IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 11, 2020, accepted March 29, 2020, date of publication April 3, 2020, date of current version April 20, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2985617

Deep Convolutional and LSTM Recurrent Neural Networks for Rolling Bearing Fault Diagnosis Under Strong Noises and Variable Loads

MEIYING QIAO[1,2], SHUHAO YAN[1], XIAXIA TANG[1], AND CHENGKUAN XU[1]
[1]School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454000, China
[2]Collaborative Innovation Center of Coal Work Safety, Jiaozuo 454000, China

Corresponding author: ShuHao Yan (1452597175@qq.com)

This work was supported in part by the National Natural Science Foundation of China, Research on the measurement technology of rotating MEMS inertia while drilling, The host: Jinxian Yang, Henan Polytechnic University, under Grant 416723639, in part by the National Natural Science Foundation of China, Theory and application of iterative learning control in quantized set-value system, The host: Xuhui Pu, Henan Polytechnic University, under Grant 61573129, in part by the Natural Science Foundation of China-Henan, Research on risk identification and evaluation of key influencing factors of human error in coal mine accidents, The host: Jianyi Lan, Henan Polytechnic University, under Grant 172102310239, and in part by the Science and Technology Project in Henan Province, Research on Online Monitoring System of Power Grid Dancing Based on Inertial Technology, The host: Jinxian Yang, Henan Polytechnic University, under Grant 172102210289.

ABSTRACT To research the problems of the rolling bearing fault diagnosis under different noises and loads, a dual-input model based on a convolutional neural network (CNN) and long-short term memory (LSTM) neural network is proposed. The model uses both time domain and frequency domain features to achieve end-to-end fault diagnosis. One-dimensional convolutional and pooling layers are utilized to extract the spatial features and retain the sequence features of the data. In addition, an LSTM layer is employed to extract the sequence features. Finally, a dense layer is applied for fault classification. To enhance recognition accuracy under different noises and loads, three techniques are applied to the proposed model, including taking time–frequency domain signals as input, using the CNN–LSTM model, and adopting the mini-batch and batch normalization methods. The Case Western Reserve University and Drivetrain Diagnostics Simulator data sets are used to construct experiments under different conditions, including varying loads and different noises. The proposed model can achieve a high fault recognition rate under variable load and noise conditions as well as satisfactory anti-noise and load adaptability.

INDEX TERMS Time-frequency features, CNN, LSTM, bearing fault diagnosis, anti-noise and variable load adaptation.

NOMENCLATURE

CNN	**Convolutional neural network**
LSTM	**Long short term memory**
WConv	**Wide convolutional layer**
TF-	**Time domain and frequency domain**
CWRU	**Case Western Reserve University**
DDS	**Drive train Diagnostics Simulator**
OSDM	**Oversampling Data Enhancement**
BN	**Batch Normalization**
STFT	**the short-time Fourier transform**
1-D	**one-dimensional**
VMD	**Variational mode decomposition**

PARAMETERS

Keep_pro	***The parameters of dropout***
Lr	***The learning rate of neural networks***
Epochs	***The training iteration***
num_classes	***The number of the model's outputs***
train-test_rate	***The rate of training set and test set***
BatchNormal	***The parameters of Batch Normalization***
Output Shapes	***The shape of model's outputs***
Numbers	***The number of the convolutional kernel***
Activation function	***The activation function of neural networks***
Padding	***The parameters of pooling layer***
γ	***scaling parameters***
β	***translation parameters***

The associate editor coordinating the review of this manuscript and approving it for publication was Dong Wang.

I. INTRODUCTION

The fault monitoring of mechanical equipment is critical and determines whether they can operate safely. Rolling bearings are the core components of mechanical equipment and are susceptible to damage. According to statistics, 30% of mechanical equipment malfunction are caused by bearing failures [1]. Failure of rolling bearings can lead to mechanical system collapse, which can cause huge economic losses and endanger the safety of personnel in serious cases. In industrial applications, rolling bearings generally operate under different loads and noises. These interference factors further complicate fault diagnosis and require fault diagnosis methods to stably and efficiently identify faults under different loads and noises and ensure the safe operation of industrial equipment [2].

A vibration signal is commonly used for fault diagnosis of rolling bearings. Such a signal is a type of time-series data that changes periodically [3]. Traditional fault diagnosis techniques use signal processing methods, such as fast Fourier transform (FFT), variational mode decomposition (VMD), and wavelet decomposition, to identify faults by determining periodic impact components of signals [4]–[6]. However, when disturbed by noise, a signal will cause the periodic impact feature to be submerged in the noise. At present, research on bearing fault diagnosis with noise interference mainly focuses on signal feature extraction methods [7]. Such schemes are often used in conjunction with denoising techniques to extract features that reflect bearing conditions and are insensitive to noise and operating conditions. To achieve intelligent fault diagnosis, these methods must be used in conjunction with a machine learning algorithm [8]. Amar *et al.* [9] proposed a novel vibration spectrum imaging feature enhancement program for low signal-to-noise ratio (SNR) conditions and implemented fault recognition through ANN. Wang and Yan [10] used an SVM to classify the working state of a bearing according to the energy of each modal component after VMD decomposition

With the development of modern industries, the sampling point and sampling period of a fault monitoring system continue to increase as well as the amount of data collected. Therefore, the ability to process large amounts of data has become a necessity for modern fault diagnosis methods [11]. Although the above method can effectively solve the problem of bearing fault diagnosis for small samples, the following phenomena still exist for large samples and complex interferences. 1) A fault diagnosis method is divided into four isolated parts, namely, denoising, feature extraction, dimension reduction, and classification. Such an approach destroys the coupling relationship between the parts and causes partial loss of fault information. 2) Artificial feature extraction requires certain prior knowledge, exhibits poor generalization ability under variable loads and noises, and experiences difficulty processing extensive data.

In recent years, deep learning algorithms have been effectively used for image recognition, speech recognition, and fault diagnosis [12]–[14]. For fault diagnosis, deep learning algorithms tend to employ end-to-end fault diagnosis schemes that use only one model for denoising, feature extraction, dimensionality reduction, and classification, which is called the adaptive feature extraction fault diagnosis method [15]. In [16], a deep CNN model with strong domain and anti-noise adaptability was established, which is an end-to-end fault diagnosis technique. In [17], an unsupervised learning method based on sparse filtering and softmax was proposed to solve the problem of fault diagnosis against the background of big data.

Among deep learning algorithms, the CNN is widely used in image processing, voice recognition, video processing, and fault recognition [18]–[20]. In bearing fault diagnosis, the input of the CNN fault diagnosis model generally employs a two-dimensional (2D) time–frequency map [21], [22] or one-dimensional (1D) time-series data [23]. Li *et al.* [24]. proposed a bearing fault diagnosis method based on STFT and DCNN, which extracts time–frequency map features through STFT and uses DCNN to achieve end-to-end fault diagnosis. Zhang *et al.* [25], developed a WDCNN method with the first layer of a wide convolution kernel, which was used in conjunction with an adaptive batch normalization (BN) algorithm to enhance the domain and anti-noise adaptability of the model.

Although the above methods perform well in dealing with fault diagnosis under slight noise, they exhibit low recognition accuracy for strong noise conditions, mainly for the following reasons. (1) The models use only time-domain or frequency-domain data as input, thereby resulting in the insufficient extraction of data features. (2) The models can fully extract spatial features of data but cannot extract temporal features.

Long short-term memory (LSTM) is a special recurrent neural network (RNN) structure that can effectively process sequence data [26]. Yildirim [27], proposed a method for processing ECG signals using a wavelet sequence and a bidirectional deep LSTM network. An LSTM network can extract the temporal features of time-series data and enhance the generalization ability of a model through a gate structure [28]. However, in the case of a large amount of data, an LSTM network will encounter difficulty extracting nonlinear characteristics of data and exhibit slow convergence speed [29].

To address the above problems, a time–frequency dual-input model based on a CNN and LSTM network (TF-WConvLSTM) is proposed in this study. The main contributions of this research are as follows:

1. An end-to-end bearing fault diagnosis model is developed. The model uses a time–frequency dual-input structure and combines a CNN and LSTM network to fully extract the spatiotemporal characteristics of bearing vibration data.

2. The algorithm uses time-domain and frequency-domain data, thereby effectively expanding the feature extraction range and improving the generalization ability of the model.

3. The algorithm performs well in noisy environments and directly processes an original noise signal without a pre-denoising method.

4. The algorithm has a strong load adaptation ability and can achieve 90%+ accuracy under different workloads and mixed variable loads.

The load adaptability and anti-noise characteristics of the proposed FT-WConvLSTM method are evaluated by a constant load test, a variable load test, and a noise test using the Case Western Reserve University (CWRU) and Drivetrain Diagnostics Simulator (DDS) data sets. The recognition accuracy of the proposed method can reach 90%+ under different workloads, and accuracy can reach 86% in different noise interference environments.

II. CONVOLUTIONAL NEURAL NETWORK AND LONG-SHORT-TERM MEMORY RECURRENT NEURAL NETWORK

A CNN is a feedforward neural network that can efficiently extract the spatial features of a sample. The CNN structure includes a convolution layer, a pooling layer, a BN layer, and an activation layer. LSTM is a special type of RNN that is used to process and predict sequence data and can fully extract the temporal features of data. In this study, a 1D CNN is employed to extract the spatial features of time-domain signals and time-frequency maps, and LSTM is utilized to extract temporal features.

A. 1D CONVOLUTIONAL OPERATION

A convolutional operation can be divided into 1D and 2D convolutions according to the dimensions of the input data. 2D convolution is typically used to process images. It is a convolution operation along the XY dimensions of an image that uses a 2D filter to extract spatial features. However, it will cause temporal features to be lost when processing sequence data. By contrast, 1D convolutional employs only convolution operations along a single dimension of data. If this dimension is the timeline, then the temporal characteristics of the data are retained. For time-series data, 1D convolution employs multiple filters to move along the timeline to extract features. After the bias is added, an activation function is utilized to obtain the feature map of the retained time feature [30]. The specific formula is as follows:

$$\begin{aligned} Q_i^{l+1}(\tau) &= f\left(\sum_{i=1}^{F^l} \omega_{ij}^l(\tau) * Q_j^l(\tau) + b_i^l\right) \\ &= f\left(\sum_{i}^{F^l}\left[\sum_{t=1}^{T^l} \omega_{ij}^l(t)\, Q_i^l(\tau - t)\right]\right), \end{aligned} \tag{1}$$

where $Q_i^{l+1}(\tau)$ is the feature map i in layer l, f represents the nonlinear function, F is the number of feature maps, W is the convolutional kernel, and T is the size of the convolutional kernel.

B. 1D POOLING OPERATION

The pooling layer lies behind the convolutional layer and employs a down-sampling operation to reduce the dimension of the feature map. We use the max-pooling operation in this research. Suppose that the output feature maps of the convolution layer are $X \in \mathbb{R}^{M,N,D}$. The max-pooling layer uses the patch of w^*d to move along a single direction in the feature map with step S and divides each feature map X^d into multiple regions R_m^d $(1 < m < M)$. Meanwhile, the patches employ a pooling operation to output the maximum value of the neurons in each region [31].

$$P_m^d = \max_{i \in R_m^d} \{x_i\}, \tag{2}$$

where X represents the activation value of each neuron in region R.

C. BATCH NORMALIZATION

The BN method [32] is a layer-by-layer normalization method for reducing the internal covariance migration phenomenon. It can normalize any intermediate layer of a neural network and reduce sample differences between layers. Thus, this technique can reduce training time and avoid the effects of gradient disappearance and gradient explosion. The BN method is commonly set between convolution and activation layers. The specific formulas are as follows:

$$\mu_{\mathrm{B}} = \frac{1}{M}\sum_{m=1}^{M} x_m, \quad \sigma_{\mathrm{B}}^2 = \frac{1}{M}\sum_{m=1}^{M}(x_m - \mu_{\mathrm{B}})^2, \tag{3}$$

$$\hat{x}_m = \frac{x_m - \mu_{\mathrm{B}}}{\sqrt{\sigma_{\mathrm{B}}^2 + \varepsilon}}, \tag{4}$$

where x_m represents the value of x over a mini-batch, $\mathrm{M} = \{x_1, x_2, \cdots, x_m\}$, μ_B represents the mean value of the mini-batch samples, and σ_B^2 is the variance of the mini-batch sample. The BN method introduces scaling parameters γ and translation parameters β to enable BN results to be adaptively extended and translated. The specific formula is as follows:

$$y_m = \gamma \hat{x}_m + \beta \equiv BN_{\gamma\beta}(x_m). \tag{5}$$

D. LONG-SHORT-TERM MEMORY OPERATION

The internal structure of a hidden layer of RNN is relatively simple. For large data volumes, the RNN model will involve gradient disappearance or gradient explosion. To solve the above problems, LSTM [33] introduces input gate, forgetting gate, and output gate structures, thereby increasing the function of information long- and short-term selection memory. A gate is an operation that includes a sigmoid network and bitwise multiplication. A sigmoid network can output a value between 0 and 1, which represents whether the input value can go through a gate.

The aim of the forgetting gate is to make the LSTM network forget previously useless information.

$$f_t = \sigma\left(W\left[x_t, h_{t-1}, C_{t-1}\right] + b_f\right). \tag{6}$$

The function of the input gate is to determine current state C_t according to current input $[h_{t-1}, x_t, C_{t-1}]$ and $C_t C_{t-1}$.

$$\begin{aligned} i_t &= \sigma\left(W\left[x_t, h_{t-1}, C_{t-1}\right] + b_c\right) \\ C_t &= f_t \cdot C_{t-1} + i_1 \cdot \tanh\left(W\left[x_t, h_{t-1}, C_{t-1}\right] + b_c\right). \end{aligned} \tag{7}$$

The output gate determines output h_t at the current moment according to current state C_t, current input x_i, and output h_{t-1} at a previous time.

$$\begin{aligned} o_t &= \sigma\left(W\left[x_t, h_{t-1}, C_t\right] + b_o\right) \\ h_t &= \tanh\left(Ct\right) \cdot o_t, \end{aligned} \tag{8}$$

where $\sigma()$ is a logistic function whose outputs are in the scale of (0,1), x_t is the input at the current time, and h_{t-1} represents the state at a previous time.

E. STFT

STFT is defined as a fixed-width Fourier window function that moves along the time axis of the signal to intercept it. The signal is intercepted into several sub-signals of equal length, with each sub-signal approximately stationary. Then, the Fourier transform of the sub-signal is used to obtain the local spectrum set at each time t, thereby forming the 2D time–frequency diagram. The formula is as follows:

$$STFT = \int_{-\infty}^{+\infty} X\left(t\right) F\left(t-\tau\right) e^{-j\omega t} d\tau, \tag{9}$$

where $X(t)$ is the original signal, and $F(t-\tau)$ is the Fourier window centered at time τ. According to [22], the width of the STFT window will affect the time–frequency resolution R of the time–frequency image. Simultaneously, R determines the information contained in the time–frequency image and includes useful fault information and noise. Therefore, the reasonable selection of T and F can make fault features significant and reduce noise interference.

$$\begin{aligned} T &= \left[\frac{X_s - Xr}{X_w - X_r}\right], \\ F &= \begin{cases} \dfrac{N_x}{2} + 1, & N_x \text{ is even number} \\ \dfrac{N_x + 1}{2}, & N_x \text{ is odd number}, \end{cases} \qquad R = T \times F. \end{aligned} \tag{10}$$

F. SNR

SNR represents the ratio of signal to noise. The smaller the SNR, the greater the proportion of noise contained in the signal.

$$SNR = 10\log_{10}\left(\frac{e_{\text{signal}}}{e_{\text{noise}}}\right), \tag{11}$$

where e_{signal} represents the energy of the signal, and e_{noise} represents the energy of the noise.

III. PROPOSED INTELLIGENT DIAGNOSIS METHOD

A. TF-WConvLSTM NETWORK

The flow of the TF-WConvLSTM bearing fault diagnosis method is shown in Fig. 1. For fault diagnosis under complex interference, including variable loads and different noises, improving the recognition accuracy of the model is a key issue. Three methods address the above issue. First, a dual-input model is established using the time and frequency domain features of a bearing vibration signal to enhance feature extraction. Second, an interference training technique

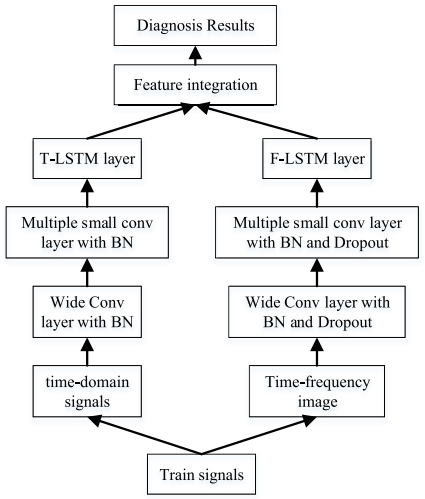

FIGURE 1. The overall framework of proposed TF-WConvLSTM.

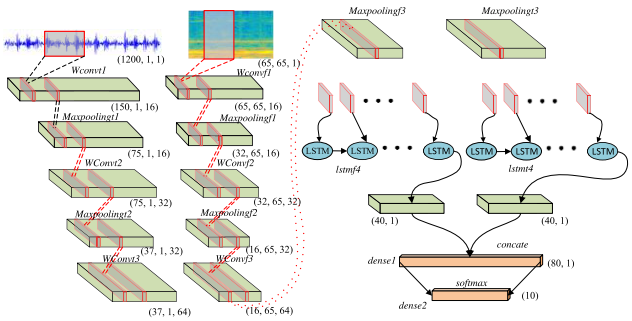

FIGURE 2. The structure of the model.

combining mini-batch and BN methods is applied, and BN is used between each convolutional and active layer. In addition, BN can be combined with the mini-batch training method to interfere with the training of the model, thereby improving its anti-interference ability. Third, a CNN–LSTM model is employed, which can effectively extract and utilize the spatiotemporal features of the input and enhance the generalization ability of the model using an LSTM gate structure. Fig. 2 shows the schematics of the model.

As shown in Fig. 2, the model is a dual-input network structure that includes a time–frequency domain. Specifically, the time-domain input is a time-series signal with a structure of [1200,1]. Meanwhile, the frequency-domain input is a time–frequency diagram with a time–frequency resolution of 65*65. The first layer of the time- and frequency-domain branches adopts a convolutional layer with a wide convolution kernel and can effectively increase the receptive field of the CNN, thereby reducing the influence of noise. After the convolution layer, a 1D pooling operation is employed to reduce the dimension of the feature map. Then, the second and third layers adopt stacked convolution and pooling operations with small convolution kernels, which increase the depth of the network model and enhance feature extraction. The fourth layer uses LSTM to extract the temporal features of the feature map. Finally, the outputs of the FLSTM and TLSTM layers are concatenated and classified by a dense

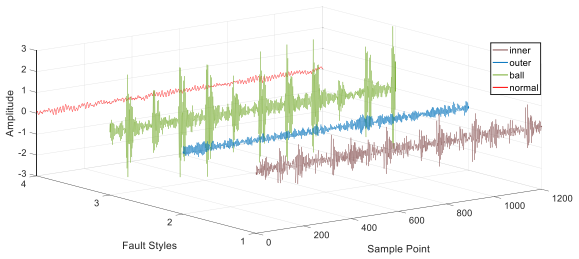

FIGURE 3. Visualization of time domain signal.

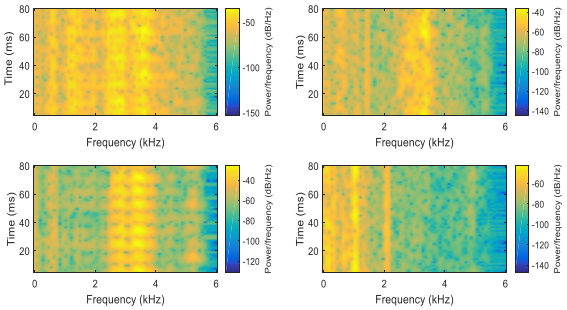

FIGURE 4. Visualization of frequency domain signal.

layer. The use of BN between the convolutional and active layers of the model serves to hasten the training process and reduce the effects of interference. Dropout is used between the convolutional and pooling layers to prevent overfitting.

B. USE OF TIME–FREQUENCY FEATURES TO IMPROVE THE MODEL

The bearing vibration signal is a time-series data that contains extensive information in time-domain and frequency-domain features. As shown in Figs. 3 and 4, periodic shock components can be clearly observed from the time domain of a signal. In the frequency domain, different types of fault signals are concentrated in various frequency bands. Therefore, time-domain and frequency-domain information are employed as input to construct a two-input fault diagnosis model, which can enhance feature extraction and anti-interference.

(1) Time-domain data use raw vibration data acquired by an accelerometer. As shown in Fig. 3, the bearing vibration signal changes periodically. One rotation of a bearing is one cycle. Therefore, the number of signal points collected by a cycle is T, which can be obtained according to the sampling frequency of the sensor and the rotational speed of the motor, that is, $T = Fs \times (60/S)$. To avoid data skew, three consecutive periods are utilized as a sample to build the time-domain data set.

(2) Frequency-domain data involve time–frequency maps transformed by STFT. According to the frequency-domain characteristics of the bearing vibration signal, the length of the sample can be determined by a one-fold fault frequency, and the width of the STFT window is determined according to a two-fold fault frequency.

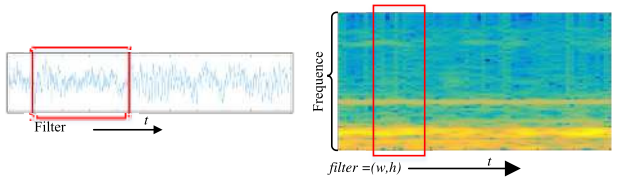

FIGURE 5. The process of 1D convolution.

The 1D convolution is applied to extract spatial features of time-domain signals and time-frequency figures. And it can retain the sequence features of the data. The specific process is shown in Fig.5, the convolution kernel of the 1D convolutional layer only moves along the t-axis of the time-domain signal and time-frequency figures to extract their spatial features.

C. CNN AND LSTM

A CNN can adaptively complete feature extraction and data dimension reduction through convolution and pooling and exhibits better generalization ability than traditional feature extraction methods. In bearing fault diagnosis, a bearing signal is generally converted into a time-domain map or a time–frequency map and then processed by a 2D CNN. However, a 2D convolutional operation can only extract the spatial features but ignores the temporal features of a signal, thereby causing the above model to perform poorly under complex interferences. The characteristics of the bearing vibration signal are considered, and a 1D convolutional operation is used to extract features along the time-axis convolution of the signal, thereby ensuring feature extraction while retaining temporal features.

LSTM has unique advantages when dealing with time-series data. However, it exhibits poor feature extraction performance for large-volume samples and requires substantial computation time to achieve satisfactory results. Therefore, this study uses a CNN–LSTM structure. First, convolutional and pooling layers are used to extract the spatial characteristics and reduce the dimensions of data. Then, an LSTM network is utilized to further extract the time characteristics of the data.

The LSTM network adopts a gate control unit structure, as shown in Formulas (6) and (7). Among the gates, the input and forgetting gates can effectively control the input of sample information and the forgetting of information. The input gate determines features i_t to be retained according to $[h_{t-1}, x_t, C_{t-1}]$, while the forgetting gate determines the features that must be forgotten in C_t. As shown in Formula (7), input gate i_t and forgetting gate f_t can be regarded as filters. Meanwhile, input $\widehat{C}_t$ and previous state C_{t-1} are filtered to screen out effective fault information to contribute to denoising. Therefore, the CNN–LSTM structure can improve the anti-noise performance of the model.

CNN network shows excellent spatial feature extraction ability, while LSTM network shows excellent sequence

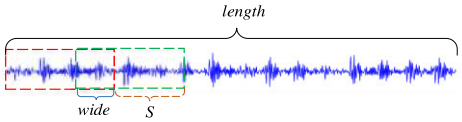

FIGURE 6. Data enhancement.

feature extraction ability. The bearing vibration signal is a kind of time series data with periodic changes. This requires that the bearing fault diagnosis model should contain both excellent temporal and spatial feature extraction capabilities. Therefore, 1D convolutional neural network and LSTM network are both used in bearing fault diagnosis model. It extracts the spatial features through 1d convolutional layer, and retains the temporal features. Then, it extracts the temporal features through LSTM network. Therefore, this model can extract the features of bearing vibration data more completely.

D. OVERSAMPLING DATA ENHANCEMENT (OSDM)

Deep learning algorithms often experience overfitting owing to few samples during training. To avoid this phenomenon, data set enhancement techniques are often used. A bearing vibration signal is a 1D time-series data. This study uses an oversampling method to enhance the data set. A sampling window, with a width of w and a repetition rate of *wide* as well as a step size of S, is resampled along the time axis. The specific process is shown in Fig. 6.

IV. VALIDATION OF PROPOSED TF-WConvLSTM

A. DATA DESCRIPTION

1) CWRU BEARING FAULT DIAGNOSIS DATA SET

The test data are derived from the simulated bearing failure test rig of the CWRU [34]. The test rig includes a 2 HP motor, an encoder, and a dynamometer. The test bearing is a SKF6205 motor bearing. A single-point fault with multiple pitting diameters is introduced by the EDM technology, and the fault diagnosis data of the drive section and fan section are collected by an acceleration sensor. The experimental data are shown in Table 1. The data of the drive section, with a load of 0 HP, 1 HP, 2 HP, and 3 HP and a sampling frequency of 12 kHz, are selected. To verify the performance of the proposed model, data of 10 different fault levels, including inner race, outer race, ball, and normal, are selected for the test. As shown in Table 1, each fault type collects 800 samples, and 1200 signal points constitute a group of samples.

2) DDS BEARING FAULT DIAGNOSIS DATA SET

To verify the performance of the model in actual conditions, an experiment is performed using the DDS test bench shown in Fig. 7.

The DDS test rig is a complete powertrain system that includes a variable speed drive motor, a primary gearbox, a secondary parallel gearbox, a torque sensor, an encoder, a programmable magnetic brake, and a sensor acquisition

FIGURE 7. Drive train Diagnostics Simulator test bench.

system. This test bench can study the vibration characteristics, noise characteristics, and diagnostic techniques of the gearbox. In this study, the rolling bearings of normal, inner-ring, outer-ring, and rolling-element faults are tested by the DDS.

B. EXPERIMENTAL SETUP

1) CONTRAST MODEL

To verify the advantages of the proposed TF–WConvLSTM model under complex interference conditions compared with traditional machine learning, feature extraction deep learning, and common single-input end-to-end fault diagnosis methods, it is compared against the SVM, CNN, WDCNN, TF-CNN, WConvLSTM, and VMD-LSTM models. WDCNN is a deep CNN method in which the first layer is a wide convolution kernel and the second layer is the proposed common convolution kernel. VMD-LSTM is a bearing fault diagnosis method proposed by our team using VMD-TEO window extraction features and deep bidirectional LSTM network modeling. The parameters of the model are shown in Table 2.

2) PARAMETERS OF PROPOSED TF–WConvLSTM MODEL

The TF–WConvLSTM network used in the experiment is a dual-input model. The time-domain and frequency-domain branches of the model consist of a wide convolution layer, two 1D convolution pooling layers, and an LSTM layer. Classification is achieved by one concatenating layer and two dense layers. The first convolution kernel has a width of 64, and the others have a width of 2. The number of LSTM cells is 100, and the number of Dense1 cells is likewise 100. A total of 10 Dense2 cells is present according to the data set sample type. The pooling layer adopts a maximum pooling layer, and BN operations are used between the convolutional and active layers. The convolutional, LSTM, and Dense2 layers sequentially use the *Relu*, *tanh*, and *softmax* functions as activation functions, which are trained by the Adam algorithm. The specific parameters are shown in Table 2. The test platform is a Win10 + matlab2016b and a Ubuntu16.04 + python3.5 + tensorflow11, and the CPU is an Intel (R) CORE(TM) i7-7500u.

TABLE 1. CWRU bearing fault data set.

Label	Failure level	Fault Diameter (mm)	Load A (HP)	Load B (HP)	Load C (HP)	Load D (HP)
1	Normal	0	0	1	2	3
2	Minor Inner Race Fault	0.18	0	1	2	3
3	Medium Inner Race Fault	0.36	0	1	2	3
4	Serious Inner Race Fault	0.54	0	1	2	3
5	Minor Outer Race Fault	0.18	0	1	2	3
6	Medium Outer Race Fault	0.36	0	1	2	3
7	Serious Outer Race Fault	0.54	0	1	2	3
8	Minor Ball Fault	0.18	0	1	2	3
9	Medium Ball Fault	0.36	0	1	2	3
10	Serious Ball Fault	0.54	0	1	2	3

TABLE 2. Parameters of the model.

Layer	CNN	VMD-LSTM	WCNN	TF-WCNN	WConvLSTM
1	Conv1/(2,2)/16	VMD	Conv1/(65,65)/16	Convt1/(65,65)/16	Conv1/(65,65)/16
2	Pool12/(2,2)	LSTM1/100	Pool12/(2,2)	Pool1t1/(2,2)	Pool12/(2,2)
3	Conv1/(2,2)/32	LSTM2/100	Conv1/(2,2)/32	Conv1f1/(2,2)/32	Conv2/(2,2)/32
4	Pool12/(2,2)	LSTM3/100	Pool12/(2,2)	Pool1f1/(2,2)	Pool12/(2,2)
4	Conv1/(2,2)/64	-	Conv1/(2,2)/64	Conv1t2/(2,2)/64	-
5	Pool12/(2,2)	-	Pool12/(2,2)	Pool1t2/(2,2)	-
6	-	-	-	Conv1f2/(2,2)/64	LSTM1/(100)
7	-	-	-	Pool1f2/(2,2)	LSTM2/(40)
8	Dense/100	Dense/100	Dense/100	Dense/100	Dense/100
9	Dense/10	Dense/10	Dense/10	Dense/10	Dense/10

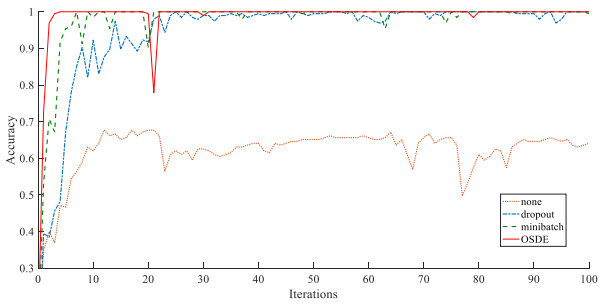

FIGURE 8. The diagram of the model training process.

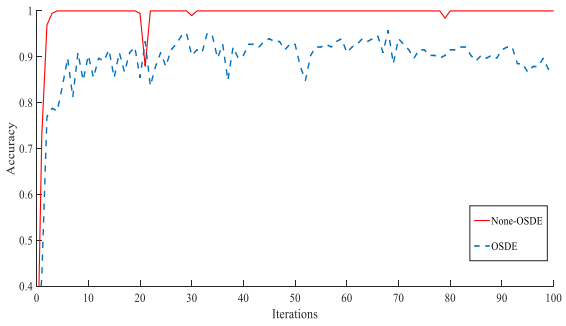

FIGURE 9. The effect of OSDE.

Fig. 8 shows the accuracy graph of training under four different structures of TF-WConvLSTM models. The accuracy of the original model fluctuates greatly during the convergence process, which can only reach 65%. Meanwhile the model has a high recognition accuracy of the training set, and the recognition accuracy of the verification set and the test set is low. This is an overfitting phenomenon caused by less training data. Therefore, to avoid this phenomenon, Dropout, BN, mini-batch methods and OSDE data enhancement methods are added to the original model. The results show that the accuracy of the model after adding dropout and BN is obviously improved, reaching 98%. However, the convergence rate of the model is slow, and the convergence process is accompanied by small fluctuations; after adopting the mini-batch training method, the convergence rate has been improved. Although the above method improves the performance of the model through structural optimization, there are still some fluctuations. Therefore, the OSDE method is used to enhance the training set. After the OSDE, the training set sample is increased to 7140 groups. The model is stable after five iterations, the accuracy is 99%.

As shown in Fig. 10, mode1 is a structure that uses dropout in both time-domain branch and frequency-domain branch. Its convergence process contains a lot of fluctuation, and the accuracy of fault identification can only reach 82%. Model2 is the structure proposed in this paper, which only uses dropout in the frequency domain branch. The convergence of the model can be achieved after 10 iterations,

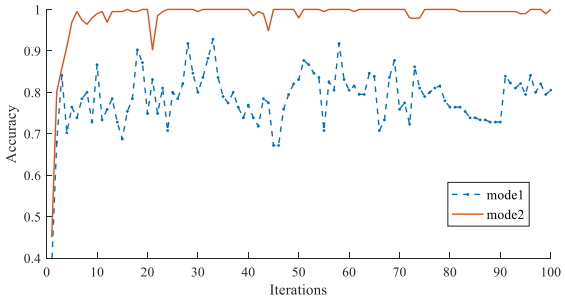

FIGURE 10. Model comparison results.

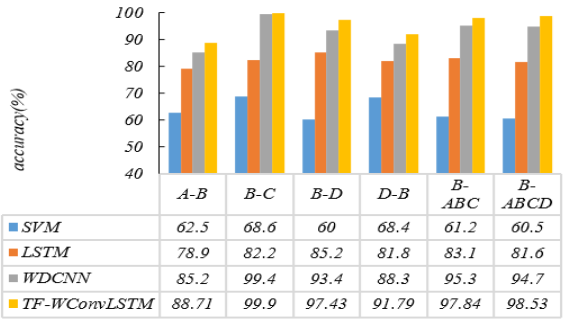

FIGURE 11. Comparison of load adaptability test results.

TABLE 5. Alter load data set.

Group	Load(HP)	Train	Test
A-B	0~1	train A	test B
B-D/D-B	1~2/2~1	train B/train D	test D/train B
B-ABC	1~012	train B	test AB
B-ABCD	1~0123	train B	test ABCD

which fault identification accuracy can reach 99%. Therefore, the structure of model2 is more suitable for bearing fault diagnosis.

Finally, an optimal TF-WConvLSTM model is determined, as shown in Table 3. The remaining hyperparameters are as follows: *Keep_pro* = *0.85, Lr* = *0.06, Epochs* = *100, num_classes* = *10, train-test_rate* = *[0.7,0.15,0.15], batch_size* = *32, BatchNorm* = *True, length* = *1200,* and *OSDE* = *7500.*

Table 4 compares the computational time of the different models. The TF-WConvLSTM model demonstrates relatively fast calculation speed. Compared with CNN networks, the proposed model shows a faster computational speed owing to the structure of the wide convolution kernel. The proposed model's feature extraction and feature dimensionality reduction abilities are better than the VMD-LSTM structure, thereby showing excellent operation speed. This is because the convolution and pooling layers are used before the LSTM layer. Compared with the WConvLSTM model, the proposed model shows higher recognition accuracy, but the calculation time is higher than before. This is due to both time-domain and frequency-domain data are extracted by the proposed model.

3) ALTER LOAD ENVIRONMENT

To verify the model under different loads, an alter load experiment is designed using the CWRU data set. The test data are shown in Table 5, where *A*, *B*, *C*, and *D* are constant load data sets with loads of 0 HP, 1 HP, 2 HP, and 3 HP, respectively. *AB* represents a data set under a two-load mixing condition, *ABC* is a data set under a three-load mix, and *ABCD* is a four-load mixed state data set.

The variable load experimental data set is constructed from the above data. Data set *A-B* simulates the load from 0 HP to 1 HP, which is constructed using the training set of A and the test set of B. Data set *B-D*\ *D-B* simulates the state at

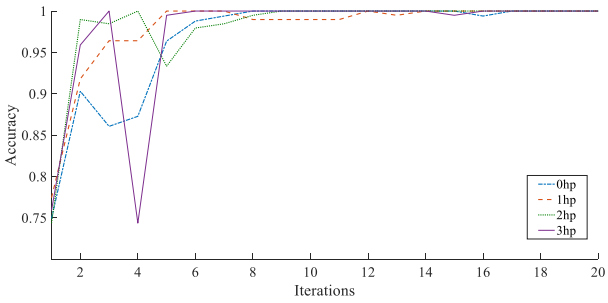

FIGURE 12. Model convergence process under different load.

which the load transitions from 1 HP to 2 HP or from 2 HP to 1 HP. Data sets *B-ABC* and *B-ABCD* simulate the state of the bearing when it is running from a constant to a variable load.

As shown in Fig. 11, the recognition accuracy of SVM can only reach between 60% and 70%. Thus, the LSTM model performs better, with the highest accuracy of 85%. The WDCNN model exhibits better load adaptability than the above model, with a recognition accuracy of 85%. By contrast, the proposed model performs well under a single variable load and demonstrates remarkable recognition accuracy for complex variable load conditions. The model's recognition accuracy is nearly over 90%. In conclusion, the TF-WConvLSTM bearing fault diagnosis model shows considerable load adaptability.

The convergence of the model under four different loads is shown in Fig. 12 below.

To verify the effect of the proposed model, the ROC curve method is used to evaluate the model. This study takes the B-D case as an example, and the specific results are shown in Fig. 13.

To visually present the feature representation of the TF-WConvLSTM model, t-SNE is employed to visualize the feature representation of the Dense1 layer. Fig. 14 shows the visualization diagram of the output by the Dense1 layer when SNR is 2. Ten types of bearing signal data achieve effective separation and are concentrated in a specific area. Thus, the model has learned effective features and realized data classification.

TABLE 3. Detailed overview of the proposed TF-WConvLSTM model.

Layers	Types	Output Shapes	Kernel Size/Stride	Kernel Numbers	Dropout	Activation function	Pad-ding	BN
1-1	Input_T	(1200,1,1)	-	-	-	-	-	N
1-2	Input_F	(65,65,1)	-	-	-	-	-	N
2-1	1D Time Wide Convolution	(150,1,16)	64*1/8	16	-	RELU	2	Y
2-2	1D Frequency Wide Convolu-tion	(65,65,16)	65*65/1	16	0.85	RELU	2	Y
3-1	T-1D-Conv1	75*1*32	2*1/1	32	-	RELU	2	Y
3-2	F-1D-Conv1	65*32*32	2*1/1	32	0.85	RELU	2	Y
4-1	T-1D-Conv2	37*1*64	2*1/1	64	-	RELU	1	Y
4-2	F-1D-Conv2	65*16*64	2*1/1	64	0.85	RELU	1	Y
5-1	T-LSTM	40*1	40	-	-	Tanh	-	N
5-2	F-LSTM	40*1	40	-		Tanh	-	N
6	Concate	80*1	-	-	-	-	-	N
7	Dense1	100*1	100	-	0.9	-	-	N
8	Dense2	10	10	-	-	Softmax	-	N

TABLE 4. The computational time.

Mode	Data type	Pretreatment time	Train-time(s)	Test-time(s)	Accuracy(%)
CNN	STFT	205.56	90.59	0.06682	87.2
VMD-LSTM	TIME	308.21	157.21	0.1176	89.6
WConvLSTM	STFT	205.56	58.23	0.0339	98.2
TF-WConvLSTM	TIME-STFT	207.16	70.12	0.0618	99.48

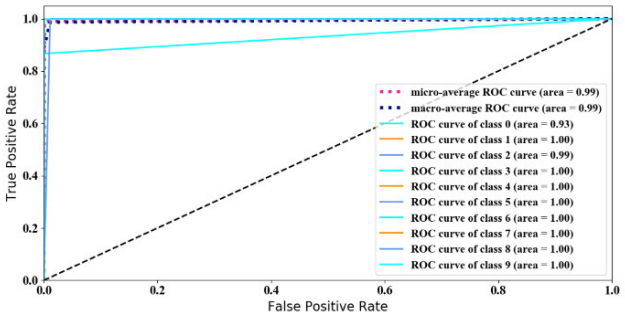

FIGURE 13. The ROC curve under load B-D.

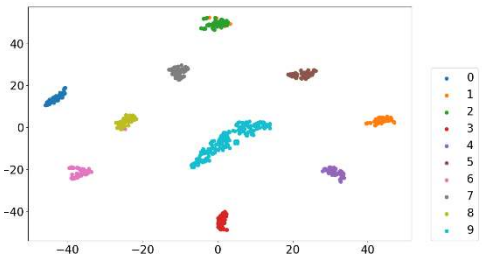

FIGURE 14. T-SNE visualization of Dense1 layer.

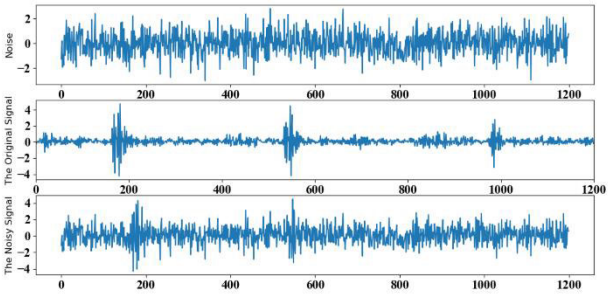

FIGURE 15. Bearing vibration signal diagram with SNR of −4.

4) NOISE TESTING

To verify the anti-noise performance of the model, noise interference experiments are designed using the CWRU data set. Gaussian white noise with different intensities is added to the original signal to simulate the bearing vibration signal under different SNR conditions.

As shown in Fig. 15, after noise with a SNR of −4 is added to the original signal, the periodic impact component of the signal weakens.

The experiment uses noise-containing signals with a SNR of $-4db{\sim}10db$ to verify the anti-noise performance of the four models. The results are shown in Fig. 16.

Although the fault recognition accuracy of the four models reaches 90% under low-noise signals, accuracy decreases under strong-noise interference. Specifically, the recognition accuracy of the WDCNN methods reaches only 70% when SNR is −4db. Moreover, the recognition accuracy of the VMD-LSTM model performs well but reaches only 75% when SNR is −4db. In contrast, the fault recognition accuracy of WConvLSTM model with single-input is higher than the above model under strong noise. However, its accuracy can only reach 79.5 in the case of SNR = −4. By contrast, the recognition accuracy of the TF-WConvLSTM model is

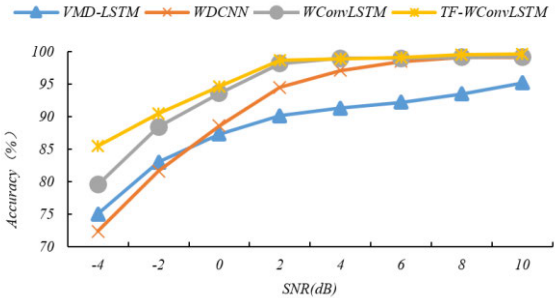

FIGURE 16. Comparison of accuracy testing on signals with different SNR values.

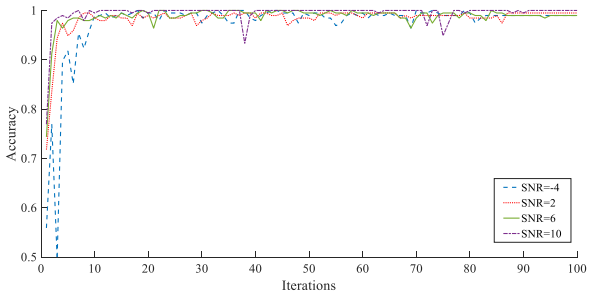

FIGURE 17. The convergence process under different noise.

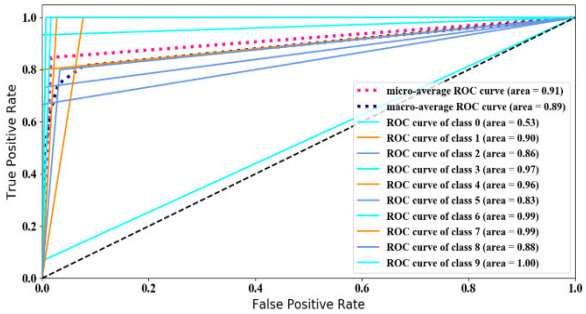

FIGURE 18. The ROC curve under SNR = −4.

consistent at 86%+ when different noise intensity signals are processed. However, under the condition of SNR < 0, the dual-input model shows better feature extraction effect, which can extract both time-domain and frequency-domain features of the model. Therefore, the dual-input model is more suitable for bearing fault diagnosis under strong noise.

The convergence of the model when SNR is −4, 2, 6, and 10 are shown in Fig. 17.

Noise can interfere with the convergence process of the model. The model convergence process shows large fluctuations when SNR is −4. In general, the model can reach convergence after 90 trainings.

The ROC curves of the model are established when SNR is −4 and 2, as shown in Figs. 18 and 19.

In industrial environments, complex disturbance conditions consisting of variable loads and strong noise often occur. Therefore, a variable load noise dataset was constructed by the CWRU data set. The *-4db ∼10DB* Gaussian white noise is

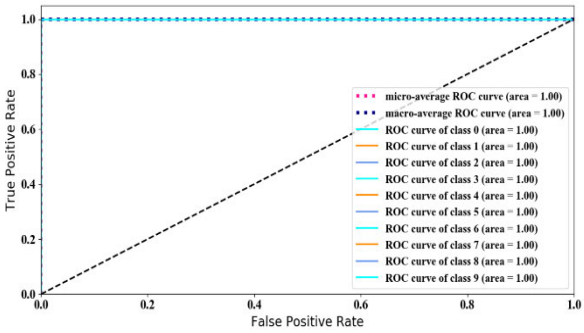

FIGURE 19. The ROC curve under SNR = 2.

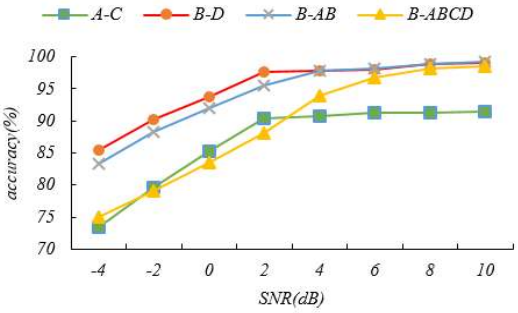

FIGURE 20. Comparison of accuracy testing on signals with different loads and noise.

added to the *A-C B-D B-AB* and *B-ABCD* four sets of variable load data sets to simulate complex conditions.

In industrial environments, complex disturbance conditions consisting of variable loads and strong noise often occur. Therefore, a variable load noise data set is constructed using the CWRU data set. *−4db ∼10DB* Gaussian white noise is added to the *A-C, B-D, B-AB,* and *B-ABCD* sets of the variable load data set to simulate complex conditions.

The results are shown in Fig. 20. The accuracy of the model exceeds 95% under slight noise. With a decrease in SNR, accuracy likewise decreases. For *B-D* and *B-AB*, accuracy is above 98% when $SNR > 4$ and can reach 85%+ under strong noise. Data set *B-ABCD* simulates the variable load state with noise. When noise is low, accuracy can reach 99%. However, as noise increases, accuracy decreases. When $SNR = -4$, accuracy will reach only 75%. In general, the model exhibits remarkable performance in bearing fault diagnosis under variable loads and noises. However, its recognition accuracy for variable loads with strong noise should be improved.

C. DDS DATA SET TEST

To verify the recognition accuracy of the model under actual working conditions, including variable speeds and noise interference, an experiment is designed using the DDS. As shown in Fig. 21, owing to the influence of noise, the periodic fault impact component cannot be clearly observed from the signal.

TABLE 6. DDS data set.

Fault style	Label	ENC Train set A	Test set A	ENC Train set B	Test set A	ENC Train set AB	ENC Test set AB
Inner Race Fault	1	400	24	300	24	700	48
Outer Race Fault	2	410	24	320	24	730	48
Ball Fault	3	390	24	330	24	720	48
Normal	0	480	24	400	24	880	48

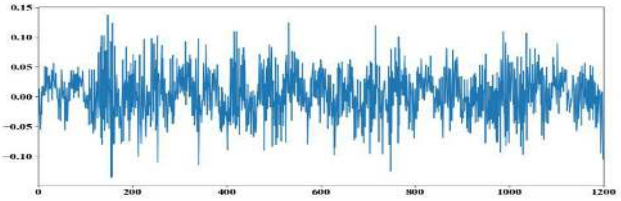

FIGURE 21. The original signal diagram of DDS.

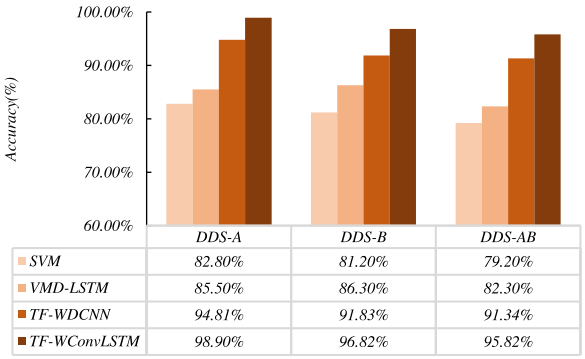

FIGURE 22. The test results of the DDS data set.

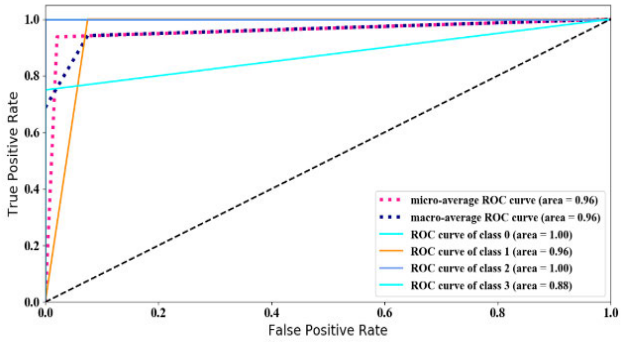

FIGURE 23. The ROC curve under DDS data set.

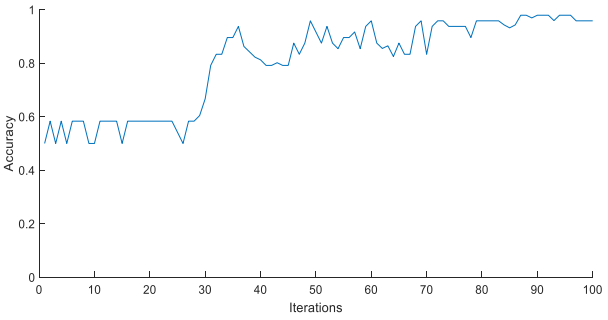

FIGURE 24. The convergence process under DDS data set.

The vibration data of a bearing with different speeds are used in the experiment. In Table 6, *A* represents a constant speed data set with a motor frequency of 35 Hz, that is, 1120*r/min*; *B* is a data set with a frequency of 45 Hz, that is, 1350r/min; and *AB* represents the variable speed data set, which consists of constant speed data sets *A* and *B* to verify the variable speed adaptability of the model.

Experimental results are shown in Fig. 22. For the *DDS-A/B* constant speed data set and DDS-AB variable speed data set, the average accuracy of SVM can reach only 82.8% while that of LSTM is under 85%. WDCNN performs well, thereby achieving approximately 90% accuracy. Conversely, the proposed TF-WConvLSTM model is much more accurate than the above method, and its accuracy is 98.0%, on average. Thus, the TF-WConvLSTM model shows better anti-noise and variable load adaptability than traditional methods. By contrast, the structure of TF-WConvLSTM model is more suitable for bearing fault diagnosis under noise and load interference

For the DDS data, the convergence process and ROC curves of the model are shown in Figs. 23 and 24.

D. DATA VISUALIZATION

To study the structure of the model as well as the learning process comprehensively, visualization technology is used to explore the key content of the structure, and the hierarchical information of deep learning is clearly expressed to analyze the performance of the TF-WConvLSTM structure. This study uses T-SNE to output the characteristics of each layer and performs dimensionality reduction visualization. As shown in Fig. 25, most of the features of the input data are effectively separated and aggregated after feature extraction by three convolutional layers, thereby indicating that the convolutional layers can effectively extract fault features. However, several features remain unseparated in the convf output; thus, the time features of the feature map are further extracted by LSTM. Feature extraction exhibits improved results after the LSTM layer. Furthermore, dense output visualization shows that the model can learn effective features and effectively distinguish fault categories.

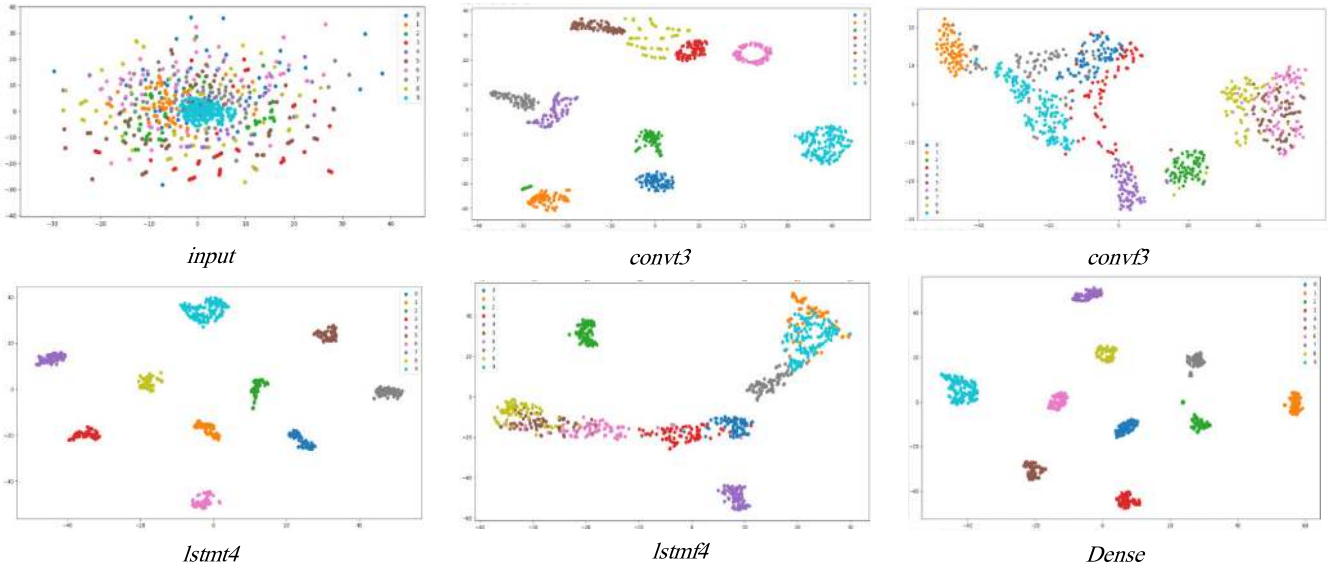

FIGURE 25. T-SNE visualization of model.

V. CONCLUSION

To address the problem of bearing fault diagnosis under variable loads and different noise interferences, an adaptable TF-WConvLSTM model is proposed. The proposed model adopts a time–frequency dual-input structure to enhance the feature extraction effect. The spatiotemporal characteristics of the model can be effectively extracted using the CNN–LSTM structure. In addition, through the LSTM gate structure, the temporal features of the model can be fully utilized, and the noise immunity of the model can be improved. With the help of data enhancement, the proposed model can achieve 90%+ accuracy under variable loads. Moreover, accuracy can reach 86%+ under noise and variable load environments. In the fourth section of the paper, an optimal model is determined experimentally, and the TF-WConvLSTM model is proven to be robust. In addition, the internal mechanism of the model is studied via network visualization. Compared with commonly used denoising preprocessing algorithms and domain adaptive algorithms that require additional information, our algorithms do not require any auxiliary algorithms.

Although the TF-WConvLSTM model exhibits remarkable anti-interference capability compared with other methods, recognition accuracy under strong noise should be improved. Therefore, the model's anti-noise and model adaptability methods must be studied further.

REFERENCES

[1] R. Abdelkader, A. Kaddour, A. Bendiabdellah, and Z. Derouiche, "Rolling bearing fault diagnosis based on an improved denoising method using the complete ensemble empirical mode decomposition and the optimized thresholding operation," *IEEE Sensors J.*, vol. 18, no. 17, pp. 7166–7172, Sep. 2018.

[2] D. Zhao, J. Li, W. Cheng, and Z. He, "Generalized demodulation transform for bearing fault diagnosis under nonstationary conditions and gear noise interferences," *Chin. J. Mech. Eng.*, vol. 32, no. 1, pp. 79–89, Dec. 2019.

[3] M. Zeng, W. Zhang, and Z. Chen, "Group-based K-SVD denoising for bearing fault diagnosis," *IEEE Sensors J.*, vol. 19, no. 15, pp. 6335–6343, Aug. 2019.

[4] X. T. Xie and S. B. Li, "Fault diagnosis of rolling bearing based on FFT and CS-SVM," *Modular Mach. Tool Autom. Manuf. Technique*, no. 4, pp. 90–94, Apr. 2019.

[5] Y. Liu, X. Wu, and T. Liu, "Feature extraction for rolling bearing incipient faults based on adaptive MOMEDA and VMD," *J. Vibrat. Shock.*, vol. 38, no. 23, pp. 219–229, 2019.

[6] Y. T. Ai and W. B. Tian, "Frequency band optimization of Morlet complex wavelet and its application in fault diagnosis of inter-shaft bearing," *J. Aerosp. Power.*, vol. 35, no. 1, pp. 153–161, Jan. 2020.

[7] S. Liu, Y. Sun, and L. Zhang, "A novel fault diagnosis method based on noise-assisted MEMD and functional neural fuzzy network for rolling element bearings," *IEEE Access*, vol. 6, pp. 27048–27068, 2018.

[8] X. L. Zhang, Q. Zhang, and X. R. Qin, "Fault diagnosis method for rolling bearing based on ITD-morphological filter and Teager energy spectrum," *Chin. J. Sci. Instrum.*, vol. 37, no. 4, pp. 788–795, Apr. 2016.

[9] M. Amar, I. Gondal, and C. Wilson, "Vibration spectrum imaging: A novel bearing fault classification approach," *IEEE Trans. Ind. Electron.*, vol. 62, no. 1, pp. 494–502, Jan. 2015.

[10] X. Wang and W. Y. Yan, "Fault diagnosis of roller bearings based on the variational mode decomposition and SVM," *J. Vibrat. shock.*, vol. 36, no. 18, pp. 252–256, Apr. 2017.

[11] J. Li, Y. B. Liu, and Y. H. Yu, "Application of convolutional neural network and kurtosis in fault dia gnosis of rolling bearing," *J. Aerosp. Power.*, vol. 34, no. 11, pp. 2423–2431, Nov. 2019.

[12] J. Yao, J. Wang, I. W. Tsang, Y. Zhang, J. Sun, C. Zhang, and R. Zhang, "Deep learning from noisy image labels with quality embedding," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 1909–1922, Apr. 2019.

[13] T. Afouras, J. S. Chung, A. Senior, O. Vinyals, and A. Zisserman, "Deep audio-visual speech recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Dec. 21, 2018, doi: 10.1109/TPAMI.2018.2889052.

[14] H. Ge, Z. Yan, W. Yu, and L. Sun, "An attention mechanism based convolutional LSTM network for video action recognition," *Multimedia Tools Appl.*, vol. 78, no. 14, pp. 20533–20556, Jul. 2019.

[15] W. Zhang, F. Zhang, W. Chen, Y. Jiang, and D. Song, "Fault state recognition of rolling bearing based fully convolutional network," *Comput. Sci. Eng.*, vol. 21, no. 5, pp. 55–63, Sep. 2019.

[16] W. Zhang, C. Li, G. Peng, Y. Chen, and Z. Zhang, "A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load," *Mech. Syst. Signal Process.*, vol. 100, pp. 439–453, Feb. 2018.

[17] Y. Lei, F. Jia, J. Lin, S. Xing, and S. X. Ding, "An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data," *IEEE Trans. Ind. Electron.*, vol. 63, no. 5, pp. 3137–3147, May 2016.
[18] H.-J. Lee and S.-G. Lee, "Arousal-valence recognition using CNN with STFT feature-combined image," *Electron. Lett.*, vol. 54, no. 3, pp. 134–136, Feb. 2018.
[19] J. Zheng, X. Cao, B. Zhang, X. Zhen, and X. Su, "Deep ensemble machine for video classification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 2, pp. 553–565, Feb. 2019.
[20] W. Sun, R. Zhao, R. Yan, S. Shao, and X. Chen, "Convolutional discriminative feature learning for induction motor fault diagnosis," *IEEE Trans. Ind. Informat.*, vol. 13, no. 3, pp. 1350–1359, Jun. 2017.
[21] Z. Zhu, G. Peng, Y. Chen, and H. Gao, "A convolutional neural network based on a capsule network with strong generalization for bearing fault diagnosis," *Neurocomputing*, vol. 323, pp. 62–75, Jan. 2019.
[22] M. Xia, T. Li, L. Xu, L. Liu, and C. W. de Silva, "Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks," *IEEE/ASME Trans. Mechatronics*, vol. 23, no. 1, pp. 101–110, Feb. 2018.
[23] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, "Real-time motor fault detection by 1-D convolutional neural networks," *IEEE Trans. Ind. Electron.*, vol. 63, no. 11, pp. 7067–7075, Nov. 2016.
[24] H. Li, Q. Zhang, X. Qin, and Y. Sun, "Fault diagnosis method for rolling bearings based on short-time Fourier transform and convolution neural network," *J. Vib. Shock*, vol. 37, no. 19, pp. 124–131, Sep. 2018.
[25] 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, Feb. 2017.
[26] W. Li, W. Nie, and Y. Su, "Human action recognition based on selected spatio-temporal features via bidirectional LSTM," *IEEE Access*, vol. 6, pp. 44211–44220, 2018.
[27] Ö. Yildirim, "A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification," *Comput. Biol. Med.*, vol. 96, pp. 189–202, May 2018.
[28] Z. Zhuang, H. Lv, J. Xu, Z. Huang, and W. Qin, "A deep learning method for bearing fault diagnosis through stacked residual dilated convolutions," *Appl. Sci.*, vol. 9, no. 9, p. 1823, 2019.
[29] Z. Li, J. Li, Y. Wang, and K. Wang, "A deep learning approach for anomaly detection based on SAE and LSTM in mechanical equipment," *Int. J. Adv. Manuf. Technol.*, vol. 103, nos. 1–4, pp. 499–510, Jul. 2019.
[30] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Proc. Adv. Neural Inf. Process. systems.*, vol. 25, no. 2, Jan. 2012, pp. 1097–1105.
[31] S. Zhai, H. Wu, A. Kumar, Y. Cheng, Y. Lu, Z. Zhang, and R. Feris, "S3Pool: Pooling with stochastic spatial sampling," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 4970–4978.
[32] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Mach. Learn.*, vol. 37. Jul. 2015, pp. 448–456.
[33] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Dec. 1997.
[34] K. A. Loparo. *Bearings Vibration Data Set Case Western Reserve University*. Accessed: Feb. 1, 2005. [Online]. Available: http://www.eecs.cwru.edu/laboratory/bearing/download.htm

MEIYING QIAO was born in 1976. She received the B.Sc. degree from the Taiyuan University of Technology, China, in 2000, the M.Sc. degree from the Jiaozuo Institute of Technology, China, in 2003, and the Ph.D. degree from the China University of Mining and Technology, China, in 2012. She is currently an Associate Professor and a Master Supervisor with Henan Polytechnic University, China. Her main research interests include machine learning, time series prediction, and fault diagnosis.

SHUHAO YAN was born in 1994. He received the B.Sc. degree from the Hebei University of Technology City College, China, in 2017. He is currently pursuing the master's degree with Henan Polytechnic University. His research interests include machine learning and fault diagnosis.

XIAXIA TANG was born in 1994. She received the B.Sc. degree from the Luoyang Institute of Technology, China, in 2018. She is currently pursuing the master's degree with Henan Polytechnic University. Her research interests include machine learning and fault diagnosis.

CHENGKUAN XU was born in 1995. He received the B.Sc. degree from Henan Polytechnic University, China, in 2018, where he is currently pursuing the master's degree. His research interests include inertial navigation and sensor fusion.