori.</td>
<td>26.14</td>
<td>27.65</td>
<td>30.84</td>
<td>37.56</td>
<td>49.83</td>
<td>69.19</td>
<td>79.22</td>
<td>86.08</td>
<td>90.36</td>
<td>92.78</td>
<td>93.72</td>
</tr>
<tr>
<td>Pro + FFT</td>
<td>40.17</td>
<td>46.24</td>
<td>51.94</td>
<td>56.98</td>
<td>64.67</td>
<td>72.13</td>
<td>79.55</td>
<td>88.16</td>
<td>90.77</td>
<td>92.67</td>
<td>94.67</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Working condition II</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM + FFT</td>
</tr>
<tr>
<td>±</td>
</tr>
<tr>
<td>ANN + FFT</td>
</tr>
<tr>
<td>WD + Ori.</td>
</tr>
<tr>
<td>±</td>
</tr>
</tbody>
</table>
According to the data in the table, the visualization curve of the average accuracy and error range of the same model under three working conditions is constructed, as shown in Figure 15. The experimental results show that the diagnostic accuracy of the noise immunity performance of FFT signal processing combined with the proposed network (Propose + FFT) model is significantly higher than that of the other five algorithm models.

In Working Condition 1, the performance of the Propose + FFT model in the low-noise environment is not much different from that of the Propose + original vibration signal model. As the noise intensity intensifies, the diagnostic accuracy of the Propose + original vibration signal model declines sharply. In contrast, the Propose + FFT model still maintains a high diagnostic accuracy, indicating that the proposed model is also suitable for strong noise environments. The diagnostic accuracy of the ANN + FFT model is low in low-noise environments. In the environment of strong noise $SNR = -10$ dB, although the diagnostic accuracy is higher than the other four models, it is still lower than that of the Propose + FFT model.

In Working Condition 2, the average accuracy of the Propose + FFT model is significantly higher than that of the other five models. Different from Working Condition 1 and Working Condition 3, the accuracy of the WDCNN+FFT model, except being lower than the proposed Propose + FFT algorithm, has higher diagnostic accuracy compared to the remaining four models. This indicates that the first layer of the $64 \times 1$ large convolutional kernel is suitable for information feature extraction in the dataset of Working Condition 3.
2, and can extract the sample feature information more adequately. In Working Condition 2, the noise immunity performance of the Propose + original vibration signal model is the same as that in Working Condition 1. The diagnostic accuracy of the model is better in the low-noise environment, while it drops sharply in the high-noise environment. This is because the addition of low noise to the signal changes the vibration signal by a small amount, and the model can still have strong diagnostic ability. However, in a strong noise environment, the vibration signal has completely lost its original characteristics. At this time, the FFT transform is used to convert the time-domain vibration signal to the frequency-domain signal, which can ignore the time information because the frequency-domain signal can retain the original vibration information. Therefore, the FFT-transformed model can still maintain a better diagnostic classification capability in the face of strong noise. This shows the necessity of signal FFT feature extraction.

In Working Condition 3, the overall trend of various algorithms is almost the same as that of Working Condition 1. Although the ANN + FFT algorithm model retains high diagnostic accuracy even in the high-noise environment, the average diagnostic accuracy in the SNR range of ~8 dB to 10 dB is still lower than that of the Propose + FFT model. In addition, the standard deviation error range of the ANN + FFT model is significantly higher than that of Propose + FFT as shown in Table 6, which also reflects the instability of the ANN + FFT model.

After the previous analysis, it can be known that the average diagnostic accuracy and error range embodied by different algorithmic models in the three operating conditions are very different, mainly due to the differences in the bearing datasets in the three operating conditions and the influence of the randomly grabbed network input sample data in the experiment. After comparison with the other five models, it can be concluded that both the average accuracy and error range of the FFT signal processing combined with the network model proposed in this paper (Propose + FFT) are better than those of the other five models under the three working conditions.

4.4. Visualization of Training Set Classification Process

In this section, the t-distributed stochastic neighbor embedding (T-SNE) technique is used to explore the classification process of small sample training data in the network model. From Table 5, it can be seen that the numbers 0 to 3 represent the types of bearing faults, respectively.

Taking the training set of Working Condition 1 as an example, one-tenth of the data volume of the training set in Table 5 is the network input data, and the T-SNE dimensionality reduction results of some network layers are shown in Figure 16.
Figure 16. Visualization of model training set classification process: (a) Initial training set; (b) FFT processing; (c) multi-scale wide convolution; (d) fourth convolution; (e) second GRU; (f) SoftMax; (g) WDCNN + Ori. (first-wide convolution); (h) WDCNN + Ori. (SoftMax); (i) WDCNN+FFT (SoftMax).

It can be seen from Figure 16 that (1) the four health states of the bearing have been well distinguished after the FFT signal feature extraction of the training data samples, (2) the further feature extraction of the FFT data by the multi-scale first layer large convolutional kernel of the convolutional neural network and the subsequent multi-layer small convolutional kernel both further enhance the linear separability of the model for the fault features, and (3) the recurrent neural network GRU further processes the current sequence information, and the proposed network model structure improves the generalization capability of the network and the bearing fault diagnosis capability in complex situations compared with the WDCNN network.

In conclusion, the combination of FFT signal processing and the improved network model proposed in this paper has a strong ability to extract indistinguishable feature information, simplify the bearing fault diagnosis problem, and improve the bearing fault diagnosis accuracy.

4.5. Visualization of Test Set Classification Process

In this section, taking Working Condition 1 as an example, the features of the test set in the original data, the data after FFT signal feature extraction, and the data in the last implied layer are reduced to two dimensions and visualized respectively by the T-SNE dimensionality reduction technique, as shown in Figure 17.
It can be seen that the original vibration signal is very disorganized and it is difficult to distinguish it correctly. After FFT, the features have been clearly distinguished after using the proposed network model. The results show that the proposed network model not only has a strong classification ability in the training set, but also still maintains a strong fault diagnosis discriminative ability in the actual test.

5. Conclusions

In this paper, the vibration dataset in the field of bearing fault diagnosis research is expanded by conducting a health state classification study on the XJTU-SY bearing dataset. After that, a deep learning network model is built to classify the divided three-condition XJTU-SY dataset for fault diagnosis research. The three parameters of RMS, kurtosis, and STFT SUM are introduced to construct the bearing health factor degradation index, and quantify each health factor in terms of trend, monotonicity, and robustness. After comparison, the RMS health factor is selected, and the first-order differential mutation point under the threshold condition is used as the failure degradation starting point of the bearing, which divides the whole life cycle of the bearing into two phases: the healthy phase and the fault phase. For the divided dataset of the XJTU-SY bearing under three working conditions, based on the WDCNN algorithm, this paper establishes a deep learning network model by introducing FFT signal feature extraction, a multi-scale first-layer wide convolution kernel, and a GRU recurrent neural network.

By comparison with the other five network models (WDCNN + original vibration data, WDCNN + FFT, Propose + original vibration data, SVM + FFT, and ANN + FFT), the results show that the network model proposed in this paper has excellent fault diagnosis capability and model noise immunity, which provides important technical support for bearing fault diagnosis in actual industrial processes.

Although the proposed method has made some achievements, there are still two limitations that need to be improved in future work. Firstly, the coverage of data in this paper is not perfect and the fault diagnosis in this paper is only for historical data due to the limited practical conditions. In the future, we can build our experimental platform to collect vibration signals. Fault diagnosis can also be extended to the scope of online detection and is not limited to the application of datasets. Secondly, in this paper, when dividing the health phase and fault phase of the bearing, the division point is taken as the collected minute unit. In future work, the state within minutes will be divided step by step, and some deeper theoretical models can be applied for health state division, such as the Hidden Markov Model and the Markovian distance.

**Author Contributions:** Conceptualization, L.S. and S.S.; methodology, L.S. and W.W.; validation, L.S. and S.S.; writing—original draft preparation, L.S. and S.G.; writing—review and editing, L.S.;
visualization, supervision, project administration, funding acquisition, S.S., C.C. and W.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This paper is supported by the National Natural Science Foundation of China (grant no. 51475129, 51675148, 51405117). This research was funded by the application and demonstration of “intelligent generation” technology for small and medium-sized enterprises—R & D and demonstration application of the chair industry internet innovation service platform based on artificial intelligence (grant no. 2020C01061).

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

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The author would like to thank Guangjie Yuan for advice in writing—review and editing.

**Conflicts of Interest:** The author of this article declares that there is no conflict of interest related to this manuscript.

**References**


31. Li, Z.; Li, Y.; Sun, Q.; Qi, B. Bearing Fault Diagnosis Method Based on Convolutional Neural Network and Knowledge Graph. Entropy 2022, 24, 1589.


34. Wei, X. Deep Learning Based Health State Assessment And Remaining Life Prediction of Rolling Bearings. Master’s Thesis, Southwest Jiaotong University, Chengdu, China, 2021.


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.