Extraction of Local and Global Features by a CNN-LSTM Network for Diagnosing Bearing Faults

Zhang Chao [1*,5], Wang Wei-zhi [1,5], Zhang Chen [1,5], Fan Bin [2], Wang Jian-guo [1,5], Fengshou Gu [3], Yu Xue [4]

1. School of Mechanical Engineering , Inner Mongolia University of Science and Technology , Baotou 014010 , China

2. College of Mechanical & Electrical Engineering, Inner Mongolia Agricultural University, Hohhot, 010018, China.

3. School of computing and Engineering, University of Huddersfield, Huddersfield, HD1 3DH, UK

4. Beijing Tianrun New Energy Investment Co., Ltd. Beijing 100000, China

5. Inner Mongolia Key Laboratory of Intelligent Diagnosis and Control of Mechatronic System, Baotou, 014010, China

Abstract: Accurate and reliable fault diagnosis is one of the key and difficult issues in mechanical condition monitoring. In recent years, Convolutional Neural Network (CNN) has been widely used in mechanical condition monitoring, which is also a great breakthrough in the field of bearing fault diagnosis. However, CNN can only extract local features of signals. The model accuracy and generalization of the original vibration signals are very low in the process of vibration signal processing only by CNN. Based on the above problems, this paper improves the traditional convolution layer of CNN, and builds the learning module (local feature learning block, LFLB) of the local characteristics. At the same time, the Long Short-Term Memory (LSTM) is introduced into the network, which is used to extract the global features. This paper proposes the new neural network —improved CNN-LSTM network. The extracted deep feature is used for fault classification. The improved CNN-LSTM network is applied to the processing of the vibration signal of the faulty bearing collected by the bearing failure laboratory of Inner Mongolia University of science and technology. The results show that the accuracy of the improved CNN-LSTM network on the same batch test set is 98.75%, which is about 24% higher than that of the traditional CNN. The proposed network is applied to the bearing data collection of Western Reserve University under the condition that the network parameters remain unchanged. The experiment shows that the improved CNN-LSTM network has better generalization than the traditional CNN.

Key words: Improved CNN-LSTM network; LFLB; LSTM; Pattern recognition; Fault diagnosis

* Corresponding author at: School of Mechanical Engineering, University of Science and Technology of the Inner Mongol, China.

E-mail address: zhanghero@imust.edu.cn (Zhang Chao).

Introduction

Bearing is one of the basic components of rotating machinery, and 45%-55% of machine shutdowns are caused by bearing faults [1]. Therefore, an effective bearing fault diagnosis method is the top priority to ensure the normal work of the machine. However, bearing fault diagnosis is always a difficult problem in the field of fault diagnosis. So far, there are two main methods about bearing fault diagnosis: vibration signal analysis method and data-driven method.

Vibration signal analysis mainly uses signal decomposition techniques, such as empirical mode decomposition and wavelet analysis, and observe the changes and trends of state indicators, or use frequency characteristics for bearing fault diagnosis. However, it is often sensitive to changes in the physical properties of mechanical systems [2]. For complex mechanical structures which especially in the strong noise environment, it is difficult to accurately detect the fault characteristic frequency or observe the abnormal change of the index.

In recent years, most of the intelligent diagnosis methods which are data-driven have been applied in the field of bearing fault diagnosis. For example, Lei et al. [3] used 10 statistical parameters to describe the healthy state of bearings and input these characteristics to the artificial neural network for fault classification. Yang et al. [4] took the component energy entropy which obtained from empirical mode decomposition as a feature and input it into the neural network for bearing fault classification. However, the above intelligent fault diagnosis methods are all based on shallow machine learning models, such as artificial neural network (ANN), k-nearest neighbor algorithm (KNN), random forest and support vector machine (SVM). Using these shallow machine learning models for bearing fault diagnosis requires artificial extraction of a series of vibration signal features (for example, variance, energy entropy, root mean square, etc.), and the extracted features are used as model input for fault classification by using classifier. Previous research results show that feature extraction is an important prerequisite for achieving ideal diagnostic accuracy. But in fact, it is difficult to artificially extract effective characteristics of bearing fault feature, basically has the following three difficult problems [5]: (1) need high level professional knowledge, (2) in difficult mechanical system to eliminate the interference of hardware and the changeable environment, the extracted features should be changeable, (3) manual extraction features may lead to low data utilization, resulting in lack of effective information. Therefore, the above classical shallow model is not suitable for bearing fault diagnosis.

In order to avoid the above problems, in recent years, deep learning methods based on pattern recognition and image processing [6] have attracted the attention of many researchers in the field of fault diagnosis, and promoted the development of fault diagnosis methods based on deep learning. Deep learning imitated human brain learning process, adopted hierarchical structure, extracted high-level features of data layer by layer through greedy learning algorithm, and then used the learned features as the input of classifier to conduct fault classification. Some well-performing deep learning networks have been proposed, such as deep belief nets (DBNs) and restricted boltzmann machine (RBM), etc. These methods have achieved good results in highly challenging recognition tasks [7], [8]. In recent years, in order to improve the efficiency of fault diagnosis, these methods have also been used in various fault diagnosis [9-11].

However, in the above deep learning network learning process, when the input is the original vibration signal or spectrum diagram, many parameters need to be optimized, resulting in a very slow network training process, and it is difficult to realize the optimal model. Based on this problem, CNN was proposed and widely concerned by various fields. Compared with other deep learning models mentioned above, CNN adopts weight sharing, local perception and sub-sampling strategies, which can significantly reduce the number of parameters to be optimized (also known as weights), greatly improving the speed of network training. Meanwhile, because CNN is insensitive to local changes in the convolution process, CNN has a strong noise reduction ability. As one of the main deep learning models, CNN has been successfully applied in many fields, such as medical imaging [12], food classification [13], stock prediction [14] and emotion recognition [15]. At present, many researchers also use CNN for fault diagnosis [16], [17], [18], and have achieved good results.

However, CNN can only extract local features of signals when processing temporal signals [18]. Bearing vibration signal is a time-varying signal. The local features extracted from such signals by CNN can loss the global information of the signal, resulting in low performance of diagnostics. To overcome the shortage, this paper puts forward the CNN-LSTM network. In the network, the inputs are the original vibration signals of bearings. The network has been created based on local feature extraction module (LFLB) of CNN. The local characteristics of the original vibration signal can be extracted from the four LFLB layers. At the same time, the long short-term memory (LSTM) network can extract the global characteristics of vibration signals. Experimental results show that the method can effectively improve the accuracy of bearing fault diagnosis, compared with the simple use of CNN, and has a certain generalization.

At present, scholars have proposed a method of bearing fault diagnosis combined with CNN-LSTM, but most of them only use a set of data sets, which are divided into training set and test set in proportion, one part is used to train the network, and the other part is used to test the accuracy of the network. degree. By improving the feature extraction part based on the traditional CNN network architecture, a batch normalization layer and an exponential linear unit layer are added to the traditional CNN convolution module, which are common with a convolution layer and a maximum pooling layer The local feature learning module is formed, which can improve some of the drawbacks of traditional CNN in extracting signal local features. The local feature extractor LFLBs is constructed by stacking 4 layers of LFLB modules, and the LFLBs are used to learn the local features in the bearing vibration signal. During the training process, the convolution kernel determines the retention and discarding of the extracted features according to the target. Therefore, compared with the traditional manual feature extraction, the local feature extraction of the deep network can be automatically completed, but the interpretability of the extracted features is relatively weak. The BN layer normalizes the features extracted by the convolutional layer, which can improve the stability of the network. Compared with the traditional CNN network, the extracted features are more representative.

In addition, Wang [19] et al. studied model-based fault diagnosis and found that the existence of unknown disturbances would cause fault alarms. In order to deal with this false alarm problem, the proposed robust fault detection method is designed. The H−/H∞ fault detection of LPV descriptor system is studied. Based on the generalized Kalman-Yakubovich-Popov lemma of the LPV description subsystem, a non-single-point structure finite frequency domain fault detection observer is designed. The sufficient conditions for the design of the fault detection observer are derived and converted into a set of LMI, which can be solved effectively. LFLB is constructed by layers that are robust to input noise. Therefore, the features learned by LFLB have certain robustness and can also avoid failure alarms caused by unknown disturbances. Liu [20] et al. studied the design of sliding-mode observers for actuator fault estimation in linear continuous-time systems under digital communication channels. This problem often occurs in a network environment. Before transmitting data through a digital communication channel, the data must be quantified. Aiming at this problem, a new descriptor sliding mode observer method is proposed. The results show that if the quantizer density is greater than $\sqrt{2}-1$, the designed observer can fully compensate for the quantization error, and the fault vector can be reconstructed with signal quantization. Syed Abu Nahian [21] et al. proposed a UIOEFIR algorithm with minimum design workload for EHA's random sensor failure and state estimation problems. This is achieved by introducing UIO into EFIR dynamics, so there is no requirement for noise statistics. For this reason, an effective sensor FDA that can be executed with any controller is developed using the proposed estimator to tolerate the position of the EHA and the failure of the head-side pressure sensor. In operation, this method effectively uses conventional control logic for trajectory tracking, and improves the estimation accuracy under time-varying measurement noise. Since the estimated amount of the design is Nopt times slower than the EKF-based method, techniques to reduce the computational cost will be studied. It can be seen that the improved CNN-LSTM neural network, because CNN is not sensitive to local changes in the convolution process, so it has a strong noise reduction ability, so that the results of fault diagnosis are more accurate. The combination of improved CNN and LSTM will also greatly reduce the cost of calculation, and at the same time, there is no need to perform any processing on the data, which has better generalization.

1 Methods to introduce

1.1 Local feature learning blocks-LFLBs

Local feature learning block (LFLB) is a substitute for CNN, aiming to improve some disadvantages of traditional CNN when extracting local features of signals. After dozens of experiments, it is concluded that four LFLB modules are the optimal solution with the strongest feature extraction ability of the network model proposed in this paper. Local feature extractor, LFLBs, is constructed by superimposing 4 layer LFLB modules, and LFLBs is used to learn local features in bearing vibration signals. The LFLB layer is composed of a convolution, a batch of normalized (BN) layer [22], an exponential linear unit layer (ELU) [23] and one of the max pooling layer (MP), whose overall structure is shown in Figure 1. The convolution layer and pooling layer are the core layers of LFLB, whose working principle is shown in Figure 2 and Figure 3. Due to the two characteristics of local spatial connectivity and sharing [24], the convolution layer reduces the number of parameters that need to be updated when training the network and speeds up the operation. At the same time, these two features enable the convolution layer to have functions similar to human visual system, so that deep features of input data can be learned. During the training process, the convolution kernel changes randomly according to the network parameters. When the feature extracted from the convolution kernel cannot represent the object effectively, the feature will be discarded. If the feature is a valid representation of the object, the feature will be retained. Therefore, local feature extraction of deep network can be completed automatically, but the interpretation of extracted features is relatively weak. BN layer normalizes the output features of convolution layer after activation, and improves the performance and stability of deep network. The conversion used in batch normalization keeps the activation mean near 0 and the activation standard deviation near 1. The ELU layer defines the output of the BN layer. Different from other activation functions, ELU has a negative value, which makes the mean value of activation approach to zero, thus speeding up the learning speed of the design network and improving the recognition accuracy. MP can reduce the noise interference to the fault feature extraction [25]. Among the common nonlinear functions, maximum pooling is the most common. It divides the input into a group of non-overlapping regions and outputs the maximum value of each sub-region [26]. Although the robustness of LFLB has not been proved mathematically, LFLB is constructed from a layer that is robust to input noise, so the characteristics learned by LFLB have certain robustness.

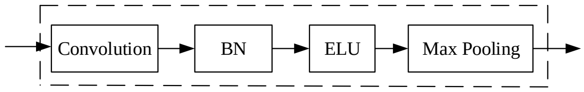

Figure 1 LFLB module

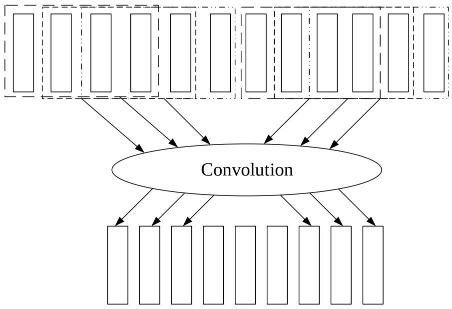

Figure 2 Convolution layer convolution process with kernel size of 4 and step size of 1

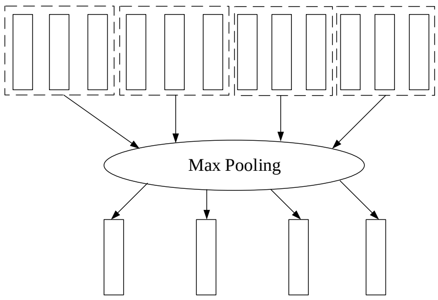

Figure 3 Maximum pooling layer pooling process with a kernel size of 3 and a step size of 3

LFLB can be adjusted according to different tasks. The differences of LFLB construction are mainly reflected in various parameters of convolution and max pooling. The convolution layer plays a role in local feature extraction. When the data is transferred to the convolution layer, it is convolved with the convolution kernel according to the width and height of the input data. Than the eigenmatrix is generated by computing the dot product between the kernel item and the input item.

The signal *x(n)* is input into convolution kernel w(n) of size l to obtain the result *z(n)*. The convolution kernel *w(n)* is randomly initialized in the experiment.

$$z(n) = x(n) * w(n) = \sum_{m=-l}^{l} x(m) \cdot w(n-m) \quad (1)$$

Then the features obtained by convolution are input into BN layer, so that these features can be normalized after activation. After normalization, the average value of these features is close to 0 and the standard deviation is close to 1. The normalized features can be expressed as:

$$z_i^l = \sigma(BN(b_i^l + \sum_j z_j^{l-1} * w_{ij}^l)) \quad (2)$$

Where, z_i^l and z_j^{l-1} represent the output feature of layer l and the j input feature of layer (l-1). w_{ij}^l is the convolution kernel between the i and j feature.Function $BN(*)$ is the activation function of batch normalization layer. Function $\sigma(*)$ is the activation function of ELU layer.

After processing by function $BN(*)$ the features obtained by the convolution layer have been normalized. Function $\sigma(*)$ is the activation function of ELU layer, which can be expressed as:

$$\sigma(x) = \begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases} \quad (3)$$

The acquired features are transmitted to the maximum pooling layer, which has the function of nonlinear down-sampling and can reduce the feature resolution. The features generated by the maximum pooling layer can be expressed as:

$$z_k^l = \max_{\forall p \in \Omega_k} z_p^l \quad (4)$$

including Ω_k representative indexes for *j* pooling areas.

1.2 Global feature extractor—LSTM

The global characteristics of bearing vibration signal refer to the timing characteristics of bearing. For timing signals like bearing vibration signals, global features are related to the whole sequence, and the detection algorithm should be applied to the whole time domain of the signal. In order to obtain global fault feature information in bearing vibration signal, LSTM is used to extract global feature. In recent years, LSTM has been applied to many fields such as language processing, audio analysis and image analysis, and achieved very good results. LSTM is a variant of the recurrent neural network (RNN), which can effectively avoid the phenomenon of gradient disappearance or gradient explosion caused by RNN. So it is often used to process and predict relatively long interval and delay sequence signals in time series [23]. LSTM is used to extract global features of signals, which can enhance the effectiveness of features and better diagnose bearing faults.

$$f_t = \sigma_g(W_f z_t^{l-1} + U_f z_t^{l-1} + b_f) \quad (5)$$

$$i_t = \sigma_g(W_i z_t^{l-1} + U_i z_{t-1}^l + b_i) \quad (6)$$

$$o_t = \sigma_g(W_o z_t^{l-1} + U_o z_{t-1}^l + b_o) \quad (7)$$

$$c_t = f_t c_{t-1} + i_t \sigma_c(W_c z_t^{l-1} + U_c z_{t-1}^l + b_c) \quad (8)$$

$$z_t^l = o_t \sigma_h c_t \quad (9)$$

Where c_t represents the state of LSTM module; W,U and b are parametric matrices and vectors; f_t, i_t and o_t are gate vectors; σ_g is sigmoid function, σ_c and σ_h are hyperbolic tangent lines; The operator F represents the Hadamard product.

1.3 CNN - LSTM network

In this paper, CNN-LSTM network is constructed to process bearing vibration signals. Local features of bearing vibration signals are extracted by superimposing four LFLBS, and global features are extracted from local features of LFLB4 output by LSTM. Its flow chart is as follows in Figure 4, in which RB stands for rolling body fault bearing，IR stands for Inner ring fault bearing, OR stands for Outer ring fault bearing，Cage stands for cage fault bearing.

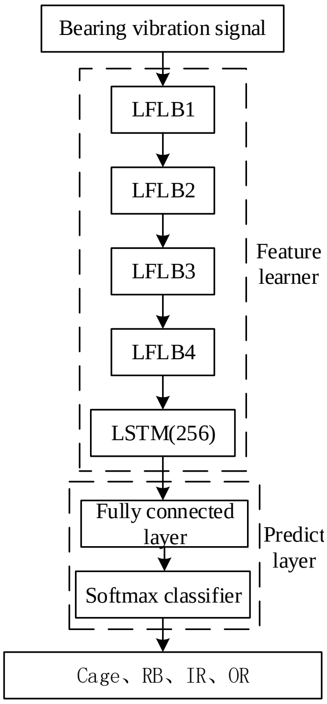

Figure 4 CNN - LSTM flow chart

The purpose of this network is to learn deep features from bearing vibration signals. Bearing vibration signals are one-dimensional, so the convolution kernel and pooling kernel of each LFLB module are one-dimensional, and the convolution kernel size of each LFLB module is 3, the step size is 1, and the filling mode is the SAME. The number of convolution kernel of LFLB1 and LFLB2 is 64, and the number of convolution kernel of LFLB3 and LFLB4 is 128. The maximum pooled kernel size and step size in each LFLB module is 4. The network LSTM layer kernel number is 256, which is used to extract global features. Network parameters are shown in Table 1. Softmax classifier is used at the top of the network, which can identify bearing fault types according to the deep features learned.

Table 1 CNN-LSTM network parameters (C is the convolution layer, P is the pooling layer)

Module		Number	Size	Stride
LFLB1	C1	64	3	1
	P1		4	4
LFLB2	C2	64	3	1
	P2		4	4
LFLB3	C3	128	3	1
	P3		4	4
LFLB4	C4	128	3	1
	P4		4	4
LSTM			256	

When bearing vibration signals are input into CNN-LSTM network, LFLBs first extracts local features from vibration signals data. Local features of the output of LFLB4 are reconstructed and input into the LSTM layer, from which the LSTM learns global features. The learning process of local and global features is shown in Figure 5 and Figure 6. Therefore, the deep fault features output from the LSTM layer contain local and timely sequence information of fault signals.

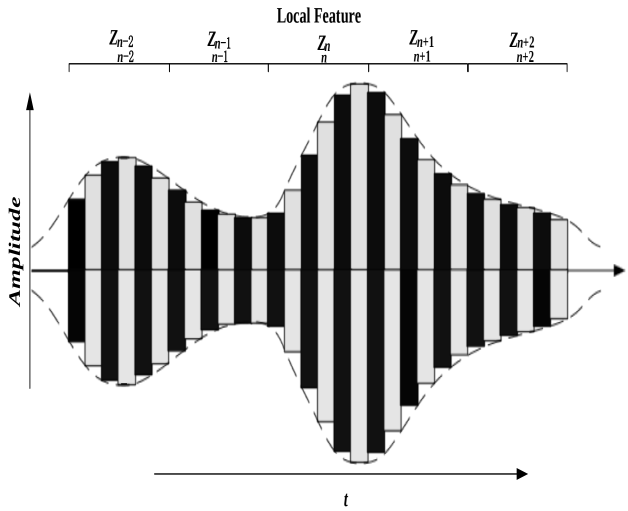

Figure 5 Local feature extraction process

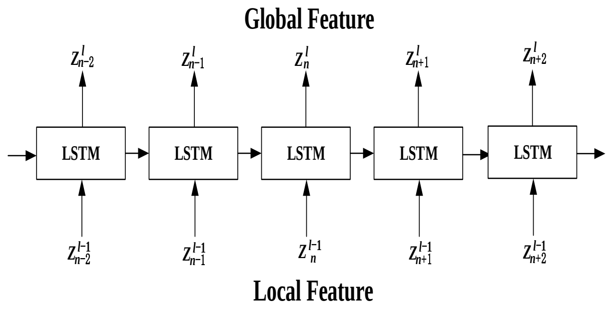

Figure 6 Global feature extraction process

The learned deep-level features are input into the full connection layer connected with the LSTM layer, and the output fault features of the full connection layer can be expressed as follows:

$$z^l = b^l + z^{l-1} \cdot w^l \tag{10}$$

Softmax is a kind of classifier, the multiple classification problems, the tag *z* contain two or more values. The Softmax function can be defined as:

$$z_i = \sum\nolimits_j h_j W_{ji} \tag{11}$$

$$soft\max(z)_i = p_i = \frac{\exp(zi)}{\sum_{j=1}^{n} \exp(z_j)} \tag{12}$$

Where z_i is the input of Softmax, h_j is the activation of the penultimate layer, and W_{ji} is the weight between the penultimate layer and the Softmax layer. Therefore, the prediction class tag is:

$$\hat{y} = \arg\max p_i \tag{13}$$

1.4 Hyperparameter optimization

It is very important to choose a set of hyperparameters for a deep architecture. The goal of hyperparameter optimization is to improve the performance of the deep network on an independent data set. Grid search and random search all have been applied to some deep learning frameworks successfully, and simplify the training of the deep architecture. When Bayesian optimization is proposed, it is shown to obtain better results in fewer experiments. In our experiments, the Bayesian optimization method is adopted to select the hyperparameters for the proposed deep networks.

Bayesian optimization is a sequential design strategy and can minimize the objective function efficiently. Hyperopt, a Python library, is used to optimize the hyperparameters in our experiments. Hyperopt defines an objective function which can be minimized, and treats it as a random function. A prior is also placed over the objective function. According to the gathered function evaluations, prior is updated to form the posterior distribution over the objective function. An acquisition function is created by using the posterior distribution. Then the hyperparameters are picked iter-atively. In order to select a suitable optimization algorithm, the distribution over the choice ('adagrad', 'adam', 'sgd', 'rmsprop') is chosen. After the training with the optimized hyperparameters,the best model is returned.

2 Bearing failure test

2.1 Data acquisition and database description

The fault bearing vibration signal was collected from the HZXT-DS-003 double-span double-rotor rolling bearing test rig at Inner Mongolia University of Science and Technology, as shown in Figure 7.

The bearings tested are 6205 deep groove ones, specified as shown in Table 2. Four types of components faults are induced manually to four bearing s respectively, which are shown in Figure 8 and expressed as OR for the fault on outer, IR for fault on inner race, Cage for fault on cage and RB for fault on ball elements respectively.

To ensure data representativeness for CNN training and testing, vibration signals were collected at six different speeds: 750, 1000, 1250, 1500, 1750 and 2000 r/min. at each speed, 800 vibration signals, each with a length of 1s were collected with a sampling rate of 12,000Hz which results in 800×6×5 datasets in total for 6 operating speeds and 5 bearing cases (including fault free).

2.2 Data processing

2.2.1 The application of traditional CNN

Traditional CNN is used to process the data set. The network consists of two convolution layers and two pooling layers, which are arranged alternately. The network architecture is shown in Figure 9, which consists of two convolution stages: CNN1 and CNN2 detailed are shown in Table 3.

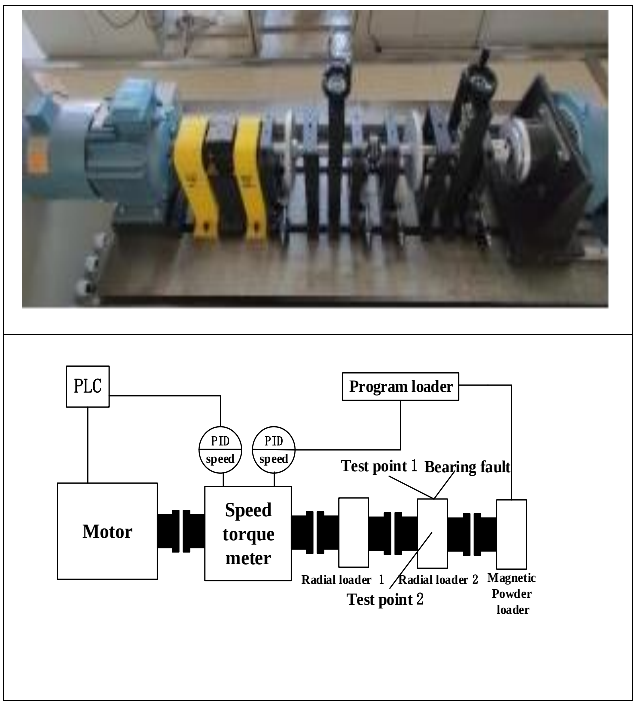

Figure 7 Hzxt-ds-003 double span double rotor rolling bearing test rig

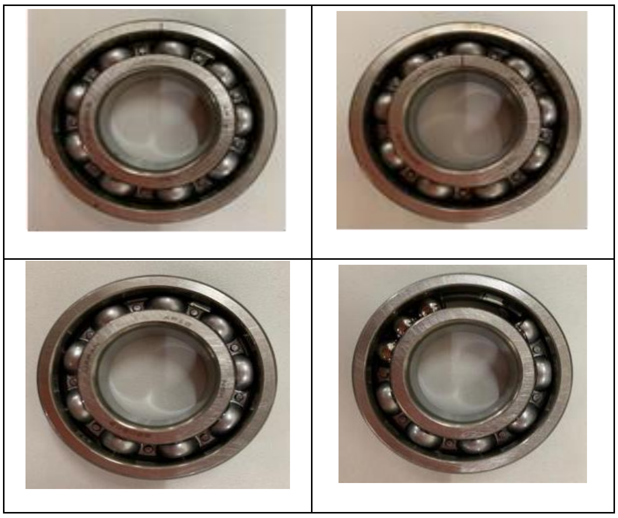

Figure 8 Failure bearing of outer ring (OR), failure bearing of inner ring (IR), failure bearing of rolling body (RB), failure bearing of cage (Cage)

Table 2 6205 Bearing parameters (D- Section bearing diameter, d- Roll segment diameter, N- Number of rollers)

D	d	N	Angle
39.04mm	7.94 mm	9	0°

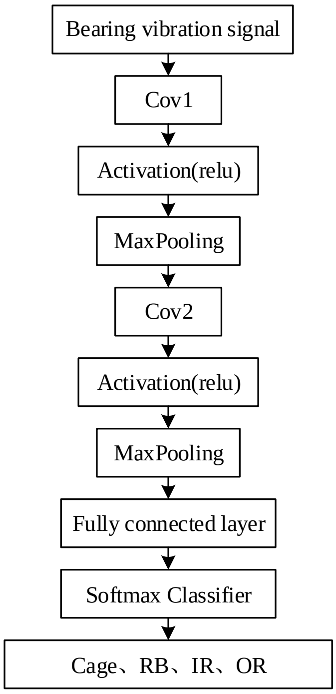

Figure 9 Traditional CNN architecture diagram

Table 3 CNN parameters

Module	number	size	Stride
CNN1	C1 32	3	1
	P1	4	4
CNN2	C2 64	3	1
	P2	4	4

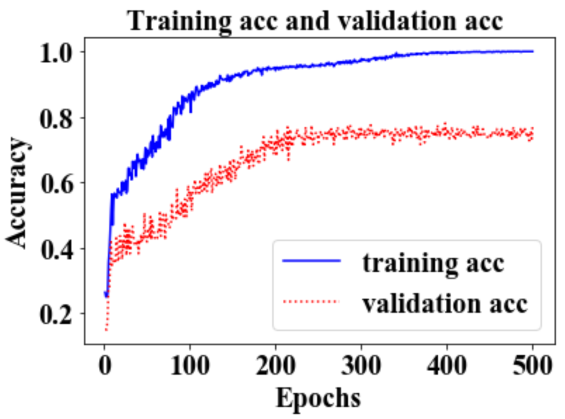

Figure 10 Changes of training accuracy and verification accuracy of traditional CNN with the number of iterations

It is divided into a training set, validation set and test set according to the ratio of 8:2:2.

Experimental results are shown in Figure 10. As can be seen from the results, using the traditional CNN to analyze the data, the number of iterations to reach the optimum is about 230, and the validation accuracy is only 76.33%. The accuracy of test set is 75.26%. Generally, the higher these two accuracy rates are, the better the network performance will be. The validation set accuracy is usually used to adjust the parameters of the network. The more accurate the test set is, the more effective the network will be. Obviously, the traditional CNN is not suitable for bearing vibration signal processing.

2.2.2 Improved CNN application on the data set

The four-layer LFLB neural network is used to process the data set. The network architecture is shown in Figure 11, and the network parameters are the same as those in Table 1. The processing results are shown in Figure 12. When the number of iterations reaches 34, the model reaches the optimal value, and the accuracy of the verification set is the highest at 83.57%. The test set is used to test the model, and the bearing fault confusion matrix obtained is shown in Table 4, and the accuracy of the test set is 80.95%. The results show that compared with the experimental results of traditional CNN, the number of iterations used to reach the optimal network is greatly reduced, indicating that the improved CNN computing cost is greatly reduced. In terms of accuracy, it also increases by about 5% from 75.26%, and the improved CNN performance is obviously better than traditional CNN.

However, the accuracy of 80.95% is far from the expected effect, and the accuracy of the outer circle is only 59.53%.

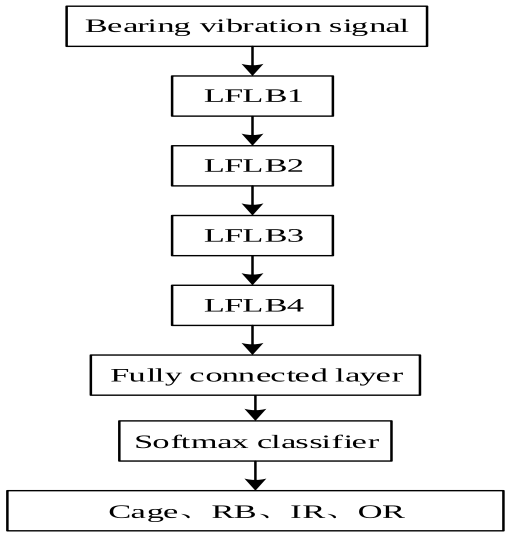

Figure 11 LFLB neural network architecture

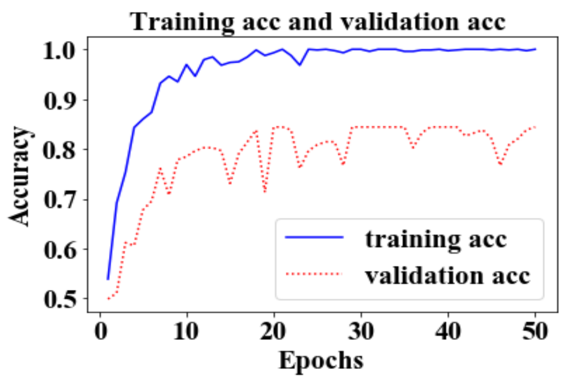

Figure 12 Changes of improved CNN training accuracy and verification accuracy with the number of iterations

Table 4 Bearing fault confusion matrix（accuracy%）

	Cage	RB	IR	OR
Cage	92.85	7.15	0	0
RE	4.76	95.24	0	0
IR	11.5	12.30	76.20	0
OR	0	0	40.47	59.53

Average accuracy =80.95%

2.2.3 The application of improved CNN-LSTM network on this data set

The improved CNN-LSTM network is used to process the dataset. Different from the improved CNN, the improved CNN-LSTM network adds the LSTM layer, which can improve the model performance by extracting global features of signals. The processing results are shown in Figure13. When the number of iterations reaches 32, the model reaches the optimal value, and the accuracy of the verification set reaches 98.75%. Test set is used to test the model, and the test accuracy is 97.45%. The bearing fault confusion matrix is shown in Table 5.

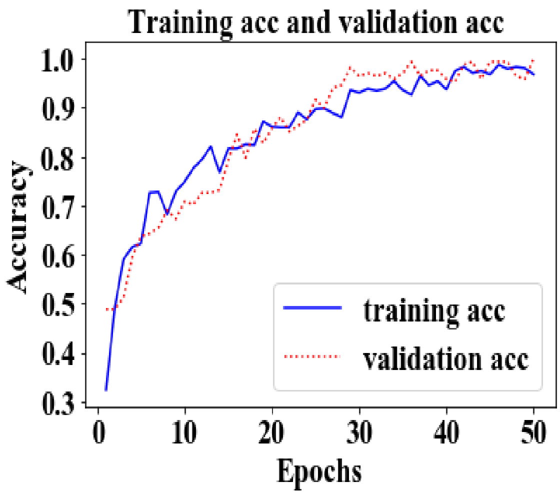

Figure 13 Changes of training and verification accuracy of CNN-LSTM neural network with the number of iterations

Table 5 Bearing fault confusion matrix（accuracy%）

	Cage	RB	IR	OR
Cage	100	0	0	0
RE	0	100	0	0
IR	0	0.36	99.64	0
OR	0	0	9.84	90.16
Average accuracy =97.45%				

From the perspective of computational cost, the number of iterations required by the model to reach the optimal value is similar to the number of iterations required by the improved CNN, which has a lower operational cost than traditional CNN. From the accuracy of test set, the improved CNN-LSTM network has a higher accuracy, which is about 16% higher than the improved CNN. Among them, the ability to identify faults in the outer circle is greatly enhanced, reaching 90.16%, which is about 30% higher than the improved CNN. It can be seen that the improved CNN-LSTM network has better performance and can be well applied to bearing fault diagnosis.

3 The generalization of improved CNN-LSTM network and its performance compared with other methods

In order to verify the generalization of the improved CNN-LSTM network, CNN-LSTM network is applied to the bearing data set of Western Reserve University. For comparison, traditional CNN is used to process the bearing data of Western Reserve University, and the performance is compared with other two advanced methods under the same data set. The bearing faults are inner ring faults, outer ring faults and rolling body faults. The test used electric discharge machining (EDM) technology to arrange a single point of failure on the outer ring of the bearing (3 o 'clock, 6 o 'clock, 12 o 'clock direction), the inner ring and the rolling body, and the failure diameter was 0. 007, 0. 014, 0. 021, 0. 028, 0. 040 inches (1 inch = 2. (54 cm). Each state was tested under loading conditions of 0, 1, 2 and 3HP(1HP = 746W), and the rotation speed was corresponding to 1797, 1772, 1750 and 1730 r/min, respectively. The sampling frequency of the signal is 12000Hz. On the premise of constant parameters, the improved CNN-LSTM network proposed in this paper is applied to the bearing vibration data collected under the condition of all new, load change and rotation speed fluctuation, which can well verify the generalization of the constructed network. Each sample of three kinds of fault bearing vibration data of Western Reserve University was cut to 10 sub-samples with a size of

12000 points. Randomly, the sub-sample set was divided into training set, verification set and test set in a ratio of 8:2:2.

3.1 The application of traditional CNN on this data set

The data set of bearing vibration signal of Western Reserve University is input into traditional CNN, and the results are shown in Figure 14. When the model is trained to the best, the accuracy of the verification set is 63.74%, and the accuracy of the test set is only 58.65%, when the test set is input for testing. It can be seen that the traditional CNN is not good at processing bearing vibration signals.

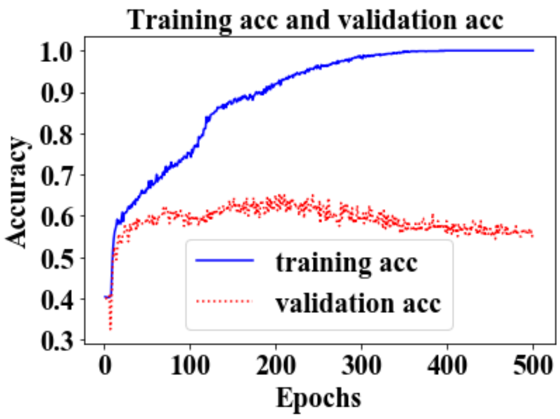

Figure 14 Changes of training accuracy and verification accuracy of traditional CNN with the number of iterations

3.2 The application of improved CNN-LSTM network on this data set

The improved CNN-LSTM network proposed in this paper is used to analyze the data set of bearing vibration signal of Western Reserve University. In order to verify its generalization, the test is carried out under the condition of hyperparameters and constant parameters. The obtained results are shown in Figure 15. When the model is optimal, the accuracy rate of the verification set can reach 98.50%, and the accuracy rate of the test set obtained from the input test set is 98.35%. It can be seen that the improved CNN-LSTM network has higher accuracy rate and certain generalization. The bearing fault confusion matrix is shown in Table 6.

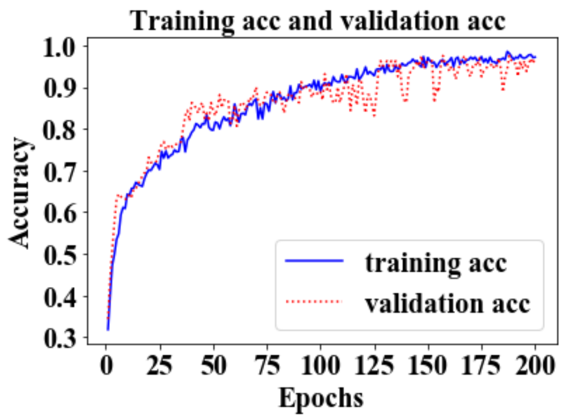

Figure 15 Training and verification accuracy of the improved CNN-SLTM neural network varies with the number of iterations

Table 6 Bearing fault confusion matrix（accuracy%）

	RB	IR	OR
RE	100	0	0
IR	0	100	0
OR	0	8.94	91.06
Average accuracy =98.35%			

3.3 Comparison with other machine learning algorithms

In order to more comprehensively verify the performance of the method in this paper, the proposed method was compared with other machine learning algorithms. For example, Xie et al. [27] preprocessed the original signal with a bandpass filter to get the optimized features, and introduced the Adam optimizer to speed up model training. Then, the improved convolutional deep confidence network was used to complete fault identification. Wang et al. [28] applied wavelet transform to bearing vibration signals, used soft threshold denoising method to filter the noise of vibration signals, and realized fault identification with radial basis (RBF) neural network. This method reduces the manual workload, improves the accuracy and calculation efficiency compared with Xie's method, and the parameters are easy to set. Compared with sparse coding and BP neural network, SAE and DBN, the above two methods have higher accuracy and better performance, and are representative to some extent.

The identification results of the three methods are shown in Table 7. It can be seen that the method proposed in this paper has higher accuracy and does not need any data processing. Compared with the other two algorithms, the method has better generalization.

Table 7 Performance comparison of three algorithms

Methods	Accuracy	Pretreatment method
The paper,	98.35%	None
Xie[27]	95.64%	Fast Fourier transform、
		Folding processing、
		Bandpass filter
Wang[28]	96.75%	Wavelet transform、
		Soft threshold denoising

4 Conclusion

Using the improved CNN-LSTM network for bearing fault diagnosis can avoid the manual feature extraction process and reduce the diagnosis cost. Moreover, because it is searching for features in massive data, it has higher generalization. In this paper, combined with the improved CNN and LSTM neural network, the local and global features of bearing vibration signals can be effectively learned. Compared with the traditional CNN network, the extracted features are more representative, and the diagnosis effect is far better than the traditional CNN neural network. Compared with the other two methods, the experimental results show that the designed network has better performance in processing bearing vibration signal data. However, there are also many shortcomings, such as the number of network layers and the selection of various parameters, which are obtained through continuous experiments and have poor explanatory ability, which is also one of the key research contents in the future.

Acknowledgments

Not applicable.

Data Availability

The experiment is carried out on Hzxt-ds-003 double span double rotor rolling bearing test rig as shown in Figure7. The experimental data can beobtained by sending an email to zhanghero@imust.edu.cn.

Authors’ Contributions

CZ was in charge of the whole trial; WZW wrote the manuscript and assisted with sampling and laboratory analyses. All authors read and approved the fnal manuscript.

Authors’ Information

Zhang chao, born in 1978, in 2012, he received a doctor's degree in engineering in Xi 'an University of Electronic Science and Technology. His research interests include fault diagnosis of rotating machinery and vibration signal processing.

Funding

This research is supported by National Natural Science Foundation of China (51565046, 51965052, 51865045), Science and Technology Plan Project of Inner Mongolia Autonomous Region, China (KJJH007).

Competing Interests

The authors declare no competing financial interests.

Author Details

1.School of Mechanical Engineering , Inner Mongolia University of Science and Technology , Baotou 014010 , China. 2.College of Mechanical & Electrical Engineering, Inner Mongolia Agricultural University, Hohhot, 010018, China. 3.School of computing and Engineering, University of Huddersfield, Huddersfield, HD1 3DH, UK. 4.Beijing Tianrun New Energy Investment Co., Ltd. Beijing 100000, China. 5.Inner Mongolia Key Laboratory of Intelligent Diagnosis and Control of Mechatronic System, Baotou, 014010, China

References:

[1] Nandi S , Toliyat H A , Li X (2005) Condition monitoring and fault diagnosis of electrical motors-a review. IEEE Transactions on Energy Conversion 20(4):719-729.

[2] Luyang Jing, Ming Zhao, Pin Li, Xiaoqiang Xu (2017) A convolutional neural network based feature learning and fault diagnosis method for the condition monitoring of gearbox. Measurement 111:1-10

[3] Lei Y , He Z , Zi Y (2011) EEMD method and WNN for fault diagnosis of locomotive roller bearings. Expert Systems with Applications 38(6):7334-7341.

[4] Yu Y , YuDejie, Junsheng C (2006) A roller bearing fault diagnosis method based on EMD energy entropy and ANN. Journal of Sound and Vibration 294(1-2):269-277.

[5] Haidong S, Hongkai J, Fuan W, Huiwei Z (2017) An enhancement deep feature fusion method for rotating machinery fault diagnosis. Knowledge-Based Systems 119:200-220.

[6] G.E. Hinton, R.R. Salakhutdinov (2006) Reducing the dimensionality of data with neural networks. Science, 313:504-507.

[7] P Vincent, H Larochelle, Y Bengio, P-A Manzagol (2008) Extracting and Composing Robust Features with Denoising Autoencoders. International Conference on Machine Learning. Helsinki, Finland.

[8] Graves A , Santiago Fernández, Gomez F (2006) Connectionist temporal classification: Labelling unsegmented sequence data with recurrent neural networks. International Conference on Machine Learning. Pittsburgh, Pennsylvania, USA.

[9] Feng J, Yaguo L , Jing L, Xin Z, Na L (2016) Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. Mechanical Systems and Signal Processing 72-73:303-315.

[10] Chen H, Wang J, Tang B, Ke X, Junyang Li (2017) An integrated approach to planetary gearbox fault diagnosis using deep belief networks. Measurement Science and Technology 28(2):025010.

[11] Yaguo L, Feng J Jing L, Saibo X, Steven X Ding (2016) An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data. IEEE Transactions on Industrial Electronics 63(5):3137-3147.

[12] Dou Q, Chen H, Yu L, Qin J, Pheng-Ann H (2017) Multi-level Contextual 3D CNNs for False Positive Reduction in Pulmonary Nodule Detection. IEEE Transactions on Biomedical Engineering 64(7):1558-1567.

[13] Ciocca G, Napoletano P, Schettini R (2017) Learning CNN-based Features for Retrieval of Food Images. New Trends in Image Analysis and Processing – ICIAP:426-434.

[14] Hoseinzade E, Haratizadeh S (2019) CNNPred: CNN-based stock market prediction using a diverse set of variables Expert Systems with Applications 129:273-285.

[15] Janfeng Z, Xia M, Lijiang C (2019) Speech emotion recognition using deep 1D & 2D CNN LSTM networks. Biomedical Signal Processing and Control 47:312-323

[16] Liang G, Yaguo L, Naipeng L, Tao Y, Ningbo L (2018) Machinery health indicator construction based on convolutional neural networks considering trend burr. Neurocomputing 292:142-150.

[17] Luyang J, Taiyong W, Ming Z, Peng W (2017) An Adaptive Multi-Sensor Data Fusion Method Based on DeepConvolutional Neural Networks for Fault Diagnosis of Planetary Gearbox. Sensors, 17(2):414-428.

[18] Xiaoxi D, Qingbo H (2017) Energy-Fluctuated Multiscale Feature Learning With Deep ConvNet for Intelligent Spindle Bearing Fault Diagnosis. IEEE Transactions on Instrumentation and Measurement 66(8):1926-1935.

[19] Zhenhua Wang , Peng Shi , Cheng-Chew Lim (2017) H−/H∞ fault detection observer in finite frequency domain for linear parameter-varying descriptor systems. Automatica 86:38–45.

[20] M Liu，L Zhang，P Shi，Y Zhao (2018) Fault estimation sliding mode observer with digital communication constraints. IEEE Transactions on Automatic Control 99:1-1.

[21] Syed Abu Nahian, Truong Quang Dinh, Hoang Vu Dao, Kyoung Kwan Ahn (2020) An Unknown Input Observer–EFIR Combined Estimator for Electrohydraulic Actuator in Sensor Fault-Tolerant Control Application. IEEE/ASME Transactions on Mechatronics 5:2208 - 2219.

[22] Ioffe S, Szegedy C (2015) Batch normalization: accelerating deep network training by reducing internal covariate shift. International Conference on International Conference on Machine Learning. Lille, France.

[23] D. Clevert, T. Unterthiner, S. Hochreiter (2016) Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs). International Conference on Learing Representations, San Juan, US.

[24] Palaz D , Collobert R , Doss M M (2013) Estimating Phoneme Class Conditional Probabilities from Raw Speech Signal using Convolutional Neural Networks. Computer Science 21:1-5.

[25] Dan C, Meier U, Schmidhuber J (2012) Multi-column deep neural networks for image classification. Computer Vision & Pattern Recognition 143: 3642-2649.

[26] Gers F A , Schmidhuber, Jürgen, Cummins F (2000) Learning to Forget: Continual Prediction with LSTM. Neural Computation 12(10):2451-2471.

[27] XIE Jiaqi，YOU Wei，SHEN Changqing (2020) Research on Bearing Fault Diagnosis Based on Improved Convolution Depth Confidence Network . Journal of Electronic Measurement and Instrumentation 2 : 36-43 .

[28] Hai-lin WANG，Chun-guang ZHANG，Chao-chen TANG (2020) Research on bearing fault diagnosis based on wave let analysis and RBF neural network． Machine Tool & Hydraulics 48(24) : 182 — 187.