

Research Article

Intelligent Fault Diagnosis of Bearing Based on Convolutional Neural Network and Bidirectional Long Short-Term Memory

Dazhang You,¹ Linbo Chen ,¹ Fei Liu ,² YePeng Zhang,¹ Wei Shang ,¹ Yameng Hu,³ and Wei Liu⁴

¹School of Mechanical Engineering, Hubei University of Technology, Wuhan, China

²School of Computer Science and Engineering, Wuhan Institute of Technology, Wuhan, China

³School of Industrial Design, Hubei University of Technology, Wuhan, China

⁴Aeronautics Computing Technology Research Institute, Xi'an, China

Correspondence should be addressed to Fei Liu; joinee@sina.com

Received 6 June 2021; Revised 16 August 2021; Accepted 18 October 2021; Published 11 November 2021

Academic Editor: Filippo Ubertini

Copyright © 2021 Dazhang You et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The traditional bearing fault diagnosis methods have complex operation processes and poor generalization ability, while the diagnosis accuracy of the existing intelligent diagnosis methods needs to be further improved. Therefore, a novel fault diagnosis approach named CNN-BLSTM for bearing is presented based on convolutional neural network (CNN) and bidirectional long short-term memory (BLSTM) in this paper. This method directly takes the collected one-dimensional raw vibration signal as input and adaptively extracts the feature information through CNN. Then, the BLSTM is used to fuse the extracted features to acquire the failure information sufficiently and prevent the model from overfitting. Finally, two different experimental datasets are used to verify the effectiveness of the method. The experimental results show that the proposed CNN-BLSTM model can accurately diagnose the fault category of bearings. It has the advantages of rapidity, stability, antinoise, and strong generalization.

1. Introduction

As a commonly used power transmission device in mechanical equipment, bearings are widely used in industrial production [1–3]. Its operation status will directly determine the useful performance of the equipment [4–7]. Once a failure occurs, it will cause unstable operation of the equipment and even cause major losses [4]. Therefore, it is of great significance to carry out effective condition monitoring and diagnosis of bearings [8–10].

At present, the research methods of fault diagnosis are mainly divided into model-based methods, experience-based methods, and intelligence-based methods [11–16]. Among them, the intelligence-based methods were also called the data-driven methods. This method uses a large amount of historical data to establish a fault diagnosis model and gets rid of the dependence on physical parameter models [17]. Compared with the model-based and experience-based diagnosis methods, the data-driven diagnosis

method has stronger operability [4]. With the continuous improvement of monitoring methods, the data in enterprise can be collected much faster and more widely than ever before, which has laid a good foundation for data-driven fault diagnosis methods [17, 18]. In recent years, with the rapid development of deep learning theory, remarkable achievements have been made in the fields of image processing, computer vision, speech recognition, and so on [19]. As a representative deep learning technology, convolutional neural network (CNN) has been widely used in data-driven fault diagnosis methods. Many scholars have proposed a lot of data-driven fault prediction methods [20–24]. Gao et al. [25] proposed a new method of fault diagnosis using a bidirectional long short-term memory- (BLSTM-) based Hilbert-Huang transform (HHT) and CNN. Liang et al. [26] proposed a fault diagnosis method based on Wavelet Transform (WT), Generative Adversarial Network (GAN), and CNN. Abdeljaber et al. [27] used one-dimensional CNN to automatically extract the optimal damage-sensitive

features directly from the original acceleration signals. Zhou et al. [28] proposed a bearing remaining life prediction and fault diagnosis method based on short-time Fourier transform (STFT) and CNN. Levent et al. [29] directly input the raw signal into one-dimensional CNN for fault classification and verified it on the Case Western Reserve University (CWRU) bearing dataset.

Although these methods can better extract fault features and perform fault diagnosis, the accuracy of diagnosis is not high, the operation process is complicated, and the running time is longer. In addition, what CNN learns is the spatial information in the receptive field. With the deepening of the network, the features learned by CNN become more and more abstract, which is prone to overfitting and further affects the judgment of test results. Therefore, it is the key to improving the accuracy of fault diagnosis and preventing the overfitting of the model without destroying the relationship between information. A new fault diagnosis method that can effectively extract fault feature information and prevent overfitting is proposed. The bearing faults at the same position are divided into one class, and the fault diagnosis model based on CNN and BLSTM is established. The effectiveness of the diagnosis method is verified by the benchmark dataset CWRU bearing data [30] and the bearing dataset measured under the time-varying speed condition of the University of Ottawa (uOttawa) in Canada [31].

The contributions of this paper can be summarized as follows:

- (1) A new feature extraction method, called CNN-BLSTM, is proposed to reduce the complexity of the diagnosis process and avoid overfitting of the model during the training process
- (2) Avoid the interference of high correlations between different size fault classes in the same position
- (3) This method has good robustness and achieves high accuracy even under strong noise interference

The rest of this paper is as follows. In Section 2, the basic principles and network structure of CNN and BLSTM are introduced. In Section 3, we describe the proposed CNN-BLSTM intelligent fault detection method in detail. In Section 4, the effectiveness of the method proposed in this paper is verified through two different experiments, and the results are compared with other intelligent fault diagnosis methods in the literature. Section 5 is the conclusion of this paper.

2. Theoretical Background

2.1. Convolutional Neural Network. As one of the representative algorithms of deep learning, CNN has powerful feature extraction ability and was widely applied in various fields such as image processing and time-series data [32]. In mechanical equipment monitoring, the data collected by sensors is usually a one-dimensional time series. Therefore, it is more suitable to use one-dimensional CNN for feature extraction. A typical CNN consists of an input layer, convolutional layers, activation layers, pooling layers, fully

connected layers, and an output layer. The structural difference between two-dimensional and one-dimensional CNN, the main difference, is the use of one-dimensional arrays instead of two-dimensional matrices in the kernel and feature maps [27].

The main principle of one-dimensional CNN is to use convolution kernel to perform convolution operation on one-dimensional signals input to the convolution layer and then extract features and sparse processing layer by layer through activation function and pooling layer. The formula can be defined as follows:

$$x_j^t = f \left(\sum_{i=1}^M x_i^{t-1} * k_{ij}^t + b_j^t \right), \quad (1)$$

where k is the convolution kernel; j is the number of cores; M is the number of input channels x_i^{t-1} ; b_j^t is the bias parameter; $*$ is the convolution operator; f is the activation function; Sigmoid, ReLU, and other activation functions are the commonly used functions. Because the Sigmoid activation function has some disadvantages—for example, the gradient disappears, and the output is not a zero mean value—the ReLU activation function can overcome these disadvantages and converge faster, but there is neuron necrosis. Therefore, this paper adopts the ELU activation function, whose formula can be expressed as follows:

$$y_j^t = f(\alpha, x_j^t) = \max \begin{cases} \alpha(e^{x_j^t} - 1), & x_j^t < 0; \alpha \geq 0, \\ x_j^t, & x_j^t \geq 0. \end{cases} \quad (2)$$

where y_j^t is the activation value of the t -th layer.

The pooling layer is usually after the convolution layer; the main purpose of pooling is to downsample, make use of the local characteristics of the data itself, optimize the network structure, and improve the robustness of the features while reducing the data dimension. Commonly used downsampling methods include stochastic-pooling, mean-pooling, and max-pooling, among which max-pooling is the most widely used, which is expressed as follows:

$$z_j^{t(l)} = \max_{(l-1)r+1 \leq w \leq rl} \{y_j^{t(w)}\}, \quad (3)$$

where w is the width of the pooling area; $y_j^{t(w)}$ represents the t -th activation value of the j -th neuron in the l -th layer; $z_j^{t(l)}$ represents the t -th value after the pooling operation in the j -th neuron in the l -th layer.

2.2. Bidirectional Long- and Short-Term Memory Neural Network. LSTM, as a variant of RNN, avoids the problems of exploding or vanishing gradients in the backpropagation of long sequence data in traditional RNN and realizes the long-term memory of useful information in state monitoring data. The LSTM unit consists of 3 components: forget gate, update gate, and output gate. The unit structure is shown in Figure 1.

The forget gate is responsible for discarding unimportant information from the unit state received by the previous unit. The update gate is responsible for the addition of information

to the cell state. The output gate is responsible for selecting useful information from the current cell state and showing it gives as an output [33]. The formula is as follows:

$$\begin{aligned}
 f_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \\
 i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \\
 \tilde{C}_t &= \tanh(W_C \cdot [h_{t-1}, x_t] + b_C), \\
 C_t &= f_t \cdot C_{t-1} + i_t \tilde{C}_t, \\
 o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \\
 h_t &= o_t * \tanh(C_t),
 \end{aligned} \tag{4}$$

where f_t , i_t , and o_t are the forget gate, update gate, and output gate, respectively; b_f , b_i , and b_o are the bias vectors of the forget gate, update gate, and output gate, respectively; C_{t-1} is the previous neuron state; h_{t-1} is the last hidden state; C_t is the current neuron state; h_t is the current hidden state; σ is the sigmoid activation function; \tanh is the hyperbolic tangent function.

Since the LSTM network can only process data in one direction when processing time series such as condition monitoring data, it cannot make full use of the data [34]. The BLSTM layer consists of LSTM layers superimposed on each other in opposite directions, which enables the model to make full use of the main feature information and improve the performance. The structure of the BLSTM unit is shown in Figure 2.

Generally, there are 1-2 fully connected layers behind the BLSTM layer. All neurons in the fully connected layer are usually connected with the feature vector nodes output from the previous layer. The output layer after the fully connected layer often uses Softmax as a classifier to effectively classify multiple targets. Assuming a task is a category problem, the output of Softmax can be expressed as follows:

$$\begin{aligned}
 O_j &= \begin{pmatrix} P(y = 1|x; \theta) \\ P(y = 2|x; \theta) \\ \dots \\ P(y = k|x; \theta) \end{pmatrix} \\
 &= \frac{1}{\sum_{j=1}^k e^{\theta^{(j)}x}} \begin{pmatrix} e^{\theta^{(1)}x} \\ e^{\theta^{(2)}x} \\ \dots \\ e^{\theta^{(k)}x} \end{pmatrix},
 \end{aligned} \tag{5}$$

where $\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(k)}$ are the parameters of the model and O_j is the final output result.

3. Proposed Method

In this paper, a fault diagnosis method of CNN-BLSTM is proposed based on the research of fault diagnosis. It makes the most of the advantages of CNN in feature extraction and

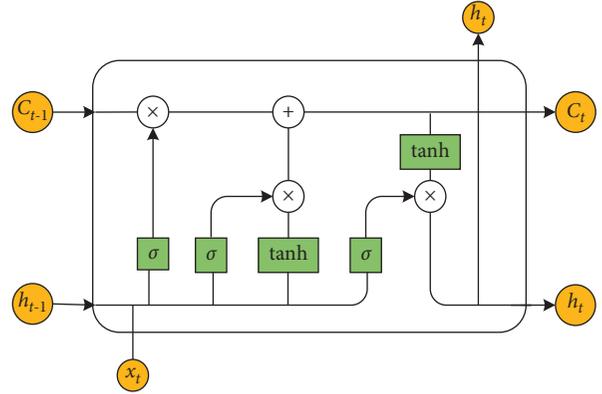


FIGURE 1: LSTM unit structure.

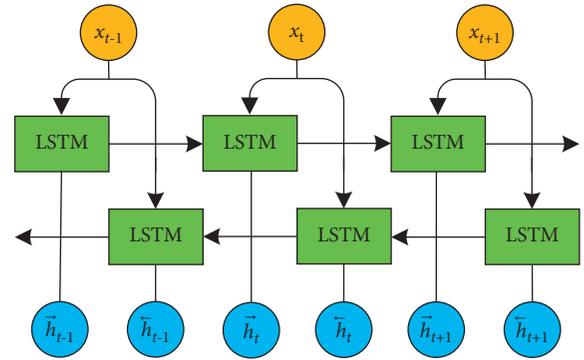


FIGURE 2: BLSTM unit structure.

BLSTM in dealing with gradient disappearance and explosion. As shown in Figure 3, the proposed CNN-BLSTM model consists of three alternating convolutional and pooling layers, a BLSTM layer, and two fully connected layers. Firstly, the main features of input signals are extracted through three alternating convolution layers and pooling layers. Then, BLSTM is used to further extract features as the input of the full connection layer. Finally, the Softmax classifier is used for fault identification and classification.

4. Experimental Validation

To verify the effectiveness of the proposed method, two experiments on different bearing datasets were conducted. The data of Case 1 was the public bearing fault dataset obtained from CWRU (<https://csegroups.case.edu/bearingdatacenter/home>). The data of Case 2 came from the bearing dataset obtained by the uOttawa under the condition of time-varying speed (<https://data.mendeley.com/datasets/v43hmbwxpm/2>).

4.1. CWRU Bearing Dataset

4.1.1. Experiment Setup and Data Description. As the benchmark dataset of fault detection, CWRU rolling bearing data has been widely used in the field of fault monitoring by many scholars [30]. The CWRU experimental platform, as shown in Figure 4, consisted of a 2-horsepower (about 1.5 KW) motor, torque sensor/decoder, power meter, and so

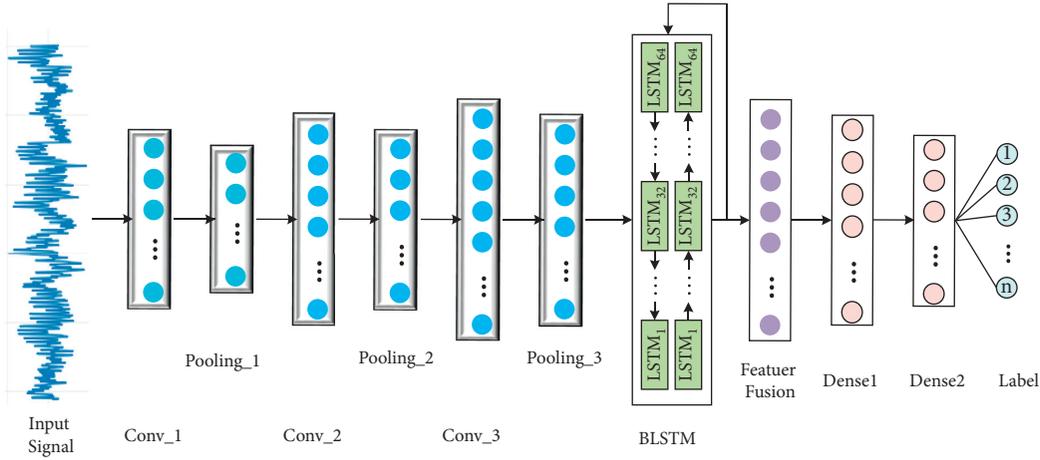


FIGURE 3: CNN-BLSTM model structure frame diagram.

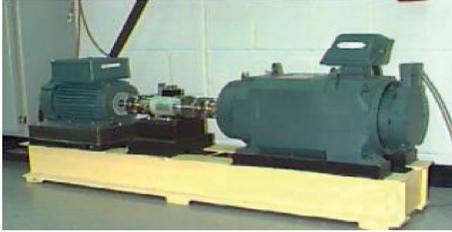


FIGURE 4: CWRU experimental platform.

on. The bearing (6205-2RS JEM SKF) to be tested was installed on the drive end of the motor to support the motor shaft, and an acceleration sensor was placed above the bearing seat of the motor drive end to collect the vibration acceleration signal of the faulty bearing. The load was about 1HP (1772 r/min). The sampling frequency was 12 kHz, and the time was 10 s. Three bearing failure types were considered: rolling, inner race, and outer race faults. Each fault type was induced in the rolling bearings by electrodischarge with different fault diameters (0.007, 0.014, and 0.021 inches, of which 1 inch = 25.4 mm) [32].

The selected fault status was specifically constituted as shown in Table 1. There are four fault types of the bearing: normal, ball fault, inner race fault, and outer race fault, reaching 10 fault conditions in total, of which the outer race fault point was 6 o'clock direction. In view of the fact that faults in the same location were often classified into one class in life, the fault type in Table 1 was 4. This paper used One-Hot-coding to label the fault type; for example, the bearing outer race fault was "label = 3" and the code was [0, 0, 0, 1]. In this experiment, in order to ensure the integrity of the fault information of the sampled data, 1024 data points of the experimental platform about two cycles were selected as a sample. This paper performed downsampling for each fault data in Table 1, and the sampling data points were 120,832. Therefore, each fault signal data could be divided into 118 samples, with a total of 1180 samples. To ensure sufficient training volume, 90% of them were randomly selected as training set samples and 10% as test set samples. In addition, 10% of the training set was randomly divided into validation sets to adjust model parameters.

4.1.2. CNN-BLSTM Model. The parameters selected by the CNN-BLSTM model through debugging are shown in Table 2. The first convolutional layer used a wider core, which could effectively suppress high-frequency noise and improve the performance of the network [35]. Therefore, the size of the first layer of convolution kernel in this paper was selected as 8, and the step size was 2 to speed up the iterative convergence efficiency. In order to prevent the overfitting problem, dropout was introduced in the second and third layers of the convolution layer, a value of 0.2. Softmax was used as a classifier and the optimizer was selected as the Adam. The loss function was classification cross-entropy and batch size is 64. In order to avoid the contingency and particularity of the experiment, this paper used the same parameters to carry out 10 independent experiments. To detect the fault state within a relatively short time, the number of each experiment iteration was 50. All experiments were run on the same PC platform, using TensorFlow to build the CNN-BLSTM model; the processor was Intel(R) Core(TM) i5-6500 CPU @ 3.22 GHz, 3.32 GHz.

4.1.3. Experimental Results and Analysis. In the intelligent fault diagnosis method, most methods either used too little data or divided faults of different sizes in the same location into different categories. Since the equipment may have a sudden destructive failure in operation, it may not be able to fully reflect the comprehensive data fault state of the industrial scale when the amount of fault data is small. In order to explore the influence of too many classifications on the diagnosis results, faults of different sizes at the same location were regarded as different classes. The data categories in Table 1 were divided into 4 and 10 categories for experimental comparison. The comparison model was the method proposed in this paper, the CNN-LSTM model framework with BLSTM layer replaced by LSTM layer and the CNN model framework without BLSTM layer. Figure 5 shows the test results of 10 repeated experiments. Tables 3 and 4 show the average accuracy and standard deviation of the test results.

Combining Figure 5 and Tables 3 and 4, we could see the following: (1) Whether it was classified into 4 or 10

TABLE 1: CWRU experimental sample.

State	Defect size (inch)	Data length	No. of samples	Load zone	Class (label)
Healthy state	—	1024	118	—	0
	0.007	1024	118	—	1
Ball fault	0.014	1024	118	—	1
	0.021	1024	118	—	1
	0.007	1024	118	—	2
Inner race fault	0.014	1024	118	—	2
	0.021	1024	118	—	2
	0.007	1024	118	6:00	3
Outer race fault	0.014	1024	118	6:00	3
	0.021	1024	118	6:00	3

TABLE 2: CNN-BLSTM model parameters applied.

Name	Filters	Kernel size/stride	Padding	Units	Activation function
Conv_1	16	8/2	Same		Elu
Pooling_1		3/-	Valid		
Conv_2	32	3/1	Same		Elu
Pooling_2		3/-	Valid		
Conv_3	64	3/1	Same		Elu
Pooling_3		3/-	Valid		
BLSTM_1				64	
Dense				32	Elu
Dense				4	Softmax

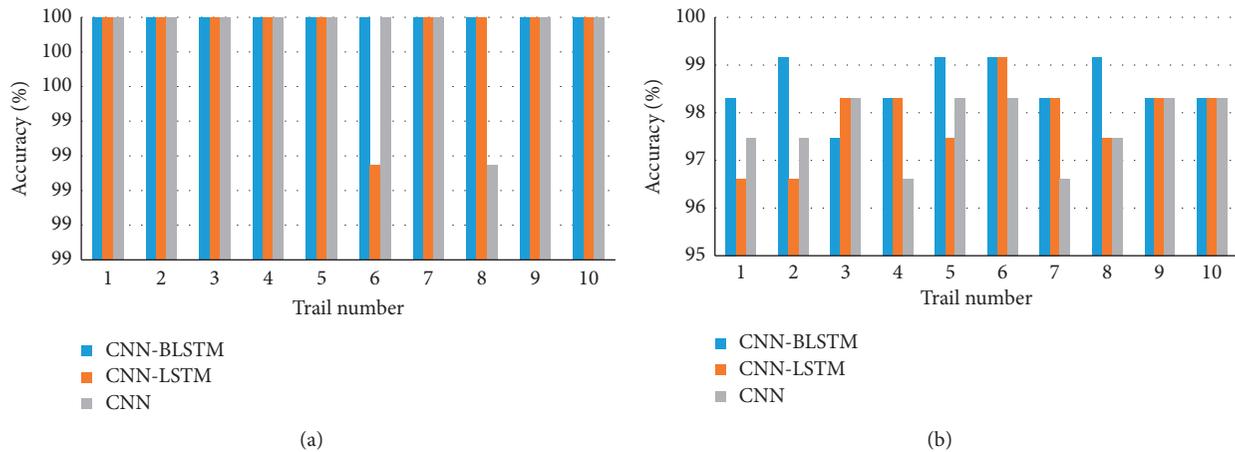


FIGURE 5: Classification test experiment results. (a) The experimental data is divided into 4 categories. (b) The experimental data is divided into 10 categories.

TABLE 3: Dividing into 4 types of test diagnosis results.

Model	Average recognition rate (%)	Standard deviation	Calculation time (seconds)
CNN	99.92	0.0025	16.119
CNN-LSTM	99.92	0.0025	22.711
CNN-BLSTM	100	0	27.183

categories, the CNN model had the lowest accuracy and the highest standard deviation, which proves that its stability was poor. (2) When the dataset was divided from 10 categories to 4 categories, both the test diagnosis results and stability were improved. (3) In the test results of 4 and 10 classes, the CNN-BLSTM model proposed in this paper had the highest average test accuracy and the lowest standard

deviation. It shows that adding the BLSTM layer to the CNN model structure can make full use of the extracted characteristic signals and improve the test accuracy, but the calculation time will increase accordingly.

As can be seen from Figure 5(b), the mode of the test result was 98.31%. Figure 6 shows the confusion matrix of the three models with an accuracy of 98.31%. The confusion

TABLE 4: Dividing into 10 types of test diagnosis results.

Model	Average recognition rate (%)	Standard deviation	Calculation time (seconds)
CNN	97.72	0.0066	16.693
CNN-LSTM	97.88	0.0078	23.016
CNN-BLSTM	98.56	0.0054	27.268

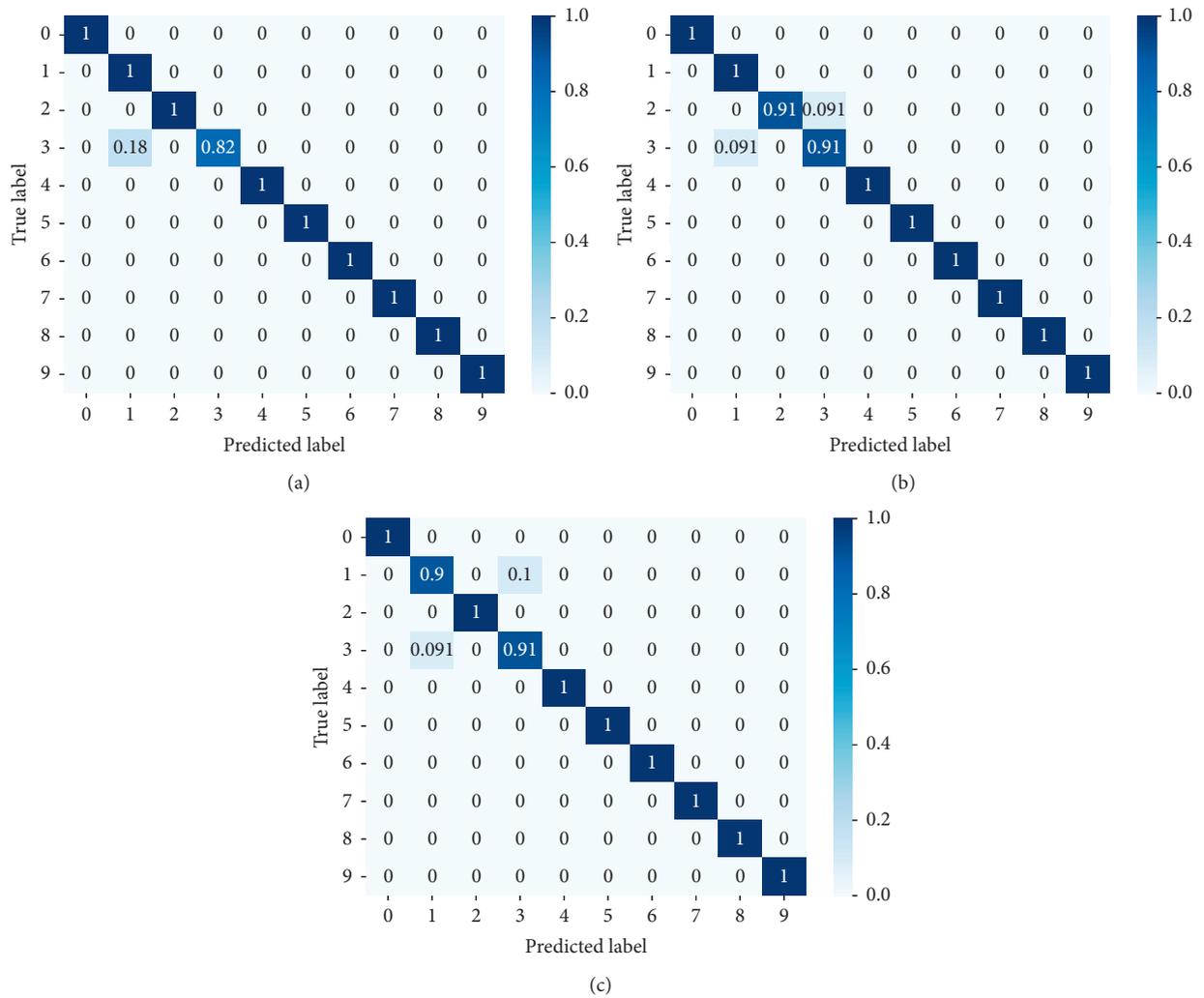


FIGURE 6: The mode of the three models in 10 classifications was 98.31%. (a) CNN-BLSTM model. (b) CNN-LSTM model. (c) CNN model.

matrix could clearly reflect the matching degree between the true label and the predicted label. As shown in Figure 6, none of the three models correctly identified the ball fault, which indicated that the high correlation between the different state categories at the same position would interfere with the fault recognition, thus affecting the accuracy of diagnosis.

In the traditional intelligent diagnosis methods of bearing, most of the methods had preprocessed the original signal, such as high-order cumulants (HOCs), WT, continuous wavelet transform (CWT), and EMD, or performed dimensionality reduction techniques to extract main features and then classify. Although selected data or manually extracted features could indeed improve the accuracy of the test results, this method required prior knowledge and destroyed the correlation between the

data. Therefore, these methods might not be able to achieve real-time condition monitoring tasks of mechanical equipment. Besides, the datasets used in most of these studies were too small to fully reflect the comprehensive data fault state of the industrial scale. Compared with the other articles, the CNN-BLSTM method proposed in this paper can effectively extract the fault features and identify the fault state in a shorter time.

Table 5 uses the same baseline dataset. We used a relatively large feature learning data framework and achieved higher accuracy without any data preprocessing or feature manipulation. Although the WT-GAN-CNN method proposed in literature [26] also had a high test accuracy, it used only 6 kinds of data types and required WT to preprocess the data. The CWT-CNN-RF method proposed in reference [40]

TABLE 5: Comparison between other methods in the literature and our proposed model.

Classifier	Data preprocessing	CWRU test accuracy	Calculation time (seconds)
KNN [36]	HOCs and WT	91.23%	
SVM [37]	WP	98.7% (4 classes)	12.914
RF [38]		98.04% (4 classes)	
WT-CNN [26]	WT	99.25% (4 classes)	
WT-GAN-CNN [26]	WT	100% (4 classes)	
Compact 1D CNN [29]		93.2% (6 classes)	
CNN-LSTM [39]		99.77% (6 classes)	419
CWT-CNN-RF [40]	CWT	99.73% (10 classes)	114.57
MSCNN-LSTM [32]		98.46% (10 classes)	
CNN-BLSTM		100% (10 classes)	27.268

has good results in 10 kinds of data, but its processing process is more complicated. Compared with the method that directly used the original signal as the input in literature [29, 32, 38, 39], the CNN-BLSTM method proposed in this paper can effectively extract the fault features and identify the fault state in a shorter time.

4.1.4. Performance under Noise Environment. In the process of collecting rotating mechanical signal data, due to the complex working environment of the equipment, some outside interference noise would inevitably be collected. In order to better simulate the unknown real noise and verify the fault detection accuracy of the proposed CNN-BLSTM model for signals containing noise, the additive white Gaussian noise (AWGN) with different standard variances was added to the original signal data to form different signal-to-noise ratios (SNR) [26]. The SNR formula was

$$\text{SNR}_{\text{db}} = 10 \log_{10} \left(\frac{P_{\text{signal}}}{P_{\text{noise}}} \right), \quad (6)$$

where P_{signal} is the power of signals and P_{noise} is noise in signals.

The experiment was implemented as follows. Gaussian white noise was added to the original vibration signal of the selected data in Table 1 to form a noisy signal. According to the above method, the SNR ranging from -4 to 6 dB noise signal was directly input into the model to test the CNN-BLSTM model proposed in this paper.

The test results were compared with several other existing methods in the literature [41], and the results as shown in Figure 7. It was clear that the test accuracy rate of fault detection increased as the SNR became decreased. When the noise condition was $\text{SNR} = -4$ dB, the performance of the other three methods as well as this paper was poor and the test accuracy was very low, while our proposed CNN-BLSTM model could still achieve an average accuracy of 97.35%. When $\text{SNR} = 0$ dB, the CNN-BLSTM model could easily achieve 100% test accuracy. In different SNR, the CNN-LSTM model also had high test accuracy, but its stability was not as good as CNN-BLSTM. In summary, the CNN-BLSTM model performed well against noise in different environments without any preprocessing.

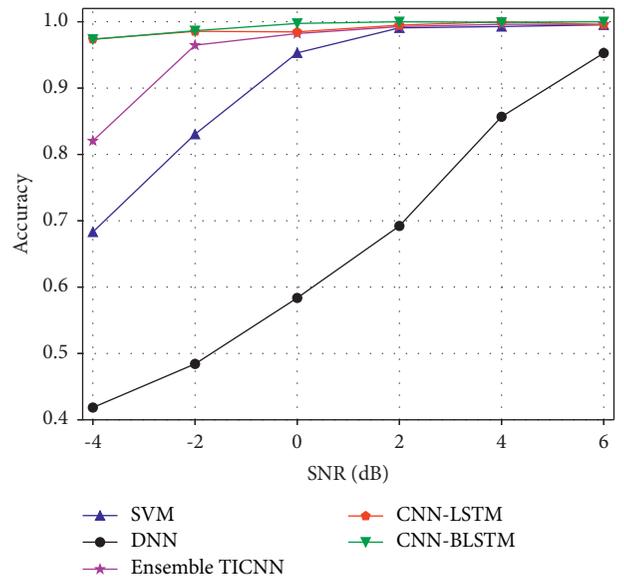


FIGURE 7: Comparison of accuracy with different SNR values.

4.2. uOttawa Bearing Dataset

4.2.1. Experiment Setup and Data Description. To further validate the effectiveness and generalization of the proposed method for bearing fault diagnosis under unknown time-varying speed conditions [42], this paper selected uOttawa bearing data for further experiments. Each dataset had two experimental settings: bearing condition and varying speed condition. There were 5 types of bearing conditions, including health, inner race fault, outer race fault, ball fault, and comprehensive fault, of which comprehensive fault contained ball fault, inner race fault, and outer race fault. There were 4 types of varying speed conditions: (i) increasing speed, (ii) decreasing speed, (iii) increasing then decreasing speed, and (iv) decreasing then increasing speed. The data was the vibration data measured by the acceleration sensor, the sampling frequency was 200,000 Hz, the sampling duration was 10 seconds, and the sampling points in a period were 1024 [31]. This paper selected 2000 data points as a sample, so each fault signal data could be divided into 1000 samples, totaling 20,000 samples.

The specific composition of selected fault states is shown in Table 6, with 5 bearing statuses and 4 varying speed

TABLE 6: uOttawa experimental sample.

Bearing condition	Variable speed condition	Data length	No. of samples	Class (label)
Healthy state	(i), (ii), (iii), and (iv)	2000	4000	0
Ball fault	(i), (ii), (iii), and (iv)	2000	4000	1
Inner race fault	(i), (ii), (iii), and (iv)	2000	4000	2
Outer race fault	(i), (ii), (iii), and (iv)	2000	4000	3
Comprehensive fault	(i), (ii), (iii), and (iv)	2000	4000	4

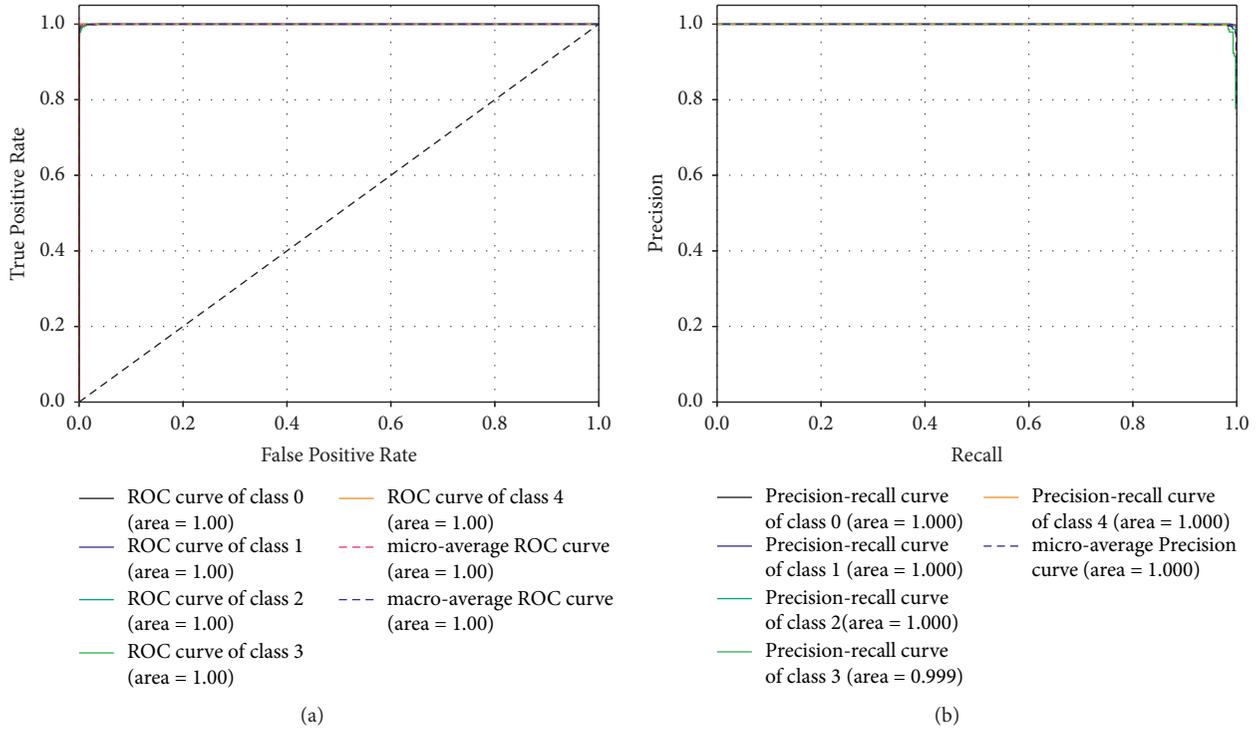


FIGURE 8: Training situation. (a) ROC curve. (b) P-R curve.

conditions, totaling 20 bearing statuses. In daily applications, faults at the same location were often classified as one type, so the fault types in the experimental samples in Table 6 could be divided into 5 types.

4.2.2. Experimental Results and Analysis. As described in Case 1, 90% of the samples were randomly selected as the training set and 10% as the test set. In addition, 10% of the training set was randomly divided into validation sets. Figure 8 shows the iterative situation of CNN-BLSTM training when the accuracy is 99.4%. Figure 8(a) is the ROC curve of the classification evaluation index; Figure 8(b) is the P-R curve, which is the Precision and Recall curve. It uses Recall as the abscissa and Precision as the ordinate.

In this paper, the CNN-BLSTM model, CNN-LSTM model, CNN model, DNN model, RF, and SVM were analyzed experimentally. Among them, the neurons of each layer of the DNN model were 128/256/512/256/32/5, and other parameters were set the same as the model proposed in this paper. The kernel function used by SVM was Radial Basis Function (RBF). RF selects the parameter n -estimators as 275 through GridSearchCV, and the others are default

parameters. The results of the 10 experiments are shown in Figure 9. Average accuracy and standard deviation are shown in Table 7. It can be seen from Figure 10 that, during the training process of the CNN-BLSTM and CNN-LSTM models, as the number of iteration steps increases, the diagnostic accuracy becomes higher and higher, and the model shows a trend of convergence. Although the CNN and DNN models take a short time to process large-scale data, they are prone to overfitting during the training process, which affects the diagnosis results.

Combining Figures 9 and 10 and Table 7, we could see that, among all the methods, the CNN-BLSTM method used in this paper had the highest diagnostic accuracy and the best robustness, with an average accuracy of 99.19%. The CNN-LSTM model had the second-highest accuracy rate, partly because LSTM could prevent overfitting. Besides, the accuracy of the CNN test used in this paper also had good results, but its accuracy was not as good as CNN-BLSTM and CNN-LSTM. It was worth noting that SVM had the lowest accuracy rate among all methods. This was because a small number of support vectors determined the final result. When processing large-scale data samples, SVM was vulnerable to interference from the strong correlation between

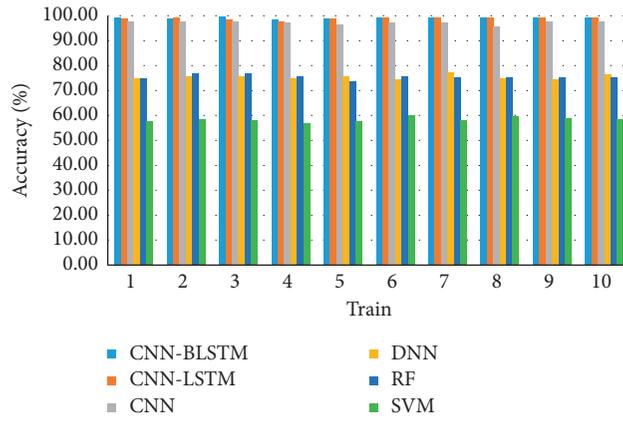
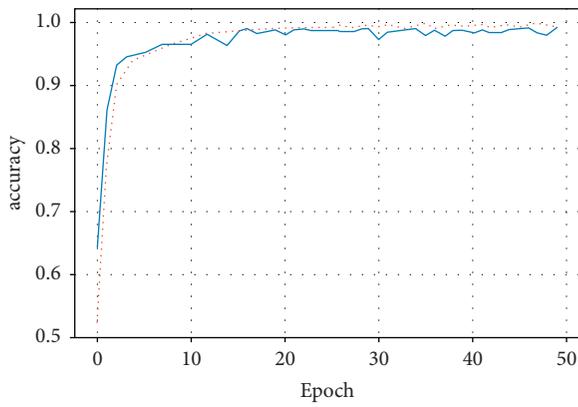


FIGURE 9: Test results.

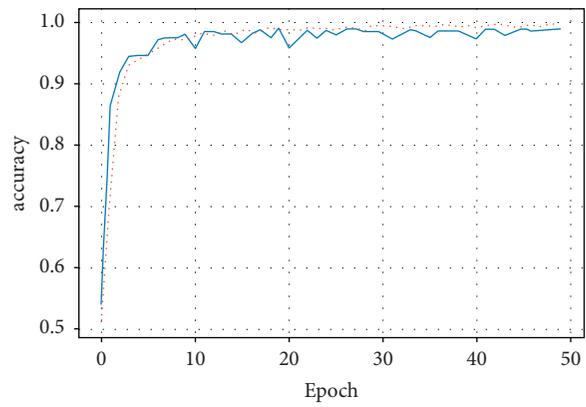
TABLE 7: Test results by different methods.

Model	Average recognition rate (%)	Standard deviation	Calculation time (seconds)
SVM	58.5	0.0087	804.729
RF	75.44	0.008	68.25
DNN	75.38	0.0068	42.94
CNN	97.29	0.0068	506.475
CNN-LSTM	99.05	0.0056	810.311
CNN-BLSTM	99.2	0.0029	817.506



..... train
 — validation

(a)



..... train
 — validation

(b)

FIGURE 10: Continued.

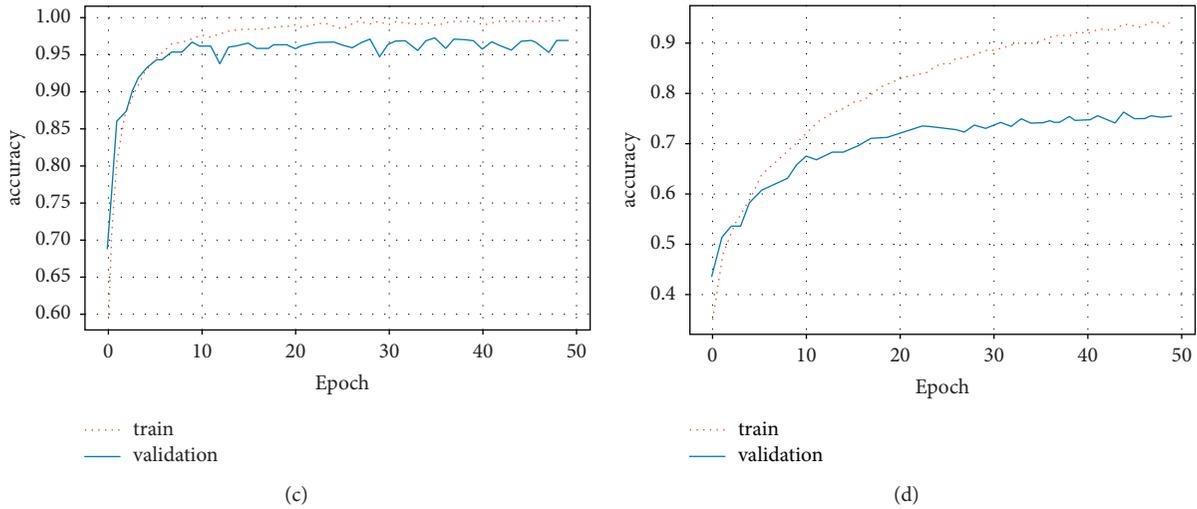


FIGURE 10: Model training iterative process. (a) CNN-BLSTM highest accuracy = 99.55%. (b) CNN-LSTM highest accuracy = 99.45%. (c) CNN highest accuracy = 97.9%. (d) DNN minimum accuracy = 77.15%.

signal data. As can be seen from the experimental results, RF is a simple and fast method, but its classification results are not as good as the proposed method.

5. Conclusion

In order to improve the diagnosis accuracy and stability of the fault diagnosis method, a feature extraction method to prevent overfitting, namely, CNN-BLSTM, is proposed in this paper. CNN-BLSTM adaptively extracts the fault features in the raw signal data through the end-to-end method and then uses the BLSTM layer to fuse the feature information to enhance stability and avoid falling into local optimum. Taking two sets of experimental bearing vibration datasets as samples, the validity and feasibility of the method proposed in this paper are verified. The results show that this method can accurately extract highly discriminative features directly from the input raw signal data. High diagnostic accuracy can still be achieved even when the working condition is changed or ambient noise is strong. The highest fault diagnostic accuracy on the CWRU and uOttawa datasets can reach 100% and 99.55%, respectively, higher than that of other existing fault diagnosis approaches. Therefore, the proposed CNN-BLSTM is an effective feature extraction method for fault recognition, and BLSTM is very suitable for the classification problem of preventing model overfitting. The proposed fault diagnosis method has higher diagnostic accuracy and meets the application requirements of robustness and generalization.

Data Availability

The data of Case 1 was the public bearing fault dataset obtained from CWRU (<https://csegroups.case.edu/bearingdatacenter/home>). The data of Case 2 came from the bearing dataset obtained by the uOttawa under the condition of time-varying speed (<https://doi.org/10.17632/v43hmbwxpm.2>).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors would like to acknowledge the <https://doi.org/10.13039/100008136> Case Western Reserve University and the <https://doi.org/10.13039/100008572> University of Ottawa for the public data of rolling bearing failure experiments on the Internet. This research was supported by the Natural Science Foundation of China (Grant no. 51875180).

References

- [1] Q. Lu, R. Yang, M. Zhong, and Y. Wang, "An improved fault diagnosis method of rotating machinery using sensitive features and RLS-BP neural network," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 4, pp. 1585–1593, 2020.
- [2] Z. He, H. Shao, J. Cheng, X. Zhao, and Y. Yang, "Support tensor machine with dynamic penalty factors and its application to the fault diagnosis of rotating machinery with unbalanced data," *Mechanical Systems and Signal Processing*, vol. 141, Article ID 106441, 2020.
- [3] L. Song, H. Wang, and P. Chen, "Vibration-based intelligent fault diagnosis for roller bearings in low-speed rotating machinery," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 8, pp. 1887–1899, 2018.
- [4] H. Zhao, H. Liu, Y. Jin, X. Dang, and W. Deng, "Feature extraction for data-driven remaining useful life prediction of rolling bearings," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–10, 2021.
- [5] T. Jin, H. Ding, H. Xia, and J. Bao, "Reliability index and Asian barrier option pricing formulas of the uncertain fractional first-hitting time model with Caputo type," *Chaos, Solitons & Fractals*, vol. 142, Article ID 110409, 2020.
- [6] X. Cai, H. Zhao, S. Shang et al., "An improved quantum-inspired cooperative co-evolution algorithm with multi-strategy and its application," *Expert Systems with Applications*, vol. 171, Article ID 114629, 2021.

- [7] W. Deng, S. Shang, X. Cai, H. Zhao, Y. Song, and J. Xu, "An improved differential evolution algorithm and its application in optimization problem," *Soft Computing*, vol. 25, no. 7, pp. 5277–5298, 2021.
- [8] Z. Chen, K. Gryllias, and W. Li, "Mechanical fault diagnosis using convolutional neural networks and extreme learning machine," *Mechanical Systems and Signal Processing*, vol. 133, Article ID 106272, 2019.
- [9] X. Zhao, M. Jia, and M. Lin, "Deep laplacian auto-encoder and its application into imbalanced fault diagnosis of rotating machinery," *Measurement*, vol. 152, Article ID 107320, 2020.
- [10] Y. Wang, J. Zhou, L. Zheng, and C. Gogu, "An end-to-end fault diagnostics method based on convolutional neural network for rotating machinery with multiple case studies," *Journal of Intelligent Manufacturing*, pp. 1–22, 2020.
- [11] M. Kordestani, M. F. Samadi, M. Saif, and K. Khorasani, "A new fault diagnosis of multifunctional spoiler system using integrated artificial neural network and discrete wavelet transform methods," *IEEE Sensors Journal*, vol. 18, no. 12, pp. 4990–5001, 2018.
- [12] F. Liu, J. Gao, and H. Liu, "The feature extraction and diagnosis of rolling bearing based on CEEMD and LDWPSO-PNN," *IEEE Access*, vol. 8, pp. 19810–19819, 2020.
- [13] Y. Hu, X. Xue, Z. Jin, and K. Peng, "Time-varying fault diagnosis for asynchronous multisensor systems based on augmented IMM and strong tracking filtering," *Journal of Control Science and Engineering*, vol. 2018, Article ID 5205698, 8 pages, 2018.
- [14] Z.-H. Zhang and G.-H. Yang, "Distributed fault detection and isolation for multiagent systems: an interval observer approach," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 6, pp. 2220–2230, 2020.
- [15] S. Luo, J. Cheng, M. Zeng, and Y. Yang, "An intelligent fault diagnosis model for rotating machinery based on multi-scale higher order singular spectrum analysis and GA-VPMCD," *Measurement*, vol. 87, pp. 38–50, 2016.
- [16] H. Wang, Z. Liu, D. Peng, and Y. Qin, "Understanding and learning discriminant features based on multiattention 1DCNN for wheelset bearing fault diagnosis," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 9, pp. 5735–5745, 2020.
- [17] L. Wen, X. Li, L. Gao, and Y. Zhang, "A new convolutional neural network-based data-driven fault diagnosis method," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 7, pp. 5990–5998, 2018.
- [18] J. Guo, X. Liu, S. Li, and Z. Wang, "Bearing intelligent fault diagnosis based on wavelet transform and convolutional neural network," *Shock and Vibration*, vol. 2020, Article ID 6380486, 14 pages, 2020.
- [19] X. Zhang, P. Han, L. Xu, F. Zhang, Y. Wang, and L. Gao, "Research on bearing fault diagnosis of wind turbine gearbox based on 1DCNN-PSO-SVM," *IEEE Access*, vol. 8, pp. 192248–192258, 2020.
- [20] Q. Ni, J. C. Ji, K. Feng, and B. Halkon, "A fault information-guided variational mode decomposition (FIVMD) method for rolling element bearings diagnosis," *Mechanical Systems and Signal Processing*, vol. 164, Article ID 108216, 2022.
- [21] H. Dongying, N. Zhao, and S. Peiming, "Gear fault feature extraction and diagnosis method under different load excitation based on EMD, PSO-SVM and fractal box dimension," *Journal of Mechanical Science and Technology*, vol. 33, no. 2, pp. 487–494, 2019.
- [22] Y. Cheng, M. Lin, J. Wu, H. Zhu, and X. Shao, "Intelligent fault diagnosis of rotating machinery based on continuous wavelet transform-local binary convolutional neural network," *Knowledge-Based Systems*, vol. 216, Article ID 106796, 2021.
- [23] J. Liu, K. Zhou, C. Yang, and G. Lu, "Imbalanced fault diagnosis of rotating machinery using autoencoder-based supergraph feature learning," *Frontiers of Mechanical Engineering*, pp. 1–11, 2021.
- [24] S. Liu, H. Jiang, Z. Wu, and X. Li, "Data synthesis using deep feature enhanced generative adversarial networks for rolling bearing imbalanced fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 163, Article ID 108139, 2022.
- [25] T. Gao, W. Sheng, M. Zhou, B. Fang, F. Luo, and J. Li, "Method for fault diagnosis of temperature-related MEMS inertial sensors by combining hilbert-huang transform and deep learning," *Sensors*, vol. 20, p. 19, 2020.
- [26] P. Liang, C. Deng, J. Wu, and Z. Yang, "Intelligent fault diagnosis of rotating machinery via wavelet transform, generative adversarial nets and convolutional neural network," *Measurement*, vol. 159, Article ID 107768, 2020.
- [27] O. Abdeljaber, O. Avci, S. Kiranyaz, M. Gabbouj, and D. J. Inman, "Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks," *Journal of Sound and Vibration*, vol. 388, pp. 154–170, 2017.
- [28] S. Zhou, M. Xiao, P. Bartos, M. Filip, and G. Geng, "Remaining useful life prediction and fault diagnosis of rolling bearings based on short-time fourier transform and convolutional neural network," *Shock and Vibration*, vol. 2020, Article ID 8857307, 14 pages, 2020.
- [29] E. Levent, I. Turker, and S. Kiranyaz, "A generic intelligent bearing fault diagnosis system using compact adaptive 1D CNN classifier," *Journal of Signal Processing Systems*, vol. 91, no. 2, pp. 179–189, 2019.
- [30] W. A. Smith and R. B. Randall, "Rolling element bearing diagnostics using the Case Western Reserve University data: a benchmark study," *Mechanical Systems and Signal Processing*, vol. 64–65, pp. 100–131, 2015.
- [31] H. Huang and N. Baddour, *Bearing Vibration Data Under Time-Varying Rotational Speed Conditions*, Mendeley Data, Amsterdam, Netherland, 2019.
- [32] X. Chen, B. Zhang, and D. Gao, "Bearing fault diagnosis base on multi-scale CNN and LSTM model," *Journal of Intelligent Manufacturing*, vol. 32, no. 4, pp. 971–987, 2020.
- [33] P. B. Mallikarjuna, M. Sreenatha, S. Manjunath, and N. C. Kundur, "Aircraft gearbox fault diagnosis system: an approach based on deep learning techniques," *Journal of Intelligent Systems*, vol. 30, no. 1, pp. 258–272, 2020.
- [34] J. Liu, Q. Li, H. Yang, Y. Han, S. Jiang, and W. Chen, "Sequence fault diagnosis for PEMFC water management subsystem using deep learning with t-SNE," *IEEE Access*, vol. 7, pp. 92009–92019, 2019.
- [35] W. Zhang, G. Peng, C. Li, Y. Chen, and Z. Zhang, "A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals," *Sensors*, vol. 17, no. 2, p. 425, 2017.
- [36] M. F. Yaqub, I. Gondal, and J. Kamruzzaman, "Inchoate fault detection framework: adaptive selection of wavelet nodes and cumulant orders," *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 3, pp. 685–695, 2012.
- [37] Q. Hu, Z. He, Z. Zhang, and Y. Zi, "Fault diagnosis of rotating machinery based on improved wavelet package transform and SVMs ensemble," *Mechanical Systems and Signal Processing*, vol. 21, no. 2, pp. 688–705, 2007.
- [38] T. Han, D. Jiang, Q. Zhao, L. Wang, and K. Yin, "Comparison of random forest, artificial neural networks and support

- vector machine for intelligent diagnosis of rotating machinery,” *Transactions of the Institute of Measurement and Control*, vol. 40, no. 8, pp. 2681–2693, 2018.
- [39] A. Khorram, M. Khalooei, and M. Rezghi, “End-to-end CNN + LSTM deep learning approach for bearing fault diagnosis,” *Applied Intelligence*, vol. 51, no. 2, pp. 736–751, 2020.
- [40] G. Xu, M. Liu, Z. Jiang, D. Söffker, and W. Shen, “bearing fault diagnosis method based on deep convolutional neural network and random forest ensemble learning,” *Sensors*, vol. 19, p. 5, 2019.
- [41] W. Zhang, L. Chuanhao, P. Gaoliang, C. Yuanhang, and Z. Zhujun, “A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load,” *Mechanical Systems and Signal Processing*, vol. 100, pp. 439–453, 2017.
- [42] H. Huang, N. Baddour, and M. Liang, “Bearing fault diagnosis under unknown time-varying rotational speed conditions via multiple time-frequency curve extraction,” *Journal of Sound and Vibration*, vol. 414, pp. 43–60, 2018.