


ORIGINAL ARTICLE

Open Access



Fault Diagnosis Method for Rotating Machinery Based on Multi-scale Features

Ruijun Liang^{1*} , Wenfeng Ran¹, Yao Chen¹ and Rupeng Zhu¹

Abstract

The vibration signals of rotating machinery usually contain various natural oscillation modes, exhibiting multi-scale features. This paper proposes a Multi-Branch one-dimensional deep Convolutional Neural Network model (MBCNN) that can extract multi-scale features from raw data hierarchically, thereby improving the diagnostic accuracy of gearbox faults in noisy environments. Meanwhile, the algorithms for multi-branch generation and algorithms of the convolution and pooling for each branch are deduced. The MBCNN integrates multiple branches with interrelated convolution kernels of different widths, and each branch can extract the high-level features of the signal. The network parameters of each branch are adjusted by the loss function, which makes the features of the branches complementary. Through the design of MBCNN, the local, global, deep layer and comprehensive information can be obtained from the raw data. On the widely used Case Western Reserve University Bearing Dataset, this paper conducted a performance comparison between the proposed MBCNN and other baselines including the shallow learning methods, 1D-CNN, and multi-scale feature learning methods. Moreover, our gearbox dataset was conducted on a fault diagnosis platform, and a series of experiments were conducted to verify the effectiveness and superiority of the MBCNN. The results indicate that the MBCNN can identify the faults in the gearbox with an accuracy of higher than 92%, and the average validation time per sample is less than 3.2 ms. In a noisy environment, the diagnostic accuracy can reach 90%. The proposed MBCNN provides an effective and intelligent detection method to identify the faults of rotating machinery in the manufacturing processes.

Keywords Fault diagnosis, Multi-scale features, Convolutional neural network, Rotating machinery, Anti-noise ability

1 Introduction

Rotating machinery is one of the most common and essential equipment in modern industry. It is used in many important machines, like gearboxes, steam turbines, gas turbines, fans, and generators. Rotating machines usually work under tough conditions and are prone to faults, so their predictive maintenance is significant for guaranteeing safe operations and reducing economic costs. Faults in the vital components, like gears and rolling-element bearings, are the main causes

of rotating machine failures [1], including damages and fractures in bearings, as well as scratches, wear, and fractures in gears.

A typical fault diagnosis is carried out in three steps: data collection, feature extraction, and fault classification. Measured data should be processed before the extraction of features of potential faults. In terms of feature extraction, several methods have been proposed, such as fast Fourier transform (FFT) [2], empirical mode decomposition (EMD) [3], wavelets multiresolution analysis (WMRA) [4], and wavelet packet analysis (WPA) [5]. Strömbergsson et al. [6] found that, in the vibration analysis of bearing faults of fan gearbox, the wavelet packet transform (WPT) could detect faults earlier and more clearly than FFT and the discrete wavelet transform (DWT). Meanwhile, many machine

*Correspondence:

Ruijun Liang
lruijun@nuaa.edu.cn

¹ College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

learning algorithms have been utilized to implement fault diagnosis, such as the K-nearest neighbor (KNN) method [7], fuzzy neural networks (FNN) [8], multi-layer perceptron (MLP) [9], and support vector machine (SVM) [10]. Gong et al. [11] used SVM at the end of the LSTM network to diagnose small faults in a multi-sensor monitoring environment. Yang et al. [12] proposed to combine the energy entropy of set empirical mode decomposition (EEMD) with an artificial neural network (ANN) for fault diagnosis of asynchronous motor. Despite the superiorities of these methods, their diagnostic accuracy is limited. They usually extract shallow features and need human intervention, like expert experiences or prior knowledge, and the process is time-consuming. Moreover, since the feature extraction and fault classification are conducted separately, the suboptimal combination of the two steps may not provide promising fault diagnosis performance.

Deep learning can integrate feature extraction and classification, and it has become an effective method for intelligent diagnosis. The convolutional neural network (CNN), one of the typical deep learning algorithms, can automatically extract local features and integrate them. Its feature extraction and generalization capability is improved with an increasing number of layers. CNN has been widely used in the fields of computer vision [13] and natural language processing. Meanwhile, some attempts using 2D-CNNs have also been conducted in fault diagnosis. For instance, Chen et al. [14] estimated the 2D cyclic spectral coherence maps of vibration signals and employed 2D-CNNs to process and classify maps to diagnose bearing faults. Pham et al. [15] utilized 2D-CNN to diagnose multi-output bearing faults, which achieved higher accuracy and efficiency than traditional CNNs. Although 2D-CNN can learn complex objects and modes and process various 2D signals (such as images and video frames), it is difficult to adapt to 1D signals. 1D-CNN performs only 1D convolutions with a simple and compact configuration, making it feasible to achieve real-time performance and low-cost hardware implementation [16]. Yan et al. [17] extended the method based on 1D-CNN to fault diagnosis of chillers. Ince et al. [18] developed an integrated fault diagnosis system that uses 1D-CNN to monitor the conditions of a motor. Wang et al. [19] proposed a one-dimensional memory-augmented convolutional neural network (1D-MACNN) and a one-dimensional memory-augmented convolutional long short-term memory (1D-MACLSTM) network, which have been successfully used in the field of structural health monitoring. These methods utilized the 1D-CNN and extracted high-level features from raw signals without involving other processing for hand-crafted feature transformation.

Though CNNs have demonstrated their capacity, current studies only focused on a fixed time scale rather than multiple scales, thus limiting their further applications. When operated at changing speeds or heavy loads, rotating machinery is vulnerable to many tiny variations, and fluctuation of instantaneous loads, faults of a component, and noise from the environment can lead to the superposition of non-stationary signals. Thus, the vibration signals of rotating machinery are complex and have multi-scale features. The features extracted from an extended time span can reflect the overall trend of the signals, while those from a shorter time span can indicate subtle local changes.

The principle of multi-scale learning is to learn features on both long-term and short-term time scales that complement each other. The multi-scale CNN (MSCNN) is developed adopting this idea. Huang et al. [20] designed a multi-scale fusion layer in an original convolutional neural network, and enhanced the ability to distinguish different fault states by fusing multi-scale information of raw signals. Jiang et al. [21] provided a multi-scale coarse-grained operation, which reduced the complexity and computation and was easier to implement than the method in Ref. [22].

The current methods that inherit MSCNN have demonstrated the capability of learning features on different time scales, but they usually use simple down-sampling and cannot learn raw signals effectively, which easily results in the loss of feature information. In this study, a fault diagnosis model called multi-branch one-dimensional convolutional neural network model (MBCNN) is proposed. Multi-branch CNN has been used in some fault diagnosis studies [23, 24], but in these studies, multiple branches of the network have the same structure, or only the optimal branch is selected for diagnosis according to the value of loss. The MBCNN model proposed in this paper can effectively learn features on multiple time scales through multiple branches with different convolutional layers and can extract features ranging from multiple time scales.

When the gearbox works, the interaction between the components and the coupling with other subsystems such as the generator make the vibration signal have various natural oscillation modes, showing multi-scale features. 1D-CNN can achieve end-to-end fault diagnosis, but it lacks multi-scale feature extraction ability.

The MBCNN proposed in this paper improves 1D-CNN with multi-scale learning. In the MBCNN, different branches adopt convolution kernels of different sizes and different convolution strides, and the first convolutional layer in each branch adopts a large convolution kernel and a large stride, thereby effectively extracting multi-scale features in the vibration signal. Moreover, the

branches containing different numbers of convolution-pooling blocks can also hierarchically extract high-level features and capture rich information for diagnosis.

2 MBCNN Model

The MBCNN works in three consecutive stages: multi-branch generation, local convolution of each branch and fully connected classification. Figure 1 illustrates the framework of MBCNN with two branches (2b-CNN). The input of the model is the raw vibration data, and the output is fault types.

2.1 Algorithm of Multi-branch Generation

The multiple branches containing features on different time scales are obtained through the convolution kernels with different widths on the first layer. Suppose the input of the l -th convolution-pooling block is $X_b^{l,c} = \{x_{b,1}^{l,c}, x_{b,2}^{l,c}, \dots, x_{b,n}^{l,c}\}^T$. The channel output in the b -th branch via a convolutional operation is $Y_b^{l,c} = \{y_{b,1}^{l,c}, y_{b,2}^{l,c}, \dots, y_{b,m}^{l,c}\}^T$, and the output length is calculated as

$$m = m_b^l = \frac{n - W_b^l + 2 \times P_b^l}{S_b^l} + 1, \tag{1}$$

where W_b^l , S_b^l and P_b^l are the width of convolution kernel, the stride of convolutional operation and the padding width, respectively; $b = 1, 2, \dots$, $l = 1, 2, \dots$, and l_b denotes the number of the layers of the b -th branch; c denotes the channel number ($c = 1, 2, \dots, c_b^l$).

The relationship between receptive fields of the adjacent pooling layers can be described as

$$R_b^{l-1} = S_b^l(d_b^l \times R_b^l - 1) + W_b^l, \tag{2}$$

where R_b^l is the receptive field of the l -th pooling layer in the b -th branch; d_b^l is the size of the pooling kernel of the l -th pooling layer in the b -th branch ($d_b^l = 2$).

Except for the first convolution layer, the parameters of other convolution layers are fixed. When $l > 1$, then $S_b^l = 1$, $W_b^l = 3$, the last pooling layer satisfies $R_b^{l_b} = 1$, thus

$$R_b^1 = 2^{(l_b-1)} \times 3 - 2. \tag{3}$$

Generally, suppose $W_b^1 = 4S_b^1$ and $d_b^1 = 2$, then the receptive field of the neurons that are fed into the fully connected layer at the input signals is

$$R_b^0 = S_b^1 \cdot (2^{l_b} \times 3 - 1). \tag{4}$$

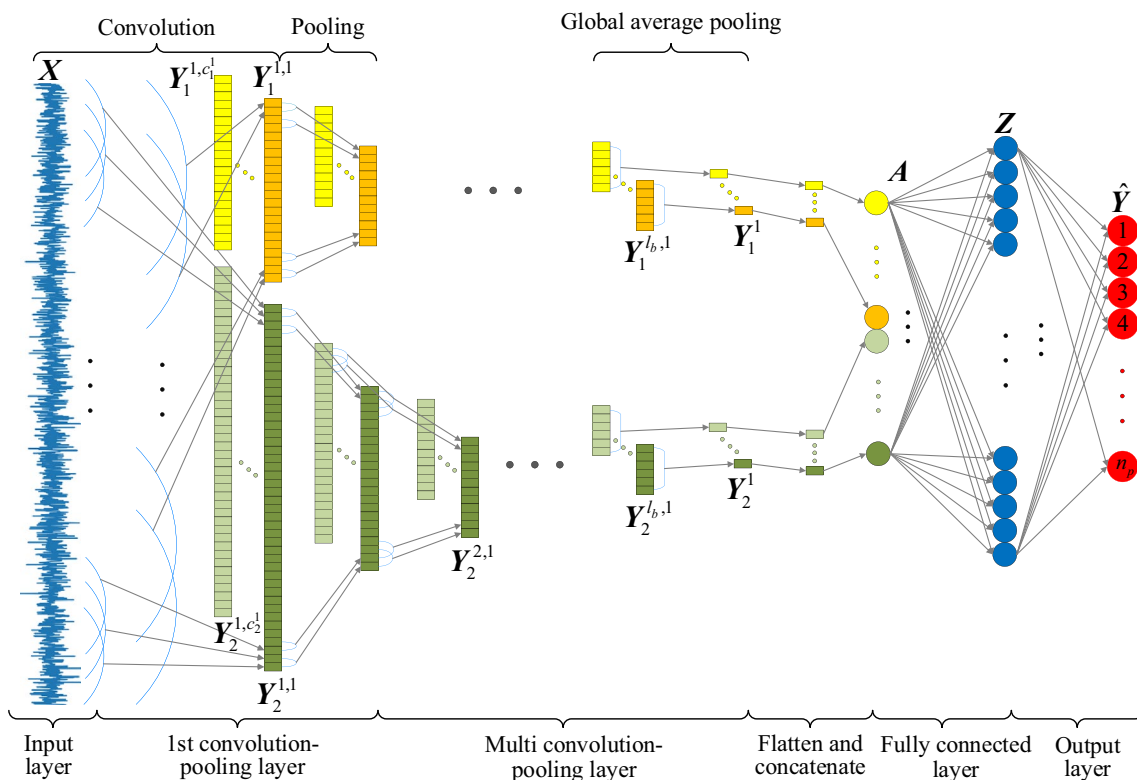


Figure 1 Framework of 2b-CNN

The MBCNN needs to learn the features that are irrelative to the phase shift of the signals, so the receptive field of the neuron which is input to the fully connected layer of the branch with the widest kernel (usually refers to the first branch), which is supposed to be greater than the number of signals in one cycle. Suppose that the number of the measured signals in one cycle is L_c and the length of the total input signal is L (usually $L = (3 \sim 4)L_c$), then

$$L_c \leq R_1^0 \leq L. \tag{5}$$

The stride of the first layer in the first branch can be calculated as

$$L_c \leq S_1^1 \cdot (2^{l_1} \times 3 - 1) \leq L. \tag{6}$$

To obtain more time-scale information through fewer parameters, this model requires the strides of the first layer in other branches to meet the following condition:

$$S_b^1 = \frac{S_{b-1}^1}{2} \quad (b = 2, 3, \dots). \tag{7}$$

Each branch of the MBCNN is generated through multiple sets of convolution operations. The convolution kernel has the same function as the window function in short-time Fourier transform (STFT). Thus, the process of multi-branch generation can be regarded as an STFT where the window widths are different and the window function is automatically adjusted according to the training data. This process will cause information duplication in the multiple channels. To reduce the possibility of overfitting, the dropout layer is adopted to randomly set part of the data that are fed into the first convolutional layer as zero.

2.2 Algorithm of Convolution and Pooling for Each Branch

The kernels for different channels in the same layer of the same branch have the same width, but the weights of the kernels are different. A channel is obtained by sliding convolution with the same kernel, and the network parameters of the convolutional layer can be reduced by letting the convolution units at different positions share the same kernel. The convolution operation is expressed as

$$y_{b,j}^{l,c} = \sum_{c'=1}^{c_b^{l-1}} \left(K_b^{l,c'} \cdot \{X_b^{l,c'}\}_j \right) = \sum_{c'=1}^{c_b^{l-1}} \sum_{j'=1}^{W_b^l} \left(k_{b,j'}^{l,c'} \times x_{b,(P_j+j')}^{l,c'} \right), \tag{8}$$

and

$$X_b^{l+1,c} = Y_b^{l,c}, \tag{9}$$

where $K_b^{l,c'}$ denotes a kernel (the size of the matrix is $c_b^{l-1} \times W_b^l$), $k_{b,j'}^{l,c'}$ is the j' -th weight of this kernel; j denotes the position of the convolution units; $\{X_b^{l,c'}\}$ is

the j -th convolution unit, $P_j = S_b^l \times (j - 1)$ and $x_{b,(P_j+j')}^{l,c'}$ is the j' -th datum in this unit.

At the first convolutional layer (when $l = 1$), $c_b^0 = 1, c' = c$, and the input $X_b^{1,c}$ in Eq. (8) is the input of the model, that is $X = \{x_1, x_2, \dots, x_L\}^T$. So $X_j = \{x_{P_j+1}, \dots, x_{P_j+W_b^1}\}^T$. The multi-branch is generated by

$$Y_b^{1,c} = K_b^{1,c} \cdot \{X_1, \dots, X_m\}. \tag{10}$$

A convolution-pooling block is also used in each branch to extract high-level features, as shown in Figure 2. The relationship between the lengths of the convolutional output channels of each branch is

$$m_{b+1}^l = 2m_b^l, \tag{11}$$

where a latter branch will have one more convolution-pooling block than its previous branch in the MBCNN model.

The batch normalization layer can speed up the convergence while suppressing the over-fitting to a certain extent.

Flattening data for multi-branches will result in excessive parameters, slow training speed, and overfitting. To tackle this problem, the model replaces the last pooling layer with a global average pooling layer, which is expressed as

$$Y_b^c = \text{average}(Y_b^{l_b,c}) = \frac{1}{m_b^{l_b}} \sum_{j=1}^{m_b^{l_b}} y_{b,j}^{l_b,c}, \tag{12}$$

where Y_b^c is the output of the average pooling layer of the c -th channel in the b -th branch.

2.3 Fully Connected Classification

The features that have been extracted through convolution and pooling from Eq. (12) are connected as

$$A^{l-1} = Y_b^c |_{(c=1,2,\dots; b=1,2,\dots)}, \tag{13}$$

where A denotes the input of the fully connected layer, and l_f denotes the number of this layer. Then the activation function of SoftMax transforms the output neurons

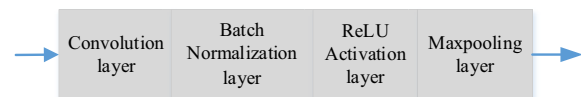


Figure 2 A convolution-pooling block

into a probability distribution with a sum of 1, and the fault type is obtained. Assume that the number of fault types is n_p , then the final output of the model is

$$\begin{aligned}\hat{y}_p &= \text{Softmax}(\mathbf{K}^{l_f} \cdot \mathbf{A}^{l_f-1} + \mathbf{B}^{l_f}) \\ &= \text{Softmax}(\mathbf{Z}^{l_f}) = \frac{e^{z_p}}{\sum_{p=1}^{n_p} e^{z_p}},\end{aligned}\quad (14)$$

where \hat{y}_p is the prediction labels of fault types; \mathbf{K} and \mathbf{B} are the parameters of fully connected layer; z_p is the logits of the p -th output neuron. The cross-entropy between the predicted label and the real label is taken as the loss function which is expressed as

$$\text{Loss} = - \sum_{p=1}^{n_p} (y_p \times \log \hat{y}_p), \quad y_p = \begin{cases} 1 & p, \\ 0 & \text{others,} \end{cases} \quad (15)$$

where y_p denotes the real labels of the fault type.

3 Experiments and Analysis

This study constructed an MBCNN with four branches (4b-CNN) for fault diagnosis of a rotating machine. The convolution stride of the first layer in the last branch (S_4^1) was initially set to 2. So, according to Eq. (7), the convolution stride of the first layer in the first branch (S_1^1) was 16, and the width of the convolution kernel (W_1^1) was 64. Substituting the number of the measured signals in one cycle (L_c) and S_1^1 into Eq. (6), it could be found that the number of layers of the first branch (l_1) is 4 or 5. The length of an input signal sample was set to 2048. Through Eq. (1), the output length was calculated, and it is 1 until the fifth convolution and pooling. Therefore, l_1 was set to 5. The number of channels was determined with the goal of reducing the loss, and it was set to 32 in the first branch.

The architecture and parameters of the 4b-CNN are shown in Figure 3.

In the following experiments, the programs were run on a computer equipped with an Intel i7-11700F 6-Core 2.50 GHz processor with 16 GB of RAM and an NVIDIA GeForce 3060 GPU. According to general parameter selection, the learning rate, the batch size, and the dropout ratio in the training were set to 0.0001, 512, and 0.1 separately. Meanwhile, a dataset was adopted to validate the model and compared it with other models based on signal processing, deep learning, CNN and MSCNN.

3.1 Fault Diagnosis for Bearing

The rolling bearing dataset [25] from Case Western Reserve University (CWRU) was used to compare the

results between this method and other methods. The fault types were the defects in rolling elements, outer ring and inner ring, and the diameters of the defects were 0.007, 0.014 and 0.021 inches. Thus, a total of nine fault types were detected in this dataset. To facilitate comparison, the data measured under load conditions of 0, 1, 2, 3 and 0–3 hp formed the sub-datasets *a*, *b*, *c*, *d*, and *e*, respectively. Similarly, 2048 pieces of data were encapsulated into a sample. Each of the sub-datasets included training samples, testing samples, and validation samples. In machine learning, if there is only training and testing, the data are generally divided at a ratio of 7:3. If there is validation and the amount of data is below ten thousand pieces, the data is generally divided at a ratio of 6:2:2. The fault labels and sample number in each sub-dataset are illustrated in Table 1.

Some researchers have applied different methods to fault diagnosis including the local feature-based gated recurrent unit (LFGRU) network [26], unsupervised feature learning (UFL) [27], energy-fluctuated multi-scale feature learning (EMFL) [28], semi-supervised distance-preserving SOM (SS-DPSOM) learning [29], and SVM optimized by inter-cluster distance (ICDSVM) [30]. Their results were compared with those of our MBCNN and the 1D-CNN, as listed in Table 2, and the structural parameters of the 1D-CNN are shown in Table 3. It can be seen that the classification accuracy of the MBCNN can reach 99.9% under a single load condition; under mixed load conditions, the MBCNN also achieved better performance, and the accuracy was even 4.17% and 2.07% higher than those of the shallow machine learning methods, respectively.

Further, the stabilities of MBCNN and 1D-CNN were compared, and the results are illustrated in Figure 4. Figure 4 shows that the MBCNN performed more stable than the 1D-CNN on fault diagnosis. This may be because the multi-branches in the MBCNN can obtain the robust features of different scales from the input signals, while the 1D-CNN only obtains the features on a single scale.

3.2 Fault Diagnosis for Gearbox

The gearbox dataset was collected from the Gearbox Dynamics Simulator (GDS) shown in Figure 5. The sampling frequency of the raw vibration signal was 12.8 kHz. It contains 15 different working conditions, where the speed range is 1600–2400 r/min with an interval of 400 r/min and the load range is 0–160 lb-in with an interval of 40 lb-in. Nine fault types for the bearing and gear are shown in Table 4. Thus, the fault diagnosis is a 10-type classification task (including a health type). The dataset

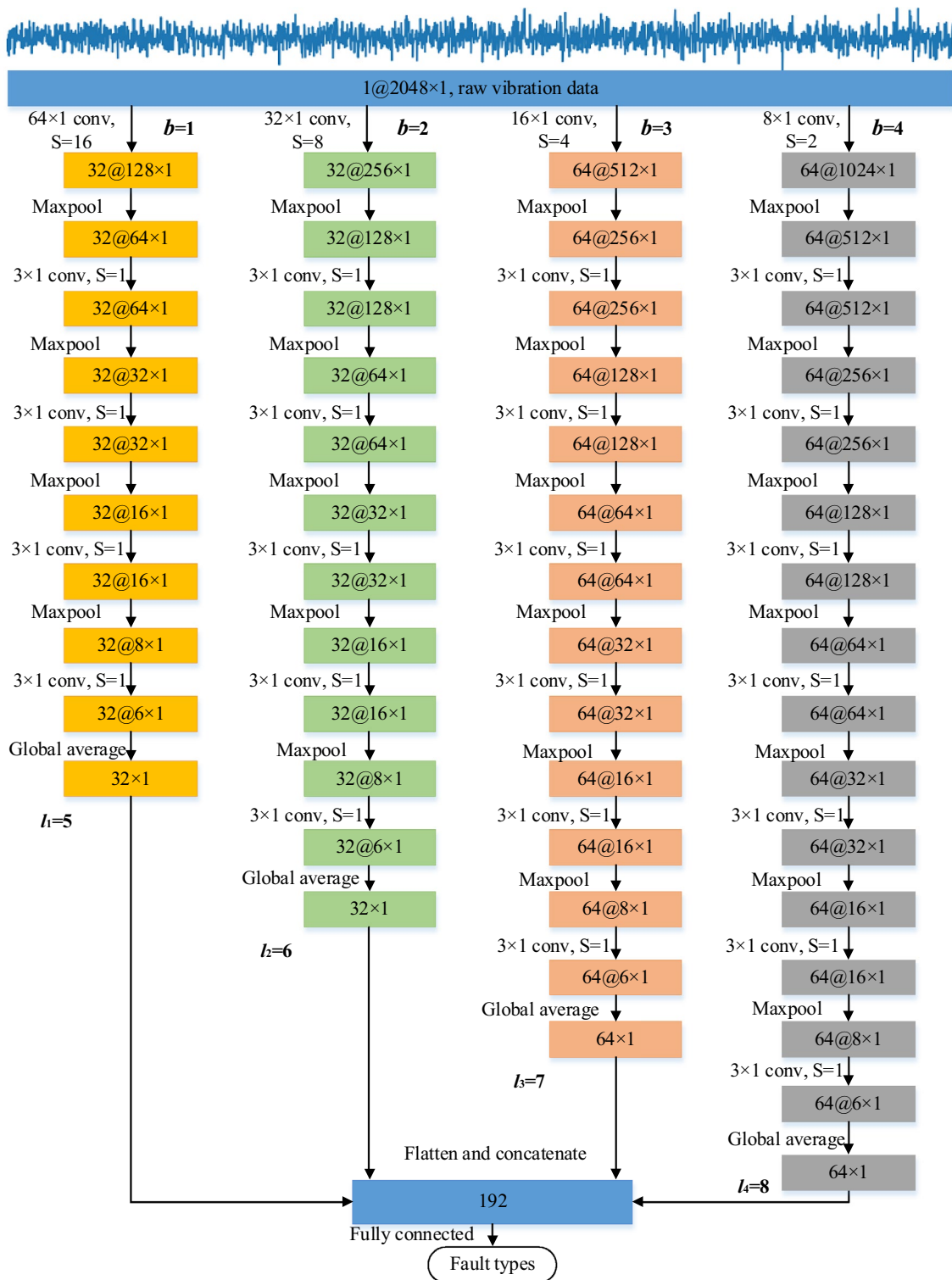


Figure 3 Architecture of a 4b-CNN

Table 1 Fault label and sample number in the bearing dataset

Fault	None	Rolling element fault			Inner ring fault			Outer ring fault			Load (hp)
		0.007	0.014	0.021	0.007	0.014	0.021	0.007	0.014	0.021	
Size (inch)	0	0.007	0.014	0.021	0.007	0.014	0.021	0.007	0.014	0.021	
Label	Normal	B ₀₀₇	B ₀₁₄	B ₀₂₁	IR ₀₀₇	IR ₀₁₄	IR ₀₂₁	OR ₀₀₇	OR ₀₁₄	OR ₀₂₁	
a	Training: testing: validation = 4500:1500:1500										0
b	Training: testing: validation = 4500:1500:1500										1
c	Training: testing: validation = 4500:1500:1500										2
d	Training: testing: validation = 4500:1500:1500										3
e	Training: testing: validation = 18000:6000:6000										0–3

Table 2 Classification accuracy of bearing fault using different methods (%)

Method	Sub-dataset				
	a	b	c	d	e
SS-DPSOM	–	–	–	–	95.8
ICDSVM	–	–	–	–	97.9
LFGRU	–	–	–	–	99.6
UFL	–	–	–	–	99.6
EMFL	98.8	98.8	99.4	99.4	99.8
1D-CNN	98.78	98.33	97.69	99.33	98.38
MBCNN	99.97	99.97	100	99.93	99.97

Table 3 Structural parameters of the 1D-CNN

Layer	Kernel / Stride	Output
Convolutional layer 1	6 × 1 × 8/2 × 1	1024 × 8
Pooling layer 1	2 × 1	512 × 8
Convolutional layer 2	6 × 1 × 16/2 × 1	256 × 16
Pooling layer 2	2 × 1	128 × 16
Convolutional layer 3	3 × 1 × 16/1 × 1	128 × 16
Pooling layer 3	2 × 1	64 × 16
Convolutional layer 4	3 × 1 × 32/1 × 1	64 × 32
Pooling layer 4	2 × 1	32 × 32
Convolutional layer 5	3 × 1 × 32/1 × 1	32 × 32
Pooling layer 5	2 × 1	16 × 32
Convolutional layer 6	3 × 1 × 64/1 × 1	16 × 64
Pooling layer 6	2 × 1	8 × 64
Convolutional layer 7	3 × 1 × 128/1 × 1	6 × 128
Pooling layer 7	2 × 1	3 × 128

has been kindly provided at <https://github.com/RuijunLiang/GDSdatasets>.

Since one-dimensional fault signals are sequent and periodic, segmenting them by an equal distance fails to represent their global information. To this end, this paper proposed overlapping segmentation by using sliding windows to refine the dataset, as shown in Figure 6. 2048

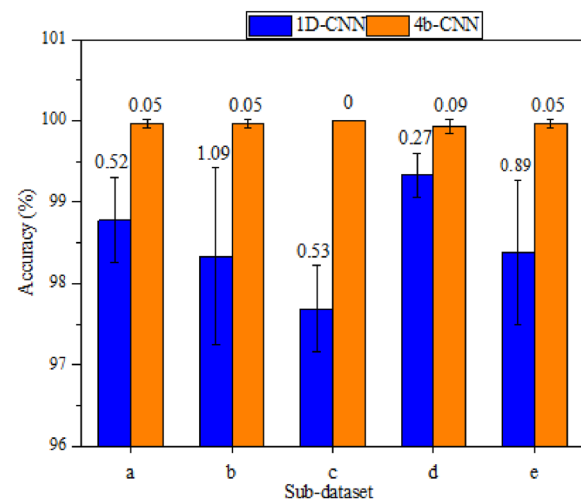


Figure 4 Stabilities of fault diagnosis by MBCNN and 1D-CNN

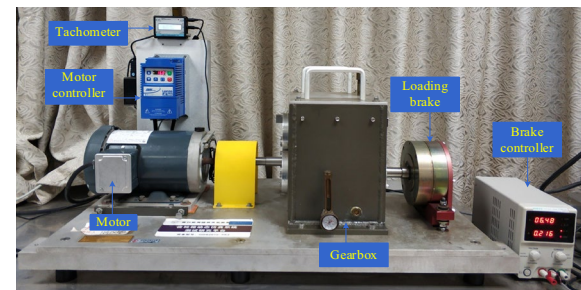


Figure 5 Gearbox dynamics simulator (GDS)

pieces of data were selected at every other 1000 pieces of data by a sliding window, and each group of the selected data was encapsulated as a sample. Then, 4500 samples were collected for each fault type including a health state, with 45000 samples in total. These samples were divided into a training dataset, a testing dataset and a validation dataset at a ratio of 6:2:2. The training dataset was under the load of 0, 80, and 160 lb-in, while the testing and validation datasets were under the load of 40, and 120 lb-in.

Table 4 Gear and bearing fault types

Fault type		Description
Gear	Miss	Missing one of feet in the gear
	Chipped	Crack occurs in the gear feet
	Root	Crack occurs in the root of gear feet
	Surface	Wear occurs in the surface of gear
	Eccentric	Geometric center of the gear does not coincide with rotating center of the shaft
Bearing	Ball	Crack occurs in the ball
	Inner	Crack occurs in the inner ring
	Outer	Crack occurs in the outer ring
	Combination	Crack occurs in the both inner and outer ring

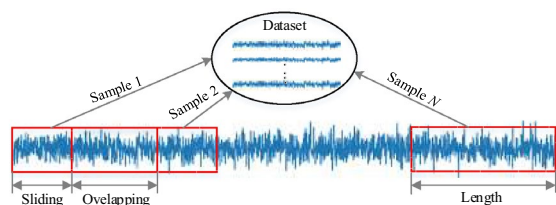


Figure 6 Dataset enhancement method

The fault label and the sample number are shown in Table 5.

To reduce the errors caused by the random selection of the initial weights, the training and the testing processes were repeated ten times, and the average classification accuracy of the multi-testing was taken as the result. The accuracy of the fault diagnosis by the 1D-CNN, and the MBCNN with various numbers of branches (2b-CNN, 3b-CNN, and 4b-CNN) is shown in Figure 7. Figure 8 displays the confusion matrixes in a certain test. From Figures 7 and 8, it can be seen that the MBCNN outperforms the 1D-CNN in accuracy by

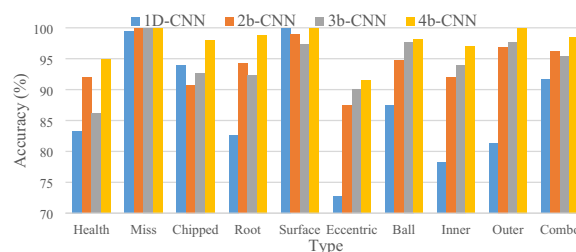


Figure 7 Classification accuracy of 1D-CNN and MBCNNs

about 14% at most. Meanwhile, the accuracy increases with the number of branches in the MBCNN, indicating that the more branches, the more multi-scale information of vibration signals can be learned.

Then, this study evaluated the training time, which was the average of over 50 epochs, and the testing time for the 1D-CNN and the MBCNNs, and the results are listed in Table 6. It can be seen that MBCNN requires more training time than 1D-CNN. This can be explained that with the increase in the number of branches, the model needs to learn more time scale information and introduce more parameters to be trained. However, when the number of branches was increased to 4, although the diagnostic accuracy was improved, the training time was significantly increased. Therefore, it is suggested to balance between time and accuracy for the fault diagnosis instead of blindly increasing the number of branches. Moreover, since the models are trained offline, the training time will not directly affect the diagnostic performance. Both the 2b-CNN and the 3b-CNN predict more accurately than the 1D-CNN, and they take nearly equal time (the difference is just 0.8 ms) to the 1D-CNN, which shows that the 2b-CNN and the 3b-CNN are applicable to diagnosis.

Table 5 Fault label and sample number in the gearbox dataset

Type		Sample number			Label	
		Training	Testing	Validation		
Health		2700	900	900	[1, 0, 0, 0, 0, 0, 0, 0, 0, 0]	
Gear	Miss	2700	900	900	[0, 1, 0, 0, 0, 0, 0, 0, 0, 0]	
	Chipped	2700	900	900	[0, 0, 1, 0, 0, 0, 0, 0, 0, 0]	
	Root	2700	900	900	[0, 0, 0, 1, 0, 0, 0, 0, 0, 0]	
	Surface	2700	900	900	[0, 0, 0, 0, 1, 0, 0, 0, 0, 0]	
	Eccentric	2700	900	900	[0, 0, 0, 0, 0, 1, 0, 0, 0, 0]	
Bearing	Ball	2700	900	900	[0, 0, 0, 0, 0, 0, 1, 0, 0, 0]	
	Inner	2700	900	900	[0, 0, 0, 0, 0, 0, 0, 1, 0, 0]	
	Outer	2700	900	900	[0, 0, 0, 0, 0, 0, 0, 0, 1, 0]	
	Combination		2700	900	900	[0, 0, 0, 0, 0, 0, 0, 0, 0, 1]

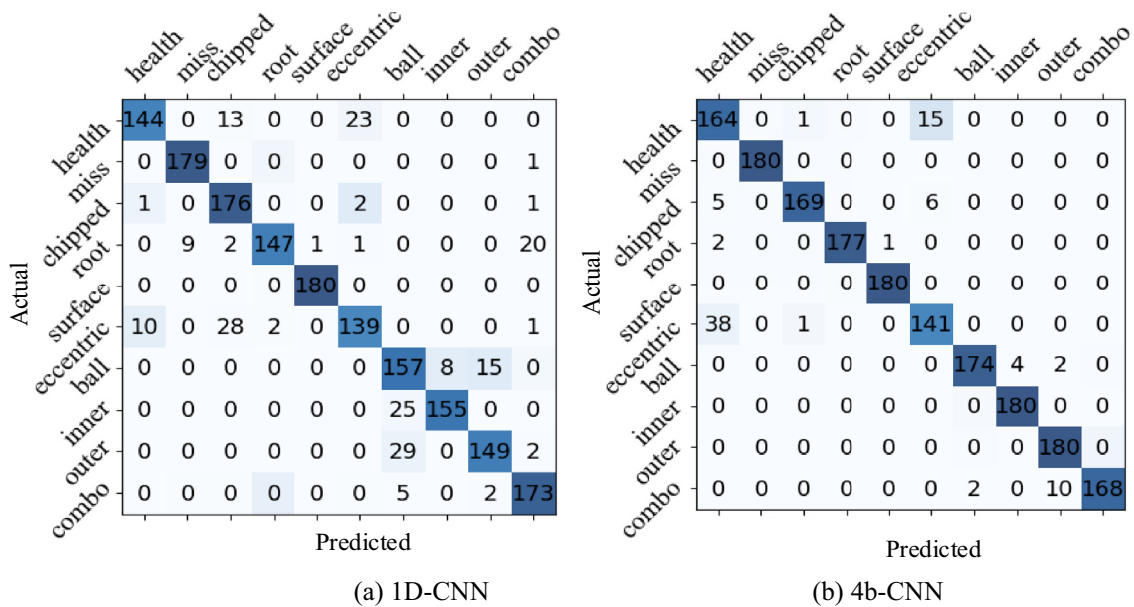


Figure 8 Confusion matrixes in a test

Table 6 Calculation time of 1D-CNN and MBCNNs

Model	Training time (s)	Validation time (ms)
1D-CNN	14.28	1.60
2b-CNN	21.75	2.44
3b-CNN	24.37	2.36
4b-CNN	47.75	3.16

4 Anti-noise Ability Test

4.1 Ensemble Learning Method

In Section 3, all data was collected in the lab. However, in practice, the environments are more complicated due to random noises. To further verify our model, an environment was simulated by adding Gaussian white noise to the bearing dataset. This study chose the sub-dataset *d* where the load was 3 hp, as described in Table 1, for testing. Figure 9a displays a raw signal sample in 10 types. Then, a noise disturbance with the signal-to-noise ratio (SNR) of -4 dB was added, and the maximum energy of the raw signal was adopted as a standard. Figure 9b shows the polluted, which look almost the same, so it is difficult to identify them and extract features.

To test the branch in MBCNN, this study used the same model parameters as those in each branch in 4b-CNN to build four 1D-CNN models, namely the 1st branch, the 2nd branch, the 3rd branch, and the 4th branch. The polluted signals with different SNRs are input in these four models, and the fault classification accuracy of the four models is listed in Table 7.

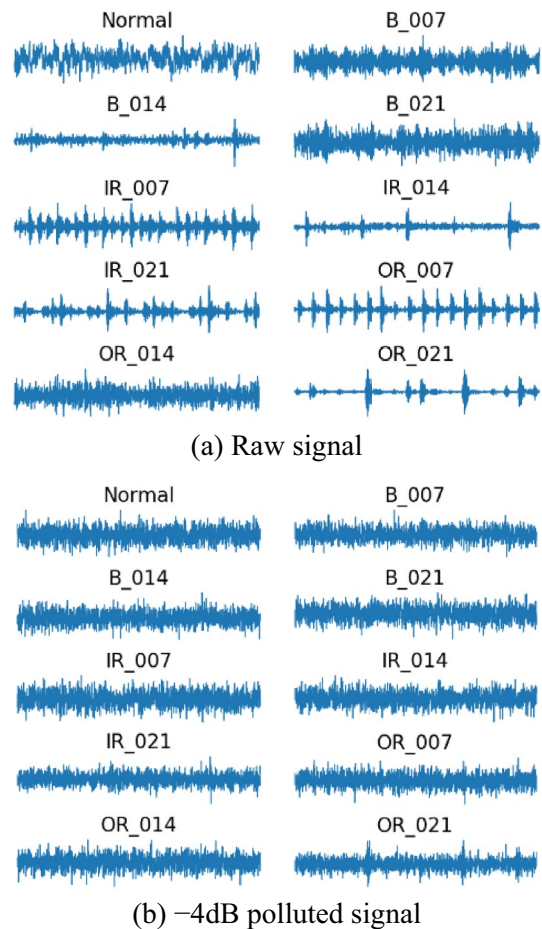


Figure 9 Raw signals and noise-polluted signals

Table 7 Classification accuracy of the different branch model (%)

Model	SNR (dB)							
	-4	-2	0	2	4	6	8	10
1st branch	80.9	81.7	82.7	85.9	86.7	91.4	96.6	98.5
2nd branch	80.9	83	84.4	85.3	87.1	92.9	97.7	98.7
3rd branch	81.5	83.6	85.2	87.2	90.5	94.7	96.2	98.7
4th branch	82.3	85.3	87.8	89.9	93.3	94.1	97.7	99.5

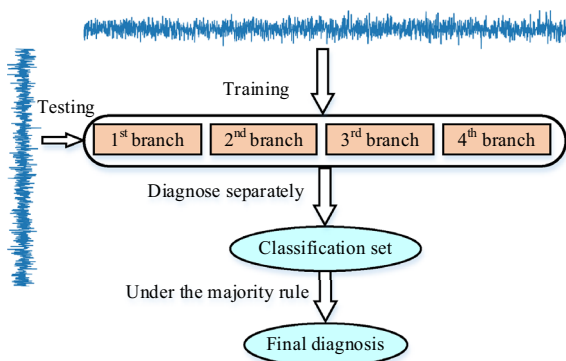


Figure 10 Ensemble learning process

It can be seen from Table 7 that all models could achieve high classification accuracy under a low noise disturbance; while all models obtain a low classification accuracy when they encounter severe noise disturbance. Thus, this study proposed an ensemble learning method to judge the fault mode from the ensemble of the four models' classification results, and it can be regarded as an embedding of 1D-CNN models in different input scales. The process of the proposed ensemble learning method is shown in Figure 10. The final diagnosis result is obtained according to the majority rule, i.e., firstly, the four models obtain four judgment results from a test sample respectively, and the final result is the majority among the four results. If the four results are different from each other or the four results are the same in pairs, the model that achieves the highest accuracy in the ensemble is preferred.

The four models and the ensemble learning method were applied to the polluted signals with different SNRs, and the results are presented in Figure 11. The ensemble learning method achieves the highest accuracy among all models, which may be because it can comprehensively judge the features learned by each model and can learn the information on different time scales

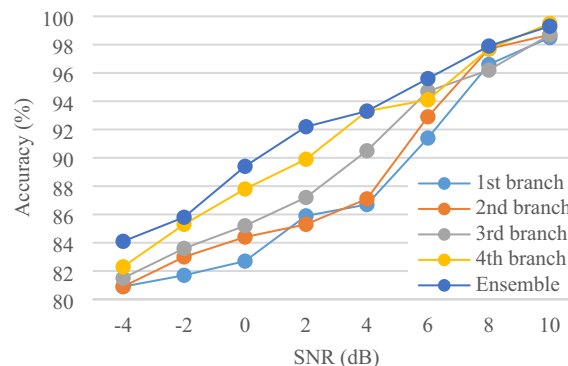


Figure 11 Classification accuracy to noise-polluted signals

4.2 MBCNN's Anti-noise Ability

The testing dataset of the polluted signals (as shown in Figure 9b) was fed into the MBCNN. The output features of the fully connected layer can be transformed into two dimensions by the t-SNE, and the visualized results are illustrated in Figure 12. It can be seen that increasing the number of branches can drive similar faults to be more concentrated while the distinct faults more separated, so the MBCNN has a strong capacity to extract features of signals, even signals with a low SNR. As the number of branches increases, the similar modes congregate closer, and the distinct modes separate further, which helps to better distinguish features of fault modes. Without any additional denoising, the MBCNN can adapt to the noise disturbance, demonstrating its potential for fault diagnosis in industrial fields.

Further, this study applied different methods including 1D-CNN and MBCNN with two, three and four branches, and the ensemble learning method to identify signals under multiple SNRs from - 4 dB to 10 dB. The results are shown in Figure 13. The superiorities of the MBCNN are obvious: (1) the classification accuracy of the MBCNNs with different numbers of branches is all significantly higher than that of the 1D-CNN under

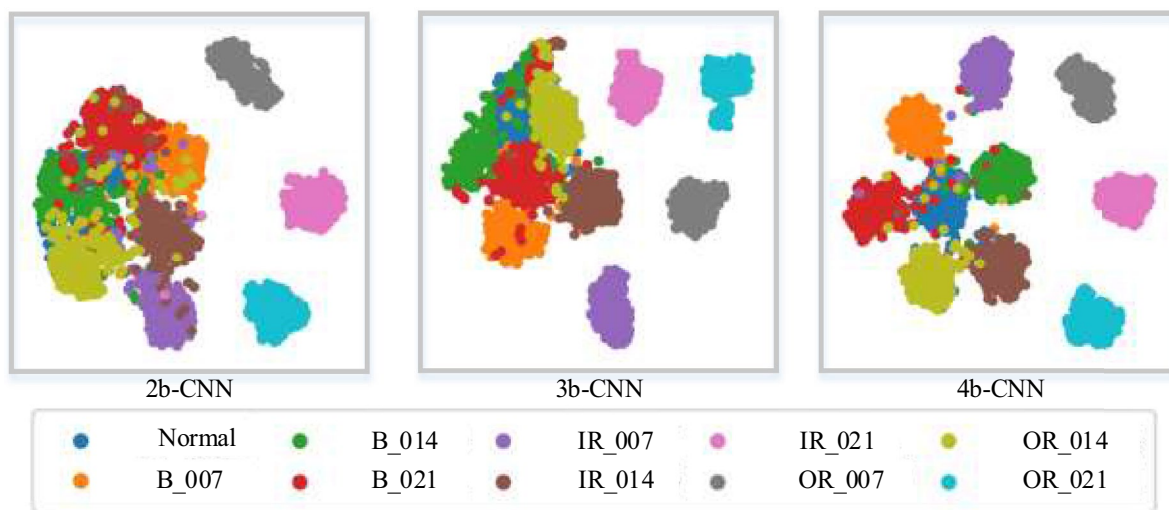


Figure 12 Visualization of the MBCNN's fault classification

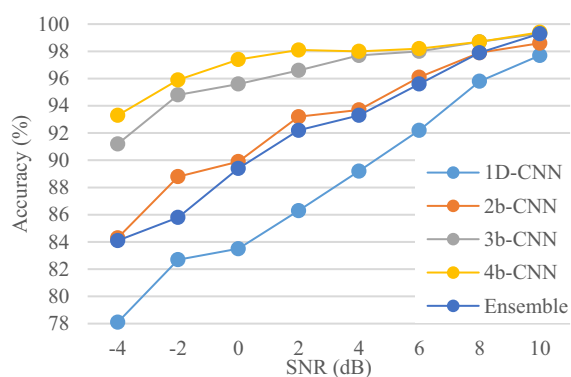


Figure 13 Classification accuracy to the noise polluted signals

different SNRs; (2) the classification accuracy of the ensemble 1D-CNN is similar to that of the 2b-CNN but lower than that of the 3b-CNN and 4b-CNN, which indicates that even the ensemble method using the features from multiple models cannot compete with the MBCNN with more branches; (3) the 3b-CNN and the 4b-CNN exhibit obvious superiority under low SNRs (−4 to 0 dB); (4) the classification accuracy of the 4b-CNN can also reach 93.3% even when the SNR is −4 dB.

The MBCNN obtained a higher classification accuracy than the ensemble learning method. This is because the ensemble learning method cannot automatically adjust the parameters of each model although it judges the learned features comprehensively and learns information on different time scales. The MBCNN not only integrates the feature information extracted by each branch but also adjusts the parameters of each branch according to the loss function and thus makes the features of branches complementary.

4.3 Comparison Between the Proposed MBCNN and the MSCNN

The proposed MBCNN shares a similar structure with the MSCNN. They both contain multiple branches that conduct a local convolution operation, and they utilize a fully connected layer to output the features of all the branches. However, there are some differences. First, the MBCNN extracts multi-scale information through different convolutional layers with different convolution kernels, while the MSCNN adopts multi-scale operations for simple down-sampling (i.e., certain features of the raw signals may be ignored). Second, in the MBCNN model, the number of layers in different branches is different, and more advanced features of the signal can be extracted. However, in the MSCNN model, all branches contain the same number of layers, thus missing high-level features.

To validate the superiority of the MBCNN model, the MBCNN with three and four branches (3b-CNN and 4b-CNN) was compared to the MSCNN with three and four scales (3S-CNN and 4S-CNN) [22] on the bearing failure dataset. The data were all polluted by noises with SNRs ranging from −4 dB to 10 dB, and the fault diagnosis accuracy shown in Figure 14 is the average result of ten times testing. It can be found that with the increase in the number of branches, more time-scale features are extracted, and the accuracy is improved. For the same branch, the classification accuracy of the MBCNN is higher than that of the MSCNN, indicating that the MBCNN can extract more features than the MSCNN. The MBCNN has obvious superiority at a low SNR (−4 dB to 2 dB), which indicates that the MBCNN can extract more high-level features from complex signals.

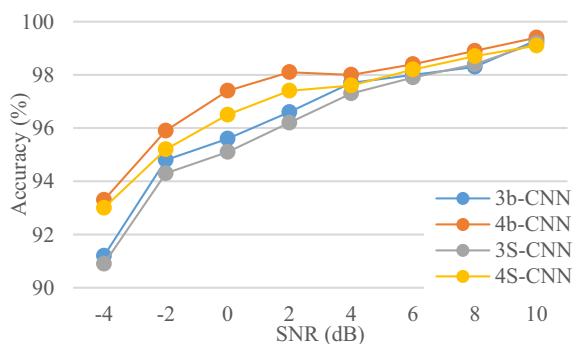
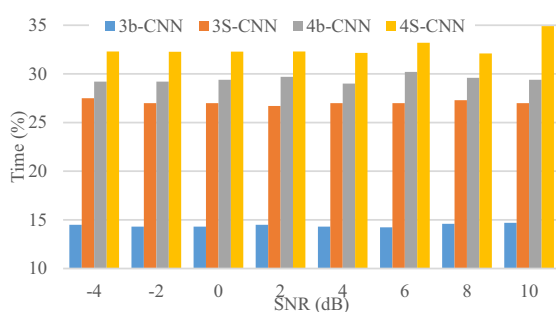
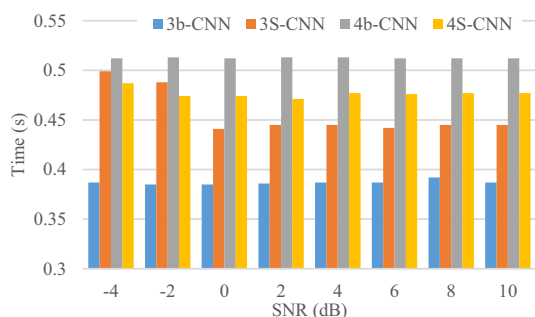


Figure 14 Fault diagnosis accuracy for noise-polluted signal



(a) Training time



(b) Testing time

Figure 15 Calculation time of the CNN models

Then, both the training time in one epoch and the testing time of one testing sample were evaluated using the above-described models including 3b-CNN, 4b-CNN, 3S-CNN and 4S-CNN, respectively. The results are given in Figure 15. For training speed, the MBCNN runs faster than the MSCNN (to the same number of branches in MBCNN with that of scales in MSCNN), because some branches of MBCNN have fewer layers. Meanwhile, the last pooling layer of each branch is a global average pooling layer, which can reduce the number of neurons to the fully connected layer. For

the testing process, both MBCNN and MSCNN could achieve fast fault diagnosis in the noise environment.

5 Conclusions

The paper proposes a CNN with multiple branches to identify complex multi-scale features of vibration signals of rotating machinery.

In terms of the structure of the MBCNN:

(1) The convolution kernels used in different branches have different widths and are related to each other, so the MBCNN can extract features containing the time scale of both a long-term and a short-term and reduce the number of model parameters.

(2) The last pooling layer of each branch is replaced with the global average pooling layer to avoid slow training speed and overfitting caused by too many model parameters.

Through experiments, the following conclusions are obtained:

(1) Comparing the ensemble learning method with the MBCNN, the former has a close classification accuracy with the 2b-CNN, but a lower classification accuracy than that of 3b-CNN and 4b-CNN. Even in low SNRs (from -4 to 0 dB), the diagnostic accuracy of the MBCNN is 7% higher. This is because the MBCNN can adjust the parameters of each branch according to the loss function and thus make the features of branches complementary, while the ensemble learning method comprehensively judges from the learned features, and the features have no information interaction.

(2) The visualization of the classification results demonstrates that increasing the number of branches in the MBCNN can make the similar faults more concentrated and the distinct faults more separated, so the features are more linearly separable, but more time is needed for training and testing in this case. The 2b-CNN and the 3b-CNN have a higher diagnostic accuracy than the 1D-CNN but they take nearly equal time (the difference is just 0.8 ms).

(3) Compared to the MSCNN, the MBCNN can not only learn more sufficient multi-scale information but also extract the higher-level features of signals from each branch. Therefore, the MBCNN can perform better on polluted signals with low SNRs (from -4 to 2 dB), and it is suitable for fault diagnosis in industries.

Acknowledgements

Not applicable.

Author contributions

RJL designed the model, analyzed the data, and wrote the original manuscript; WFR and YC developed the program and acquired the data; RPZ discussed the results. All authors read and approved the final manuscript.

Authors' Information

Ruijun Liang, born in 1974, is currently an associate professor at *College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, China*. She received her PhD from *Nanjing University of Aeronautics and Astronautics, China*, in 2007. Her research interests include intelligent manufacturing technology and equipment, intelligent detection and control. Wenfeng Ran, born in 1997, is currently a master candidate at *College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, China*.

Yao Chen, born in 1998, is currently a master candidate at *College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, China*.

Rupeng Zhu, born in 1959, is currently a professor at *College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, China*. He received his PhD degree from *Nanjing University of Aeronautics and Astronautics, China*, in 2000. His main research interest is mechanical transmission design theory and technology.

Funding

Supported by Transformation Program of Scientific and Technological Achievements of Jiangsu Province (Grant No. BA2022012) and National Natural Science Foundation of China (Grant No. 52205108)

Data availability

The datasets supporting the conclusions of this article are included within the article.

Declarations

Competing interests

The authors declare no competing financial interests.

Received: 20 January 2022 Revised: 18 October 2023 Accepted: 22 October 2023

Published online: 27 November 2023

References

- [1] Z Y Chen, K Gryllias, W H Li. Intelligent fault diagnosis for rotary machinery using transferable convolutional neural network. *IEEE Transactions on Industrial Informatics*, 2020, 16(1): 339-349.
- [2] N R Safin, V A Prakht, V A Dmitrievskii, et al. Stator current fault diagnosis of induction motor bearings based on the fast Fourier transform. *Russian Electrical Engineering*, 2016, 87(12): 661-665.
- [3] C Grover, N Turk. Rolling element bearing fault diagnosis using empirical mode decomposition and Hjorth parameters. *Procedia Computer Science*, 2020, 167: 1484-1494.
- [4] I Moumene, N Ouelaa. Application of the wavelets multiresolution analysis and the high-frequency resonance technique for gears and bearings faults diagnosis. *The International Journal of Advanced Manufacturing Technology*, 2016, 83(8): 1315-1339.
- [5] Y S Wang, N N Liu, H Guo, et al. An engine-fault-diagnosis system based on sound intensity analysis and wavelet packet pre-processing neural network. *Engineering Applications of Artificial Intelligence*, 2020, 94: 103765.
- [6] D Strömbergsson, P Marklund, K Berglund, et al. Bearing monitoring in the wind turbine drivetrain: A comparative study of the FFT and wavelet transforms. *Wind Energy*, 2020, 23(6): 1381-1393.
- [7] S Samanta, J N Bera, G Sarkar. KNN based fault diagnosis system for induction motor. *2016 2nd International Conference on Control, Instrumentation, Energy & Communication (CIEC)*, Kolkata, India, January 28-30, 2016: 304-308.
- [8] K Zhou, J Tang. Harnessing fuzzy neural network for gear fault diagnosis with limited data labels. *The International Journal of Advanced Manufacturing Technology*, 2021, 115(4): 1005-1019.
- [9] S Souahlia, K Bacha, A Chaari. MLP neural network-based decision for power transformers fault diagnosis using an improved combination of Rogers and Doernenburg ratios DGA. *International Journal of Electrical Power & Energy Systems*, 2012, 43(1): 1346-1353.
- [10] R Medina, J C Macancela, P Lucero, et al. Vibration signal analysis using symbolic dynamics for gearbox fault diagnosis. *The International Journal of Advanced Manufacturing Technology*, 2019, 104: 2195-2214.
- [11] W F Gong, H Chen, D W Wang. Fast fault diagnosis method of marine rotating machinery with multi-sensor monitoring based on improved LSTM-SVM. *Journal of Ship Mechanics*, 2021, 25(9): 1239-1250. (in Chinese)
- [12] Z S Yang, C Z Kong, X Rong, et al. Fault diagnosis of mine asynchronous motor based on EEMD energy entropy and ANN. *Micromotors*, 2021, 54(8): 23-27, 61. (in Chinese)
- [13] A Krizhevsky, I Sutskever, G E Hinton. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 2017, 60(6): 84-90.
- [14] Z Y Chen, A Mauricio, W H Li, et al. A deep learning method for bearing fault diagnosis based on cyclic spectral coherence and convolutional neural networks. *Mechanical Systems and Signal Processing*, 2020, 140: 106683.
- [15] M T Pham, J M Kim, C H Kim. 2D CNN-based multi-output diagnosis for compound bearing faults under variable rotational speeds. *Machines*, 2021, 9(9): 199.
- [16] S Kiranyaz, O Avci, O Abdeljaber, et al. 1D convolutional neural networks and applications: A survey. *Mechanical Systems and Signal Processing*, 2021, 151: 107398.
- [17] K Yan, X K Zhou. Chiller faults detection and diagnosis with sensor network and adaptive 1D CNN. *Digital Communications and Networks*, 2022, 8(4): 531-539.
- [18] T Ince, S Kiranyaz, L Eren, et al. Real-time motor fault detection by 1-D convolutional neural networks. *IEEE Transactions on Industrial Electronics*, 2016, 63(11): 7067-7075.
- [19] F R Wang, G B Song. A novel percussion-based method for multi-bolt looseness detection using one-dimensional memory augmented convolutional long short-term memory networks. *Mechanical Systems and Signal Processing*, 2021, 161: 107955.
- [20] W Y Huang, J S Cheng, Y Yang, et al. An improved deep convolutional neural network with multi-scale information for bearing fault diagnosis. *Neurocomputing*, 2019, 359: 77-92.
- [21] G Q Jiang, H B He, J Yan, et al. Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox. *IEEE Transactions on Industrial Electronics*, 2019, 66(4): 3196-3207.
- [22] T L Gu, Z H Sun, C Z Bin, et al. Fault diagnosis of rolling bearing based on multi-scale convolutional neural network. *Machinery Design & Manufacture*, 2022(5): 20-23. (in Chinese)
- [23] H P Zhao, Z W Mao, J J Zhang, et al. Multi-branch convolutional neural networks with integrated cross-entropy for fault diagnosis in diesel engines. *Measurement Science and Technology*, 2021, 32(4): 045103.
- [24] J Q Zhang, B M Xu, Z Y Wang, et al. An FSK-MBCNN based method for compound fault diagnosis in wind turbine gearboxes. *Measurement*, 2021, 172: 108933.
- [25] Case Western Reserve University bearing data center. Seeded fault test data. 2008[2020-04-08]. <https://engineering.case.edu/bearingdatacenter/download-data-file>.
- [26] R Zhao, D Z Wang, R Q Yan, et al. Machine health monitoring using local feature-based gated recurrent unit networks. *IEEE Transactions on Industrial Electronics*, 2018, 65(2): 1539-1548.
- [27] Y G Lei, F Jia, J Lin, et al. An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data. *IEEE Transactions on Industrial Electronics*, 2016, 63(5): 3137-3147.
- [28] X X Ding, Q B He. Energy-fluctuated multiscale feature learning with deep ConvNet for intelligent spindle bearing fault diagnosis. *IEEE Transactions on Instrumentation and Measurement*, 2017, 66(8): 1926-1935.
- [29] W H Li, S H Zhang, G L He. Semisupervised distance-preserving self-organizing map for machine-defect detection and classification. *IEEE Transactions on Instrumentation and Measurement*, 2013, 62(5): 869-879.
- [30] X Y Zhang, Y T Liang, J Z Zhou, et al. A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM. *Measurement*, 2015, 69: 164-179.