IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 1, 2020, accepted July 19, 2020, date of publication July 27, 2020, date of current version August 6, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3012053

Rolling Bearing Fault Diagnosis Based on Convolutional Neural Network and Support Vector Machine

LAOHU YUAN[1], DONGSHAN LIAN[1], XUE KANG[1], YUANQIANG CHEN[1], AND KEJIA ZHAI[1]

College of Aerospace Engineering, Shenyang Aerospace University, Shenyang 110136, China

Corresponding author: Laohu Yuan (ylhhit@126.com)

This work was supported in part by the National Natural Science Foundation.

ABSTRACT Rolling bearings are one of the essential components in rotating machinery. Efficient bearing fault diagnosis is necessary to ensure the regular operation of the mechanical system. Traditional fault diagnosis methods usually rely on a complex artificial feature extraction process, which requires a lot of human expertise. Emerging deep learning methods can reduce the dependence of the feature extraction process on manual intervention effectively. However, its training requires a large number of fault signals, which is difficult to obtain in actual engineering. In this paper, a rolling bearing fault diagnosis method based on Convolutional Neural Network and Support Vector Machine is proposed to solve the above problems. Firstly, the Continuous Wavelet Transform is used to convert one-dimensional original vibration signals into two-dimensional time-frequency images. Secondly, the obtained time-frequency images are input for training the constructed model. Finally, the diagnosis of the fault location and severity is completed. The method is verified on the CWRU data set and the MFPT data set. The results demonstrate that the proposed method achieves higher diagnostic accuracy and stability than other advanced techniques.

INDEX TERMS Convolutional neural network, continuous wavelet transform, fault diagnosis, rolling bearing, support vector machine.

I. INTRODUCTION

Since some industrial machines need to work continuously in harsh environments, failures of critical components such as bearings often occur. As one of the basic elements of many industrial machinery, the working state of rolling bearings has a great influence on the operation of the entire equipment [1]. Therefore, the research on the fault diagnosis technology of rolling bearings is significant for the safety of the production process and the reduction of economic losses. With the development of machine learning, many typical intelligent methods have been successfully used in fault diagnosis research, mainly including of two stages: signal feature extraction and fault classification [2].

The vibration signals of bearings usually contain sufficient fault information, but most of them are nonlinear and non-stationary. Therefore, signal feature extraction is a crucial step [3]. Time-frequency analysis is a powerful tool in signal processing, which can analyze the time domain and frequency domain as a whole. Commonly used time-frequency analysis methods include Empirical Mode Decomposition (EMD) [4], Short-Time Fourier Transform (STFT) [5], and Wavelet Transform (WT) [6]. The signal can be adaptively decomposed by EMD into intrinsic modal function components with different scales. However, there is a problem of modal confusion in this process. Although STFT has the ability to realize the time-frequency analysis of the signal, it cannot adequately reflect the sudden change of the vibration signal because its time resolution is fixed. WT is a time-frequency analysis technology, of which time window can shrink as the frequency of the signal increases, and vice versa. WT expands STFT and effectively makes up for its shortcomings, so it is widely used. Yan *et al.* have summarized the application of Continuous Wavelet Transform (CWT), Discrete Wavelet Transform (DWT), Wavelet Packet Transform (WPT), and Second-Generation Wavelet Transform (SGWT) in the field of fault diagnosis [7]. On the other hand, fault identification also plays a vital role in

The associate editor coordinating the review of this manuscript and approving it for publication was Yan-Jun Liu.

fault diagnosis because it cannot meet the requirements of data batch processing only by fault extraction. Traditional fault recognition tools include Bayesian classifier [8], Artificial Neural Network (ANN) [9]–[10], and Support Vector Machine (SVM). Under the condition that there are enough training samples, the first two can distinguish the fault types effectively. However, a large number of available failure samples are hard to obtain in actual work. SVM can achieve effective classification through a small amount of samples due to its strong network generalization ability, good generality, and high classification accuracy. Therefore, SVM has been widely used in mechanical fault diagnosis research [11]–[14]. However, SVM performs poorly on redundant data because its shallow structure has some difficulties in learning in-depth features [15].

Recently, deep learning has become a vital research direction and gradually applied to various fields [16]–[19]. The deep learning model is composed of multi-layer neural networks, which can extract and learn the in-depth features of the input signal. Deep learning method can handle the complex and high-dimensional problems in the massive data that cannot be solved by shallow learning [20]. Due to its high efficiency, plasticity and universality, scholars have applied many deep learning models to the research of fault diagnosis, such as Long Short-Term Memory (LSTM) [21], Deep Belief Network (DBN) [22], Deep Auto-encoder (DAE) [23], Gated Recurrent Unit Network (GRUN) [24], and Convolutional Neural Network (CNN). Among them, CNN has more sophisticated applications in image processing, including image classification [25], target positioning [26], and face recognition [27]. It has received more attention in the study of rolling bearing fault diagnosis. In [28], a deep CNN structural model that can automatically classify rolling bearing faults is established. In [29], a method for rolling bearing fault diagnosis based on Cyclic Spectral Coherence (CSCoh) and CNN is proposed, which improves the fault recognition performance. To simplify the network architecture, Hierarchical Symbolic Analysis (HSA) and CNN are combined for bearing fault diagnosis in [30]. However, the above method relies on large quantities of samples that can be used for training, which is difficult to obtain in actual engineering. Moreover, the construction time of the deep network is relatively long.

Considering the advantages and disadvantages of all the above work, CNN and SVM are combined to build a deep neural network framework CNN-SVM to diagnose bearing faults in this paper. In the first place, CWT is used for preliminary feature extraction. Then, the CNN-SVM network model is constructed using the method of transfer learning. In the end, the preliminary extracted features are used to train the network model to achieve bearing fault classification. This method makes full use of the excellent feature extraction capability of CNN and the exceptional classification performance of SVM, which solves the problem of in-depth signal feature extraction and the difficulty of obtaining massive samples in practical work.

This paper is organized as follows: Section 2 is dedicated to the theory of CWT, CNN, and SVM. In Section 3, the proposed CNN-SVM model is presented. In Section 4, the complete experimental procedure and analysis of the results are introduced. Finally, the conclusion is given in Section 5.

II. THEORETICAL BACKGROUND

In this paper, an intelligent diagnosis method of rolling bearing faults is proposed. Firstly, we convert the original vibration signals into time-frequency images using CWT. Then, CNN is applied to extract the in-depth features of the time-frequency images. Finally, the classifier SVM is trained using the extracted features. The fundamental theories of CWT, CNN, and SVM are introduced as follows.

A. CWT FOR TIME-FREQUENCY ANALYSIS

The CWT time-frequency analysis method performs multi-scale refinement on the signal through scaling and translation operations. Therefore, CWT can automatically adapt to the requirements of time-frequency signal analysis, clearly describing the change of signal frequency with time. [7]. Here, CWT is used for preliminary feature extraction, converting the original 1-D time-domain signals into 2-D time-frequency images. The conversion process is shown in Fig. 1 [31].

1) CONTINUOUS WAVELET TRANSFORM

CWT is a method to obtain characteristic signal information, which can be used for the processing and analysis of nonlinear signals. Its algorithm is relatively mature, and the basic definition can be expressed as [32]:

$$W_{\varphi}(a,b) = \frac{1}{\sqrt{a}} \int x(t)\varphi^*(\frac{t-b}{a})dt, \quad a > 0 \tag{1}$$

where a represents the scale parameter, b represents the time or translation parameter, $x(t)$ represents the original one-dimensional data signal, φ represents the wavelet function with scale a and position offset b, and φ^* is the complex conjugate of φ.

2) THE OPTIMAL WAVELET BASIS

An optimal wavelet basis function (WBF) is essential in any signal processing using CWT. For the same signal, we will get different results by using different WBF analysis. WBF should be selected based on two aspects, including the general properties of wavelet and the characteristics of the analyzed object. Its general principles include orthogonality, tight support, symmetry, and smoothness. When we use CWT for pulse signal processing, the closer the shape of WBF is to the signal waveform, the more features can be obtained. The formula of the similarity between the two is as follows [33]:

$$\delta = \sum_{i=1}^{k} \alpha_i \frac{m_i^2}{s_i} \tag{2}$$

where δ represents the similarity coefficient, m_i represents the maximum value of each peak after WBF is taken as the

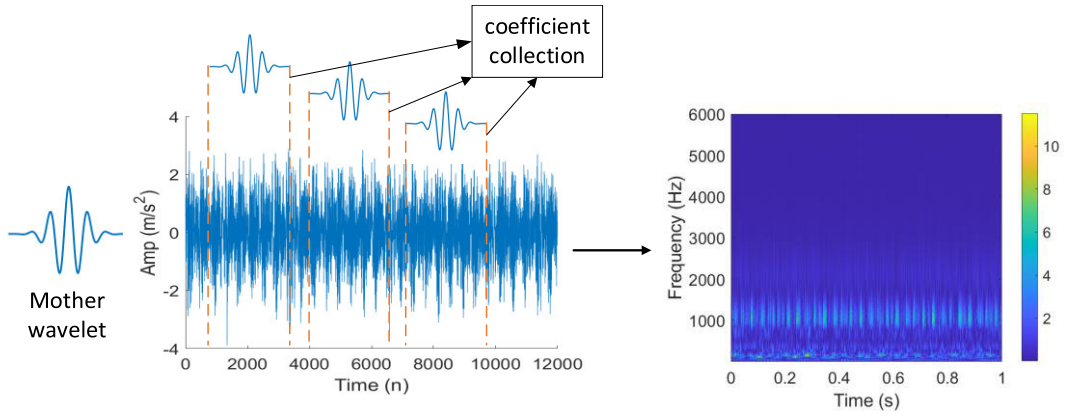

FIGURE 1. The conversion process from time-domain signal to time-frequency image.

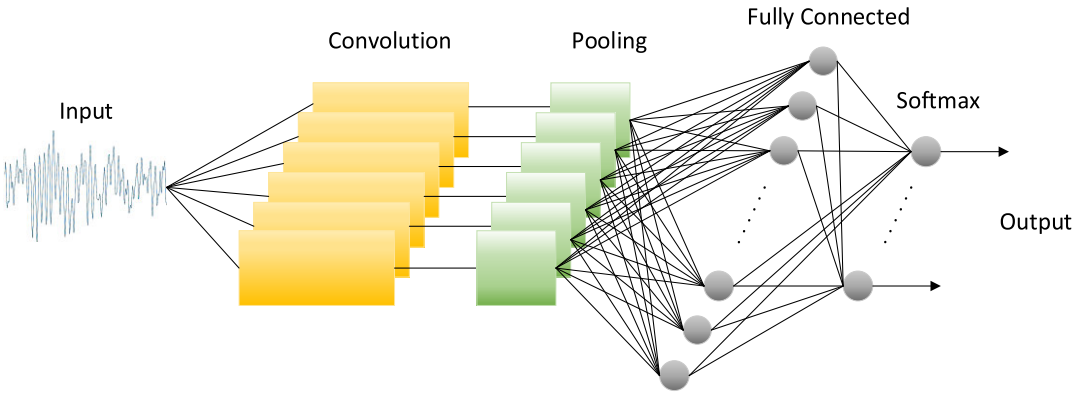

FIGURE 2. The architecture of the CNN model.

absolute value, s_i represents the area covered by each peak after WBF is taken as the absolute value, α_i represents the weighted coefficient of each peak after WBF is taken as the absolute value, $\alpha_i = \frac{m_i}{\max(m_i)}$, and k is the number of peaks after WBF is taken as the absolute value.

B. CONVOLUTIONAL NEURAL NETWORK

As a common method for extracting data features in deep learning models, CNN has made remarkable achievements in image recognition research [34]. The overall structure of CNN is demonstrated as Fig. 2. Its internal hidden layer structure is mainly composed of convolution layer, pooling layer, and fully-connected layer.

1) CONVOLUTION LAYER

In the convolution layer, a more advanced feature representation is obtained through convolution operations. The number of network parameters is decreased and the complexity of the model is reduced by performing local feature extraction on input data. The convolution formula can be defined as follows:

$$x_j^l = f\left(\sum_{i \in M_j} x_i^{l-1} k_{ij}^l + b_j^l\right) \tag{3}$$

where x_j^l is the output of layer l, x_i^{l-1} represents the output of layer $l-1$, that is, the input of layer l, M_j is the feature set of layer $l-1$, k_{ij}^l represents the weight matrix, b_j^l represents the network bias, and $f(\cdot)$ represents the activation function.

2) POOLING LAYER

In the pooling layer, the data is down-sampled by calculating the local average or maximum value, which reduces the network calculation complexity and retains the most critical features, thereby improving the efficiency of feature extraction. The calculation method can be expressed as:

$$x_j^l = f\left(\beta_j^l down\left(x_j^{l-1}\right) + b_j^l\right) \tag{4}$$

where $down(\cdot)$ is the down-sampling function, β represents the weight of the network.

3) FULLY-CONNECTED LAYER

The image features are input into the fully connected layer after being alternately transferred through convolutional layers and pooling layers. In the fully-connected layer, the deep feature information with category distinction is integrated, and the mapping relationship between the extracted features and sample types is constructed. The mathematical formula

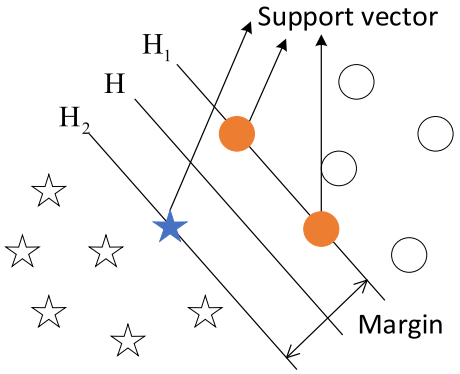

FIGURE 3. Schematic diagram of the SVM classification principle.

can be expressed as:

$$y^k = f\left(\omega^k x^{k-1} + b^k\right) \tag{5}$$

where k represents the network of layer k, y^k represents the output of the fully connected layer, x^{k-1} is the input of the fully connected layer, ω^k is the weight coefficient, b^k represents the network offset.

C. SUPPORT VECTOR MACHINE

SVM is a commonly used classifier that shows excellent advantages in resolving small sample, nonlinear, and high-dimensional pattern recognition problems. As indicated in Fig. 3, the basic model of SVM is to find an optimal hyperplane on the feature space to divide the dataset. In Fig. 3, H is the classification surface, H_1 and H_2 are straight lines parallel to H and passing the data points closest to it in the two types of data. The data point is called support vector, and the distance between H_1 or H_2 and H is called the geometric separation. Therefore, maximizing the geometric interval becomes the training goal of SVM [35].

Suppose a linearly categorizable training sample dataset, $D = \{(x_i, y_i), i = 1, \ldots, n\}$, where x_i represents K-dimensional column vector, that is, $x \in R^{K\times 1}$, y_i is the category of x_i, and $y_i \in \{-1, 1\}$. The total decision function is as follows:

$$y_i = sgn\left((\omega \cdot x) + b\right) \tag{6}$$

where sgn represents a symbolic function, ω represents the weight vector, and b is the deviation.

Therefore, the problem of maximizing the geometric interval can be transformed into the solving of the quadratic programming problem:

$$\min = \left\{\frac{1}{2}\|\omega\|^2\right\} \tag{7}$$

the constraint is:

$$y_i\left(\omega \cdot x_i + b\right) \geq 1, \quad i = 1, \ldots, n. \tag{8}$$

TABLE 1. Architecture and parameters of ResNet-18.

Layer name	Output size	Learnables
conv1	112×112	7×7,64, stride 2
conv2_x	56×56	3×3 max pool, stride 2
		$\begin{bmatrix} 3\times 3,64 \\ 3\times 3,64 \end{bmatrix} \times 2$
conv3_x	28×28	$\begin{bmatrix} 3\times 3,128 \\ 3\times 3,128 \end{bmatrix} \times 2$
conv4_x	14×14	$\begin{bmatrix} 3\times 3,256 \\ 3\times 3,256 \end{bmatrix} \times 2$
conv5_x	7×7	$\begin{bmatrix} 3\times 3,512 \\ 3\times 3,512 \end{bmatrix} \times 2$
	1×1	average pool, 1000-d fc, softmax
FLOPs		1.8×10^9

To solve the above problem, the Lagrange function is introduced:

$$L\left(\omega, b, \alpha\right) = \frac{1}{2}\|\omega\|^2 - \sum_{i=1}^{n}\alpha_i y_i\left(\omega \cdot x_i + b\right) + \sum_{i=1}^{n}\alpha_i \tag{9}$$

where α_i represents the Lagrange multiplier, $\alpha_i \geq 0$. The total decision function can be expressed as:

$$f\left(x\right) = sgn\left(\sum_{i=1}^{n}\alpha_i y_i\left(x_i \cdot x\right) + b\right) \tag{10}$$

However, the samples in the original space are often nonlinear and inseparable. Therefore, the introduction of an appropriate kernel function is necessary. Finally, the basic model of SVM can be obtained:

$$f\left(x\right) = sgn\left(\sum_{i=1}^{n}\alpha_i y_i K\left(x_i, x\right) + b\right) \tag{11}$$

where K represents the kernel function. The commonly used kernel functions include linear, polynomial, S-shaped, and Gaussian radial basis kernel function. In this paper, the Gaussian radial basis function is selected.

III. THE PROPOSED CNN-SVM MODEL

In this paper, a network architecture based on CNN and SVM is established for rolling bearing fault diagnosis, which we call CNN-SVM model. This chapter mainly includes the construction of CNN-SVM model and bearing fault diagnosis based on the proposed model.

A. DESIGN OF CNN-SVM

In this section, the CNN-SVM model is constructed, of which structure is indicated as Fig. 4. The model uses transfer learning of the pre-trained ResNet-18 network to extract 2-D image features, and uses the extracted features to train the classifier SVM, so as to achieve the purpose of image classification.

In this system, the method of transfer learning in deep learning is used, and the ResNet-18 network is selected for feature extraction. Its structure and parameters are shown in Table 1 [36]. Applying the relevant knowledge that has

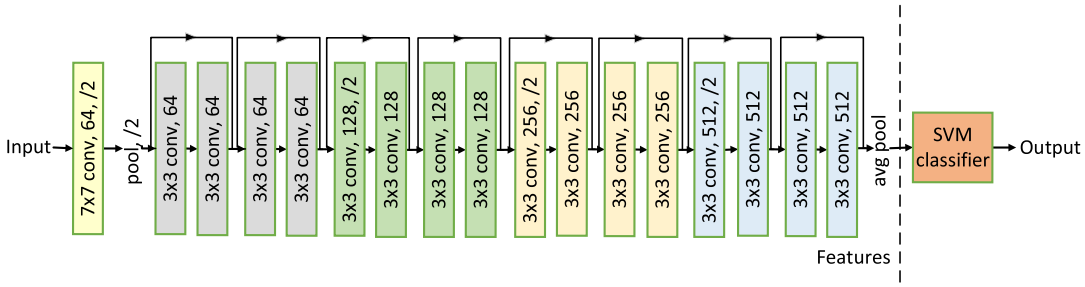

FIGURE 4. The architecture of the proposed CNN-SVM model.

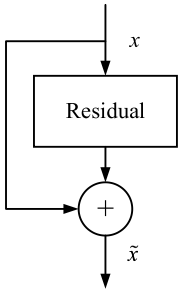

FIGURE 5. The residual structure.

been learned in the pre-training network directly to the target field can effectively solve the problem of complicated and time-consuming manual construction of CNN and the problem of insufficient data samples obtained in actual engineering. The ResNet-18 network adopts the idea of residual learning to guard against the problem of gradient disappearance to a certain extent. A typical residual structure is shown in Fig. 5.

SVM classifies different states by finding the best hyperplane that separates different data elements, and possesses exceptional generalization ability and high classification accuracy. In this study, the Softmax layer in the CNN model is replaced by an SVM classifier. After the CNN is finished trained, the extracted image features that output from the fully-connected layer are used to train the SVM classifier. Finally, the trained SVM-CNN model can be used for image type recognition.

B. THE PROPOSED METHOD FOR BEARING FAULT DIAGNOSIS

Since the amount of available bearing fault data in practical engineering is small and the feature extraction of traditional intelligent diagnosis methods is more complicated, a method for fault detection of rolling bearings based on CNN-SVM is proposes in this study. This method can take advantage of the superiority of CNN in data feature obtaining and the classification and generalization capabilities of SVM to improve diagnostic performance. The corresponding flow chart of the proposed method is demonstrated as Fig. 6, and the steps are as follows:

Step 1: The vibration signals of faulty bearings are collected on two different experimental platforms.

Step 2: First, the vibration signals are split into segments. Then, we use CWT to convert the collected vibration signals into time-frequency images with a size of 224 × 224 ×3. Finally, the time-frequency images are partitioned into training samples and test samples, and a label is set for each sample.

Step 3: First, load the pretrained ResNet-18 network. Then, input the samples processed in step 2 to the model and get the high-level feature representations of the training and test images from the fully connected layer. Finally, use the features extracted from the training samples as predictor variables to train the SVM.

Step 4: Test samples are input to the trained model to verify the validity of the diagnosis.

IV. EXPERIMENTS

To verify the flexibility and utility of the constructed model in bearing fault diagnosis, two open-source datasets are used for research in this paper, including Case Western Reserve University (CWRU) Bearing Data Center dataset [37] and the Society for Mechanical Failure Prevention Technology (MFPT) dataset [38].

A. CWRU DATA

1) DATA DESCRIPTION

The CWRU data set is one of the most widely used datasets in bearing fault diagnosis. As illustrated in Fig. 7, the experiment platform consists of four parts: a driven motor, a torque sensor, a dynamometer and an electronic controller. SKF 6205-2RS JEM deep groove ball bearings are used in the experiment. The sampling frequency of the drive end bearing is 12 kHz and 48 kHz, and the sampling frequency of the fan end bearing is 12 kHz. The single-point damage method of EDM is used to set the damage diameters of 0.1778, 0.3556, 0.5334, and 0.7112 mm on the inner ring and rolling element respectively, and the damage diameters of the outer ring at 6 o'clock are set to 0.1778, 0.3556, and 0.5334 mm. The vibration data of bearings in normal status and 11 types of fault status are collected.

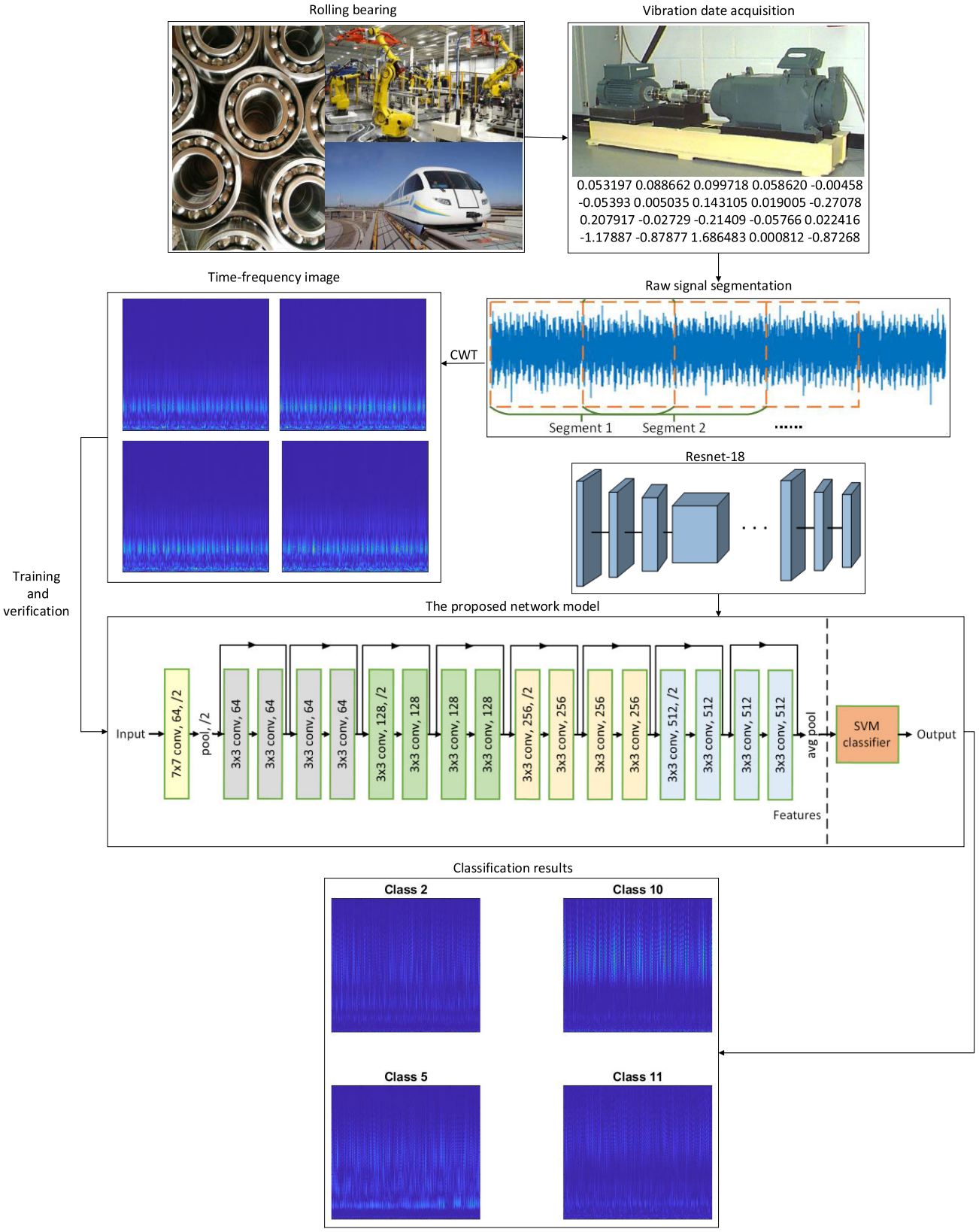

FIGURE 6. The flowchart of the experiment.

In this study, the signals collected at the frequency of 12kHz under 12 health conditions of the drive end bearing is selected. In this way, different fault locations and damage severities of bearings can be better simulated.

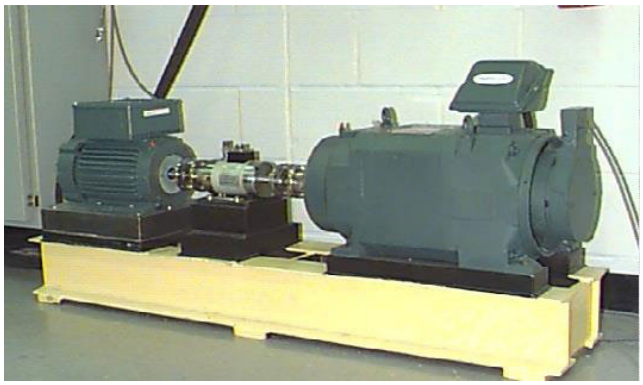

FIGURE 7. CWRU bearing test rig.

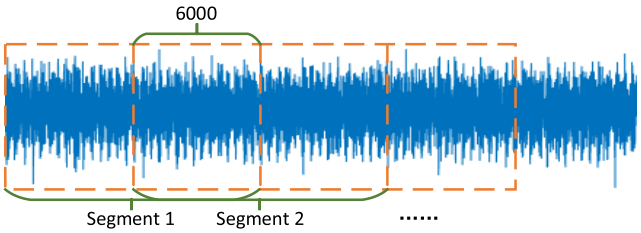

FIGURE 8. Signal segmentation.

TABLE 2. Sample description of the CWRU dataset.

Fault types	Fault diameters (mm)	Number of samples	Lable
Normal status	0	76	Class 1
Fault in inner race	0.1778	76	Class 2
	0.3556	76	Class 3
	0.5334	76	Class 4
	0.7112	76	Class 5
Fault in ball	0.1778	76	Class 6
	0.3556	76	Class 7
	0.5334	76	Class 8
	0.7112	76	Class 9
Fault in outer race	0.1778	76	Class 10
	0.3556	76	Class 11
	0.5334	76	Class 12

TABLE 3. Similarity coefficients of five commonly used wavelet basis functions.

Wavelet basis functions	Db10	Morlet	Meyer	Sym8	Coif5
Similarity coefficients	7.4970	7.2148	6.6082	6.3626	6.3298

2) DATA PROCESSING

In this experiment, 120000 data points are taken from the data set under four load conditions of 12 health states of the bearing, respectively. In order to make each sample contain enough information, 12000 data points are taken as the sample length. As shown in Fig. 8, 6000 data points are repeated between adjacent parts for a division to obtain more experimental samples. In this way, each group of fault data can be divided into 19 segments. Therefore, 19 time-frequency images in 48 states can be obtained through CWT. By merging the load types under 12 health conditions, 76 time-frequency images for each health state are obtained. Finally, a sample set containing 912 time-frequency images is obtained, as shown in Table 2. Among them, 70% of the images under each fault type are randomly chosen as training samples, and the remaining 30% of the images are used as test samples. In Fig. 9 (a)-(l), the time-frequency images of the 12 health conditions in Table 2 are shown in order.

Db10, Meyer, Sym8, Coif5, Morlet are several commonly used wavelet basis functions, and their similarity coefficients to bearing fault characteristics are listed in Table 3 [33]. According to the Table 3, the similarity coefficient of db10 wavelet is the largest. However, Db10 wavelet has an asymmetric shape, poor smoothness, and complicated theoretical analysis. Morlet wavelet is a single-frequency sine function under Gaussian envelope, of which the similarity coefficient is second only to Db10 and shape is symmetrical and smooth. Therefore, Morlet wavelet is chosen here as the WBF of CWT for the generation of time-frequency images of vibration signals.

3) RESULTS OF THE EXPERIMENT

First, the proposed CNN-SVM model is trained by the constructed training samples of the 12 bearing fault states. After the training is completed, the remaining test samples are used to verify the trained model. To prove the stability of the proposed model, the experiment is repeated ten times. The diagnosis accuracy and training time of 10 trials are revealed in Table 4. The average classification accuracy is 98.75%, and the highest diagnosis accuracy is 99.29%. A fault diagnosis result is shown in Fig. 10.

B. MFPT DATA

1) DATA DESCRIPTION

The MFPT dataset provided by the Society for Machinery Failure Prevention Technology is also applied for the analysis and research of rolling bearing faults. The MFPT dataset consists of three sets of experimental bearing vibration data and three actual fault data. Among them, the three sets of experimental bearing vibration data include baseline set, inner race fault set, and outer race fault set. The baseline dataset contains three files, and the data in each file are obtained by sampling at a frequency of 97656 Hz for 6 seconds under a load of 270 pounds. The inner race fault dataset contains seven files, which are respectively obtained by sampling at 48828 Hz for 3 seconds under seven load conditions, including 0, 50, 100, 150, 200, 250, and 300 pounds. The outer race fault dataset contains seven files, which are respectively obtained by sampling at 48828 Hz for 3 seconds under seven load conditions, including 25, 50, 100, 150, 200, 250, and 300 pounds. The data points are from a single channel radial accelerometer. The experiment bearings of the MFPT data set is also deep

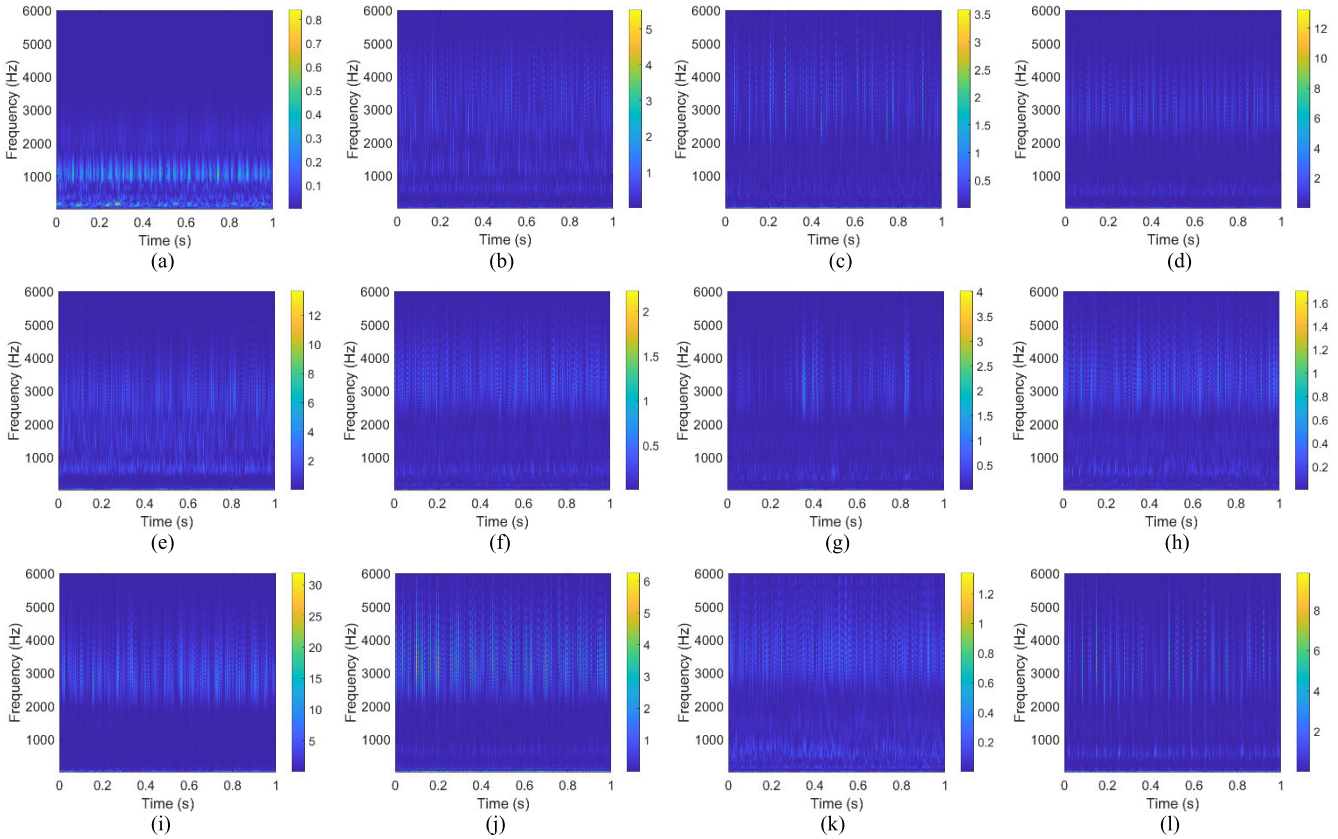

FIGURE 9. Time-frequency images of twelve health conditions on CWRU dataset: (a) Class 1; (b) Class 2; (c) Class 3; (d) Class 4; (e) Class 5; (f) Class 6; (g) Class 7; (h) Class 8; (i) Class 9; (j) Class 10; (k) Class 11; (l) Class 12.

TABLE 4. Diagnostic results of 10 trials on CWRU dataset.

Trial number	1	2	3	4	5	6	7	8	9	10	Average
Accuracy (%)	99.29	99.29	98.21	98.21	98.21	98.57	98.57	98.57	99.29	99.29	98.75
Training time (s)	35.10	35.59	34.45	37.66	35.26	34.46	36.38	33.88	37.73	37.72	35.82

groove ball bearing with a 31.62 mm pitch diameter, a 5.97 mm ball diameter, a 0° contact angle, and an element number of 8.

In this study, the three sets of bearing vibration data collected at the test rig are selected.

2) DATA PROCESSING

In this experiment, all the data points in the three fault data sets are used, and 48828 data points are taken as the sample length. In order to obtain more experimental samples, 24414 data points are repeated between adjacent parts. The baseline set is down-sampled to 48828 Hz to match other fault sets. The three files in the baseline set are divided into 22 segments respectively, the seven inner fault files are divided into five segments respectively, and the seven outer fault files are also divided into five segments respectively. Through the CWT, a corresponding number of time-frequency images in each state can be obtained. By merging the load conditions

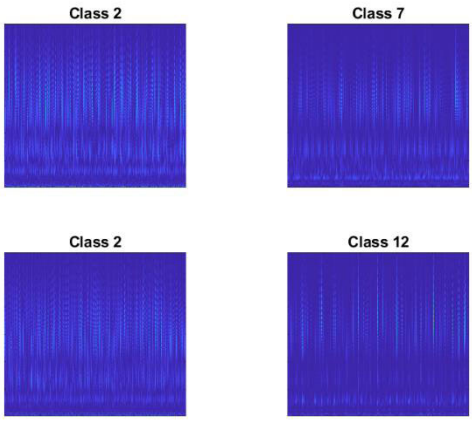

FIGURE 10. Classification results on CWRU dataset.

under the three fault types, 66 time-frequency images for the baseline set, 35 time-frequency images for the inner race and 35 time-frequency images for the outer race are obtained.

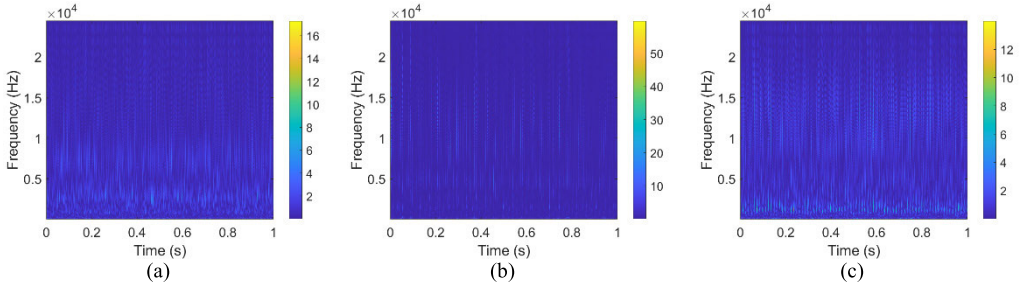

FIGURE 11. Time-frequency images of three health conditions on MFPT dataset: (a) Class 1; (b) Class 2; (c) Class 3.

TABLE 5. Sample description of the MFPT dataset.

Fault types	Load (lb)	Number of samples	Class
Normal status	270	66	1
Fault in inner race	0, 50, 100, 150, 200, 250, 300	35	2
Fault in outer race	25, 50, 100, 150, 200, 250, 300	35	3

Finally, a sample set containing 136 time-frequency images is obtained, as shown in Table 5. Among them, 70% of the images in each state are randomly chosen as training samples, and the remaining 30% of the images are used as test samples. In Fig. 11 (a)-(c), the time-frequency images of the three health conditions in Table 5 are shown in sequence.

3) RESULTS OF THE EXPERIMENT

First, the CNN-SVM model is trained by using the constructed training samples of the three bearing fault types. After the training is completed, the trained model is verified by the test sample. To prove the stability of the proposed model, the experiment is repeated ten times. The diagnosis accuracy and training time of 10 trials are revealed in Table 6. The average classification accuracy is 98.89%, and the highest diagnosis accuracy reaches 100%. A fault diagnosis result is indicated in Fig. 12.

C. COMPARISON WITH OTHER METHODS

To prove the superiority of the proposed model, we compare it with standard CNN and standard SVM models. The inputs of standard CNN and SVM are also the time-frequency images, and the classifier of the standard CNN is Softmax. In order to ensure the fairness of the experiment and the reliability of the results, the test is repeated five times on these two models by using the CWRU dataset similarly. The diagnosis accuracy of each method is shown in Fig. 13. The diagnosis accuracy and average training time of the three methods are summarized in Table 7.

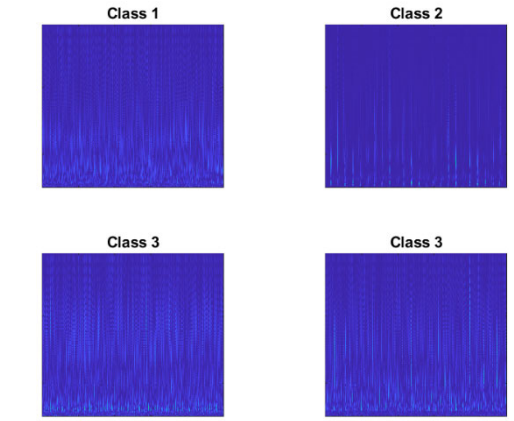

FIGURE 12. Classification results on MFPT dataset.

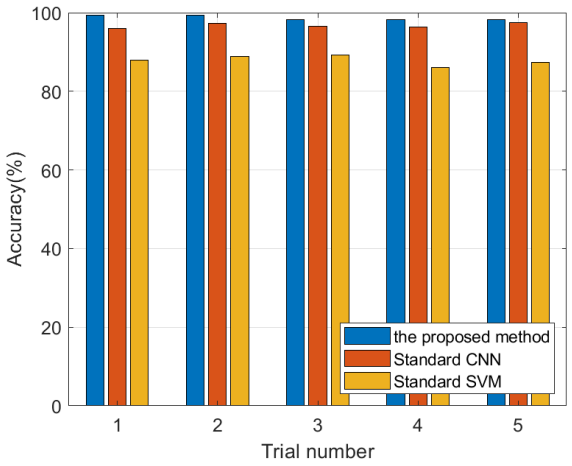

FIGURE 13. Classification accuracy of three different methods.

TABLE 6. Diagnostic results of 10 trials on MFPT dataset.

Trial number	1	2	3	4	5	6	7	8	9	10	Average
Accuracy (%)	96.3	96.3	100	100	100	100	96.3	100	100	100	98.89
Training time (s)	17.07	16.15	16.16	16.43	16.46	16.95	16.27	16.68	16.76	16.63	16.56

TABLE 7. Accuracy and average training time of the three models.

Models	Classifier	Accuracy (%)	Average training time (s)
The proposed model	SVM	98.75±0.46	35.68
Standard CNN	Softmax	96.72±0.58	730
Standard SVM	SVM	87.88±1.15	20.41

TABLE 8. Average accuracy of the proposed method and several other methods.

Categories	Methods	Average accuracy (%)
Shallow learning	PNN-SFAM [9]	97.82
	BPNN [39]	77.86
Deep learning	CNN-HMM [39]	98.125
	DAFD [40]	94.73
	DGNN [41]	97.81
	The proposed method	98.75

As can be seen, the diagnosis accuracy of the proposed method in 5 trials is higher than that of the other two standard models, and the training time is considerably shortened compared to the standard CNN model. The results show that SVM is better than the default classifier, i.e., Softmax. It is proved that the proposed model can effectively solve the difficulty of extracting in-depth features of miscellaneous data by SVM and the challenge of meeting the needs of massive samples for the training of CNN, and has excellent stability.

In recent years, there are many works about deep learning-based fault diagnosis methods. Thus, to further illustrate the innovation of the proposed method in the field of bearing fault diagnosis, we compare it with existing advanced techniques such as PNN-SFAM [9], BPNN [39], CNN-HMM [39], DAFD [40], and DGNN [41]. The average diagnosis accuracy of each method is listed in Table 8.

The details of the different methods are listed below. In [9], the PNN is introduced to classify rolling bearing states in healthy and not healthy. If the PNN decision shows that the processed state is not healthy, then the SFAM is used to classify 7 types of faults. In [39], for the BPNN method, 50 units in hidden layer, learning rate is 0.2, momentum is 0.05, and iteration number is 800. 12 types of health conditions are classified by using the learned features of the experimental data as inputs. For the CNN-HMM method, a 50×50 matrix of each sample is constructed, learning rate is 1, and maximum iteration is 100. 12 health conditions are classified by using raw data without any signal preprocessing as inputs. In [40], a deep neural network model with domain adaptability, called DAFD, has been proposed for fault diagnosis. The classification model is directly trained by using the labeled samples of 4 health conditions collected without a motor load, and then the data collected under a motor load of 3hp is classified. In [41], a fault diagnosis method based on deep generative neural networks (DGNN) has been proposed. The model is trained by using the samples of 10 health conditions collected under a motor load of 0hp, and then complete classification for the 10 health conditions under the working situation with a motor load of 1, 2, and 3hp.

The architecture and parameters of ResNet-18 which selected by the proposed method for feature extraction are given in Table 1. The selection of other hyperparameters is as follows. The mini batch size is 256. The learning rate starts from 0.1 and is divided by 10 when the error reaches a plateau, and the models are trained for up to 60×104 iterations, the weight decay is 0.0001, and the momentum equals to 0.9 [36]. The accuracies of the 5 contrastive methods are 97.82%, 77.86%, 98.125%, 94.73%, and 97.81%, respectively.

By comparing the results with other methods, it can be easily seen that the proposed method achieves a higher diagnosis accuracy, which further shows the effectiveness of the proposed method.

V. CONCLUSION

In this study, a new deep neural network model CNN-SVM is built for fault diagnosis of rolling bearings. First, we use the CWT to construct the time-frequency images of vibration signals. Then, the obtained images are input to the proposed model for training. Finally, the diagnosis of fault location and severity of the rolling bearing is completed. The experiments indicate that the diagnostic accuracy of this method can reach 98.75% for the CWRU dataset and 98.89% for the MFPT dataset which verify the flexibility and practicability of the constructed model. By comparing with the standard CNN and standard SVM, it is showed that the proposed model is able to resolve the difficulty of deep feature extraction in the traditional method and the small sample problem in actual engineering and has excellent stability. By comparing with advanced methods such as PNN-SFAM, BPNN, CNN-HMM, DAFD, and DGNN, the effectiveness of the proposed method is further verified.

However, for relatively noisy data sources, the accuracy of the proposed method still needs to be enhanced. Therefore, the structure of the proposed model needs further improvement in the future. Moreover, in this paper, only single-fault bearing vibration signals are used for model training, and no compound fault samples are created to simulate the actual situation, which is a certain challenge for the application of the proposed model in practical engineering. It has also become a research direction for us in the future.

REFERENCES

[1] D.-T. Hoang and H.-J. Kang, "A survey on deep learning based bearing fault diagnosis," *Neurocomputing*, vol. 335, pp. 327–335, Mar. 2019.
[2] M. Kang, M. R. Islam, J. Kim, J.-M. Kim, and M. Pecht, "A hybrid feature selection scheme for reducing diagnostic performance deterioration caused by outliers in data-driven diagnostics," *IEEE Trans. Ind. Electron.*, vol. 63, no. 5, pp. 3299–3310, May 2016.
[3] Q. Hu, X.-S. Si, A.-S. Qin, Y.-R. Lv, and Q.-H. Zhang, "Machinery fault diagnosis scheme using redefined dimensionless indicators and mRMR feature selection," *IEEE Access*, vol. 8, pp. 40313–40326, 2020.
[4] X. Yu, F. Dong, E. Ding, S. Wu, and C. Fan, "Rolling bearing fault diagnosis using modified LFDA and EMD with sensitive feature selection," *IEEE Access*, vol. 6, pp. 3715–3730, 2018.

[5] D. Zhong, W. Guo, and D. He, "An intelligent fault diagnosis method based on STFT and convolutional neural network for bearings under variable working conditions," in *Proc. Prognostics Syst. Health Manage. Conf. (PHM-Qingdao)*, Qingdao, China, Oct. 2019, pp. 1–6.
[6] D. Verstraete, A. Ferrada, E. L. Droguett, V. Meruane, and M. Modarres, "Deep learning enabled fault diagnosis using time-frequency image analysis of rolling element bearings," *Shock Vib.*, vol. 2017, Oct. 2017, Art. no. 5067651.
[7] R. Yan, R. X. Gao, and X. Chen, "Wavelets for fault diagnosis of rotary machines: A review with applications," *Signal Process.*, vol. 96, pp. 1–15, Mar. 2014.
[8] V. Muralidharan and V. Sugumaran, "A comparative study of Naïve Bayes classifier and Bayes net classifier for fault diagnosis of monoblock centrifugal pump using wavelet analysis," *Appl. Soft Comput.*, vol. 12, no. 8, pp. 2023–2029, Aug. 2012.
[9] J. Ben Ali, L. Saidi, A. Mouelhi, B. Chebel-Morello, and F. Fnaiech, "Linear feature selection and classification using PNN and SFAM neural networks for a nearly online diagnosis of bearing naturally progressing degradations," *Eng. Appl. Artif. Intell.*, vol. 42, pp. 67–81, Jun. 2015.
[10] J. Li, X. Yao, X. Wang, Q. Yu, and Y. Zhang, "Multiscale local features learning based on BP neural network for rolling bearing intelligent fault diagnosis," *Measurement*, vol. 153, Mar. 2020, Art. no. 107419.
[11] Z. Huo, Y. Zhang, L. Shu, and M. Gallimore, "A new bearing fault diagnosis method based on Fine-to-Coarse multiscale permutation entropy, Laplacian score and SVM," *IEEE Access*, vol. 7, pp. 17050–17066, 2019.
[12] X. Yan and M. Jia, "A novel optimized SVM classification algorithm with multi-domain feature and its application to fault diagnosis of rolling bearing," *Neurocomputing*, vol. 313, pp. 47–64, Nov. 2018.
[13] Y. Li, Y. Yang, X. Wang, B. Liu, and X. Liang, "Early fault diagnosis of rolling bearings based on hierarchical symbol dynamic entropy and binary tree support vector machine," *J. Sound Vib.*, vol. 428, pp. 72–86, Aug. 2018.
[14] X. Zhang, Y. Liang, J. Zhou, and Y. Zang, "A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM," *Measurement*, vol. 69, pp. 164–179, Jun. 2015.
[15] S. Ahlawat and A. Choudhary, "Hybrid CNN-SVM classifier for handwritten digit recognition," *Procedia Comput. Sci.*, vol. 167, pp. 2554–2560, Jan. 2020.
[16] H. Rathore, A. K. Al-Ali, A. Mohamed, X. Du, and M. Guizani, "A novel deep learning strategy for classifying different attack patterns for deep brain implants," *IEEE Access*, vol. 7, pp. 24154–24164, 2019.
[17] S. Zhang, X. Pan, Y. Cui, X. Zhao, and L. Liu, "Learning affective video features for facial expression recognition via hybrid deep learning," *IEEE Access*, vol. 7, pp. 32297–32304, 2019.
[18] Y. Li, T. Zhang, S. Sun, and X. Gao, "Accelerating flash calculation through deep learning methods," *J. Comput. Phys.*, vol. 394, pp. 153–165, Oct. 2019.
[19] C. Zhang, H. Zhang, J. Qiao, D. Yuan, and M. Zhang, "Deep transfer learning for intelligent cellular traffic prediction based on cross-domain big data," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 6, pp. 1389–1401, Jun. 2019.
[20] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015.
[21] M. Qiao, S. Yan, X. Tang, and C. Xu, "Deep convolutional and LSTM recurrent neural networks for rolling bearing fault diagnosis under strong noises and variable loads," *IEEE Access*, vol. 8, pp. 66257–66269, 2020.
[22] H. Shao, H. Jiang, X. Li, and T. Liang, "Rolling bearing fault detection using continuous deep belief network with locally linear embedding," *Comput. Ind.*, vol. 96, pp. 27–39, Apr. 2018.
[23] S. Haidong, J. Hongkai, L. Xingqiu, and W. Shuaipeng, "Intelligent fault diagnosis of rolling bearing using deep wavelet auto-encoder with extreme learning machine," *Knowl.-Based Syst.*, vol. 140, pp. 1–14, Jan. 2018.
[24] X. Li, H. Jiang, X. Xiong, and H. Shao, "Rolling bearing health prognosis using a modified health index based hierarchical gated recurrent unit network," *Mechanism Mach. Theory*, vol. 133, pp. 229–249, Mar. 2019.
[25] A. Abu Mallouh, Z. Qawaqneh, and B. D. Barkana, "Utilizing CNNs and transfer learning of pre-trained models for age range classification from unconstrained face images," *Image Vis. Comput.*, vol. 88, pp. 41–51, Aug. 2019.
[26] G. Xu, X. Su, W. Liu, and C. Xiu, "Target detection method based on improved particle search and convolution neural network," *IEEE Access*, vol. 7, pp. 25972–25979, 2019.
[27] S. Banerjee and S. Das, "Mutual variation of information on transfer-CNN for face recognition with degraded probe samples," *Neurocomputing*, vol. 310, pp. 229–315, Oct. 2018.
[28] D.-T. Hoang and H.-J. Kang, "Rolling element bearing fault diagnosis using convolutional neural network and vibration image," *Cognit. Syst. Res.*, vol. 53, pp. 42–50, Jan. 2019.
[29] Z. Chen, A. Mauricio, W. Li, and K. Gryllias, "A deep learning method for bearing fault diagnosis based on cyclic spectral coherence and convolutional neural networks," *Mech. Syst. Signal Process.*, vol. 140, Jun. 2020, Art. no. 106683.
[30] Y. Yang, H. Zheng, Y. Li, M. Xu, and Y. Chen, "A fault diagnosis scheme for rotating machinery using hierarchical symbolic analysis and convolutional neural network," *ISA Trans.*, vol. 91, pp. 235–252, Aug. 2019.
[31] R. Chen, X. Huang, L. Yang, X. Xu, X. Zhang, and Y. Zhang, "Intelligent fault diagnosis method of planetary gearboxes based on convolution neural network and discrete wavelet transform," *Comput. Ind.*, vol. 106, pp. 48–59, Apr. 2019.
[32] L. Huang and J. Wang, "Forecasting energy fluctuation model by wavelet decomposition and stochastic recurrent wavelet neural network," *Neurocomputing*, vol. 309, pp. 70–82, Oct. 2018.
[33] P. Liang, C. Deng, J. Wu, and Z. Yang, "Intelligent fault diagnosis of rotating machinery via wavelet transform, generative adversarial nets and convolutional neural network," *Measurement*, vol. 159, Jul. 2020, Art. no. 107768.
[34] P. Liang, C. Deng, J. Wu, Z. Yang, J. Zhu, and Z. Zhang, "Compound fault diagnosis of gearboxes via multi-label convolutional neural network and wavelet transform," *Comput. Ind.*, vol. 113, Dec. 2019, Art. no. 103132.
[35] Z. Lv, B. Tang, Y. Zhou, and C. Zhou, "A novel method for mechanical fault diagnosis based on variational mode decomposition and multikernel support vector machine," *Shock Vib.*, vol. 2016, pp. 1–11, Jan. 2016.
[36] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.
[37] CWRU Dataset. *Case Western Reserve University Bearing Data Center*. Accessed: Jul. 20, 2012. [Online]. Available: http://csegroups.case.edu/bearingdatacenter/pages/download-data-file
[38] MFPT Dataset. *Society for Machinery Failure Prevention Technology*. Accessed: 2012. [Online]. Available: https://mfpt.org/fault-data-sets/
[39] S. Wang, J. Xiang, Y. Zhong, and Y. Zhou, "Convolutional neural network-based hidden Markov models for rolling element bearing fault identification," *Knowl.-Based Syst.*, vol. 144, pp. 65–76, Mar. 2018.
[40] W. Lu, B. Liang, Y. Cheng, D. Meng, J. Yang, and T. Zhang, "Deep model based domain adaptation for fault diagnosis," *IEEE Trans. Ind. Electron.*, vol. 64, no. 3, pp. 2296–2305, Mar. 2017.
[41] X. Li, W. Zhang, and Q. Ding, "Cross-domain fault diagnosis of rolling element bearings using deep generative neural networks," *IEEE Trans. Ind. Electron.*, vol. 66, no. 7, pp. 5525–5534, Jul. 2019.

LAOHU YUAN received the B.S. degree in automation from Air Force Engineering University, Xi'an, China, in 2002, and the M.S. and Ph.D. degrees in aircraft design from the Harbin Institute of Technology, Harbin, China, in 2004 and 2010, respectively. He is currently a Professor Shenyang Aerospace University, Shenyang, China. His research interests include intelligent systems and intelligent control.

DONGSHAN LIAN received the B.S. degree in automation from the Hunan University of Science and Technology, Xiangtan, China, in 2018. She is currently pursuing the M.S. degree in aircraft design and engineering with Shenyang Aerospace University, Shenyang, China. Her research interests are in intelligent systems and the application of machine learning to fault diagnosis.

XUE KANG received the B.S. degree in aircraft design and engineering from Shenyang Aerospace University, Shenyang, China, in 2018, where she is currently pursuing the M.S. degree in aircraft design. Her research interests are in intelligent systems and the application of machine learning to fault diagnosis.

YUANQIANG CHEN received the B.S. degree in aircraft design and engineering from Shenyang Aerospace University, Shenyang, China, in 2019, where he is currently pursuing the M.S. degree in aircraft design. His research interest includes intelligent control.

KEJIA ZHAI received the B.S. degree in traffic and transport from the Anyang Institute of Technology, Anyang, China, in 2019. He is currently pursuing the M.S. degree in aircraft design and engineering with Shenyang Aerospace University, Shenyang, China. His research interests are computer vision and artificial intelligence.