SCIENCE CHINA
Technological Sciences

•Article•

November 2022 Vol.65 No.11: 2551–2563
https://doi.org/10.1007/s11431-022-2109-4

A multi-scale convolutional neural network for bearing compound fault diagnosis under various noise conditions

JIN YanRui, QIN ChengJin*, ZHANG ZhiNan, TAO JianFeng & LIU ChengLiang

School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

Received April 22, 2022; accepted June 10, 2022; published online September 15, 2022

Recently, with the urgent demand for data-driven approaches in practical industrial scenarios, the deep learning diagnosis model in noise environments has attracted increasing attention. However, the existing research has two limitations: (1) the complex and changeable environmental noise, which cannot ensure the high-performance diagnosis of the model in different noise domains and (2) the possibility of multiple faults occurring simultaneously, which brings challenges to the model diagnosis. This paper presents a novel anti-noise multi-scale convolutional neural network (AM-CNN) for solving the issue of compound fault diagnosis under different intensity noises. First, we propose a residual pre-processing block according to the principle of noise superposition to process the input information and present the residual loss to construct a new loss function. Additionally, considering the strong coupling of input information, we design a multi-scale convolution block to realize multi-scale feature extraction for enhancing the proposed model's robustness and effectiveness. Finally, a multi-label classifier is utilized to simultaneously distinguish multiple bearing faults. The proposed AM-CNN is verified under our collected compound fault dataset. On average, AM-CNN improves 39.93% accuracy and 25.84% F1-macro under the no-noise working condition and 45.67% accuracy and 27.72% F1-macro under different intensity noise working conditions compared with the existing methods. Furthermore, the experimental results show that AM-CNN can achieve good cross-domain performance with 100% accuracy and 100% F1-macro. Thus, AM-CNN has the potential to be an accurate and stable fault diagnosis tool.

anti-noise, residual pre-processing block, bearing compound fault, multi-label classifier, multi-scale convolution feature extraction

Citation: Jin Y R, Qin C J, Zhang Z N, et al. A multi-scale convolutional neural network for bearing compound fault diagnosis under various noise conditions. Sci China Tech Sci, 2022, 65: 2551–2563, https://doi.org/10.1007/s11431-022-2109-4

1 Introduction

As important industrial equipment, rotating machinery is widely used in manufacturing, transportation, metallurgy, aerospace, and other fields [1–4]. The rolling bearing is one of the most critical components in rotating machinery. Rolling friction occurs from the sliding friction between its shaft seat and running shaft, which is a kind of precision mechanical element to reduce friction loss. The rolling bearing diagnosis is important in achieving high-quality and low-cost maintenance of industrial equipment [5]. In the industrial environment, the fault types of rolling bearing mainly include single fault and compound fault. A "compound fault" refers to two or more faults occurring simultaneously during operation. Due to the complex and changeable working conditions of practical application environments, a compound fault is more common and more harmful [6,7]. Therefore, compared with a single fault diagnosis, the accurate location of the compound fault is more challenging. For this reason, it has attracted more attention from researchers in recent years.

Generally, signal processing and artificial intelligence (AI) methods are the two main schemes to solve the compound fault diagnosis. Hilbert-Huang transform [8], empirical mode

*Corresponding author (email: qinchengjin@sjtu.edu.cn)

decomposition [9–13], wavelet transform [14–16], sparse decomposition [17], and wavelet packet transform have been proposed as signal processing methods to address the problem of detecting compound fault under practical applications.

AI methods contain traditional machine learning (ML) approaches and deep learning (DL) approaches for diagnosing compound faults. Feature extraction and feature classification are two stages of traditional ML approaches that require numerous expert knowledge. Varying manual features often lead to different diagnostic performances. Liu et al. [18] extracted a time-frequency dictionary and used the support vector machine approach to classify them. Li et al. [19] detected compound fault from locomotive bearings based on the distance preserving projection. Guo et al. [20] combined frequency, time, and time-frequency domain features for monitoring bearing conditions based on a deep neural network. Shao et al. [21] proposed a loss function to build an autoencoder and used a fish swarm algorithm to classify bearing faults. With the development of AI methods, DL methods are now applied to fault diagnosis, which can automatically extract numerous representations without extra manual knowledge. Chen et al. [22] extracted time-frequency and statistical measurement features from input signals and built a convolutional neural network (CNN) to achieve good diagnosis performance. Huang et al. [23] utilized different scale filters to acquire more useful inputs and designed a multi-scale cascade CNN for classifying input classification. Further, Guo et al. [24] presented an adaptative CNN-based model with a hierarchical learning rate to determine bearing faults and achieve good performance. Cheng et al. [25] used continuous wavelet transform and CNN for intelligent diagnosing faults of rotating machines.

The above results show that DL methods can achieve better results in bearing fault diagnosis than the traditional ML methods. However, due to the complex and changeable actual working scenarios, the bearing vibration signal can easily be polluted by noise, which makes feature extraction a great challenge. Meanwhile, unlike the single fault, it is more difficult to accurately extract and locate compound fault features, which brings great difficulty to the task of fault classification. Thus, this paper presents a novel anti-noise multi-scale convolutional neural network (AM-CNN) for overcoming the challenges of compound fault diagnosis under different levels of intensity noise. On the one hand, due to the different intensity of noises in actual working conditions, the vibration signal of a rolling bearing can be seriously polluted. DL methods cannot easily extract meaningful deep features from vibration signals combined with noise. Therefore, to process the input information, a residual pre-processing block is proposed according to the principle of noise superposition. Meanwhile, the residual loss is presented to construct a new loss function. On the other hand, the vibration signal is usually nonlinear and characterized by uncertainty, coupling, and high complexity. Additionally, the characteristic frequency caused by compound fault changes greatly, thus leading to the distribution of fault characteristics on different scales. Thus, the vibration signal has multi-scale characteristics and contains complex characteristics at different time scales. This paper designs a multi-scale convolution block to realize multi-scale feature extraction. Multiple convolution layers with different branches and convolution kernel sizes are utilized to extract different time scales features, thereby enhancing the robustness of the network model to the learning of compound fault features. Finally, a multi-label classifier is used to simultaneously distinguish multiple bearing faults for replacing the Softmax classifier.

The main contributions of our proposed method are summarized as follows.

(1) By combining CNN with the idea of residual learning, this paper proposes a residual pre-processing block to effectively extract denoise information under noisy working conditions. Then, we design a novel loss function to update block parameters during backpropagation, aiming at acquiring clean input from different intensity noises.

(2) With the aim of improving the model performance, a multi-scale convolution block is applied based on the idea of multi-scale learning to learn multi-scale characteristics from vibration signals.

(3) The proposed AM-CNN is an end-to-end intelligent diagnosis approach for obtaining domain-invariant features, which is not only suitable for compound fault diagnosis but also has good cross-domain capability in different noisy environments.

The remainder of this paper is arranged as follows. Sect. 2 describes the details of our experimental bearing compound fault dataset. The methodology and details are introduced in Sect. 3. Sect. 4 presents the experimental design and results. Finally, the conclusions are presented in Sect. 5.

2 The experimental dataset

For validating the generalization performance and effectiveness of AM-CNN, this paper designs a bearing compound fault test to build a bearing compound fault dataset. Then, bearing vibration signals are sampled at 50 kHz, and the amount of each fault is 512 samples.

Moreover, the dataset used in this paper includes normal state, three single faults and four compound faults. And then, single faults include inner ring fault, outer ring fault, and roller fault of bearing data from the test rig. Further, four compound faults include outer ring with inner ring fault, outer ring with roller fault, inner ring with roller fault, and outer ring with inner ring and roller fault.

Further, Table 1 presents the details of the dataset, and Figure 1 lists the vibration signals with different signal faults and compound faults.

3 Methodology

3.1 Overview of AM-CNN

Figure 2 presents an overview of the AM-CNN. First, due to the different intensity noises in actual working conditions, we acquire bearing vibration signals from the dataset acquisition system. Then, we use a short-time Fourier transform (STFT) to obtain time-frequency information as inputs of the AM-CNN. Additionally, inspired by the ideas of convolution layer and residual learning, we design the proposed feature extractor for extracting deep features automatically. Finally, we use a multi-label classifier to detect multiple bearing faults simultaneously.

3.2 Pre-processing

Similar to a previous article [26], we segment the vibration signal into segments with 5120 sampling points. Then, inspired by the idea of the article, we use STFT to vibration segments for yielding input information. Figure 3 shows an example of the time-frequency input and its source vibration signal.

Table 1 Details of the proposed bearing fault dataset

Type	Label	Amount
Normal	N	512
Outer fault	Out	512
Inner fault	In	512
Roller fault	R	512
Outer with inner fault	OI	512
Outer with roller fault	OR	512
Inner with roller fault	IR	512
Outer with inner and roller fault	OIR	512

3.3 Residual pre-processing block

With the development of DL technologies, many novel processing approaches have been proposed, among which CNN is a common type. A previous study [27] first proposed CNN, and Lecun et al. [28] further developed it. Due to its strong representative ability, CNN has been applied to industrial equipment operation and maintenance [29–31], handwriting classification [32], medical diagnosis [33,34], and fault diagnosis [35,36].

The main purpose of CNN is to represent input information automatically. Eq. (1) shows the principle of the convolution operation.

$$y_i = \mathrm{Relu}\left(\sum_{j=i}^{i+k} w_j x_j + b\right), \tag{1}$$

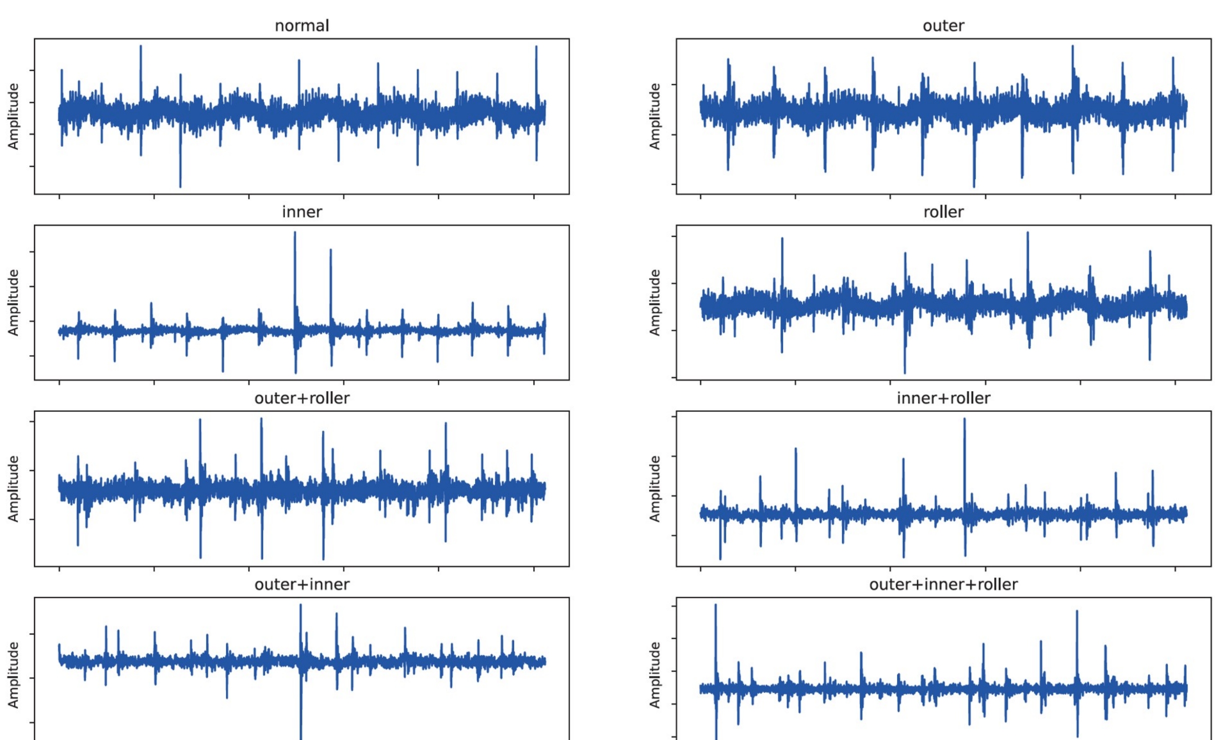

Figure 1 (Color online) Vibration signals with different signal faults and compound faults.

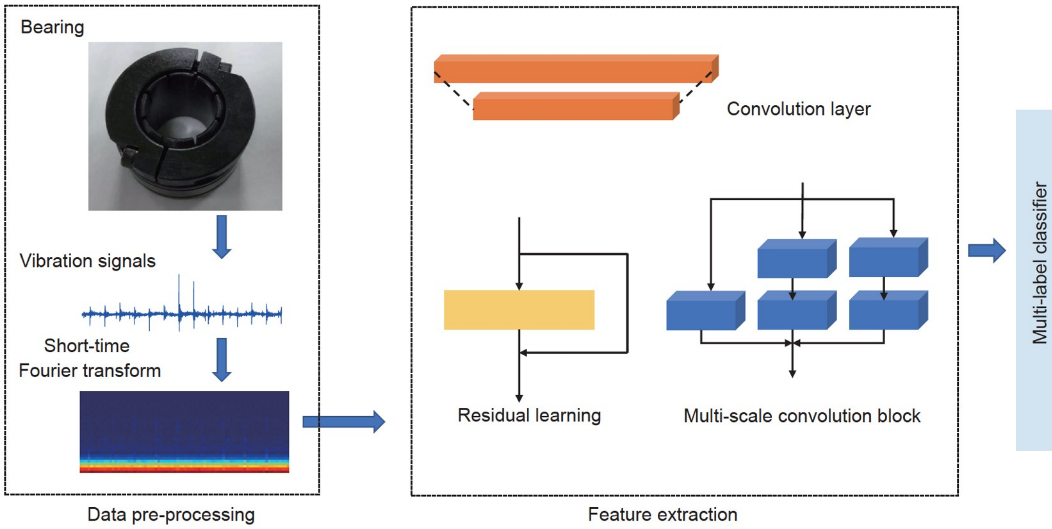

Figure 2 (Color online) Overview of the AM-CNN.

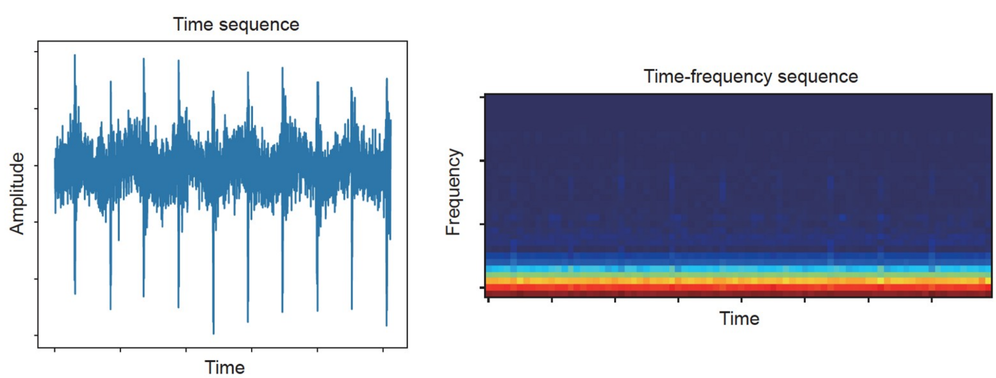

Figure 3 (Color online) The time-frequency input and its source vibration signal.

where Relu(·) means ReLu activation function, and k represents the stride and convolution kernel size.

With numerous training input signals, CNN can automatically obtain representative features to achieve end-to-end learning. Thus, in this paper, we use the convolutional layer as the basic feature extraction method for designing subsequent blocks.

Due to the complexity of actual working scenarios, the collected vibration signals are typically accompanied by numerous background noises. Hence, the achievement of noise reduction of input information is very important. Traditional denoising methods, such as the sparse model, the gradient model, and the Markov random field model, mainly establish an a priori knowledge model for input information. These methods have the following two disadvantages: (1) the involvement of complex optimization problems in the test phase, which is time-consuming, and (2) the nonconvex nature of the model and involvement of several manually selected parameters. Thus, to overcome the abovementioned disadvantages, we propose the residual pre-processing block. We consider that the collected vibration signal consists of two parts, namely, actual vibration signal and noise, as shown in eq. (2):

$$y = x + \text{noise}, \tag{2}$$

where y and x refer to the collected vibration signal and the actual vibration signal, respectively. By training a residual mapping function $F(y) \approx$ noise with residual learning, and then $x = y - F(y)$, we can input information after denoising. Figure 4 presents the proposed residual pre-processing block inspired by a previous article [37]. The residual pre-processing block does not directly output the denoised image but designs the proposed block as a prediction residual information, that is, the difference between noise observation and potentially clean information. In other words, the proposed block implicitly removes the potential cleaning information through the operation in the hidden layer. Furthermore, to enable the block to conduct end-to-end

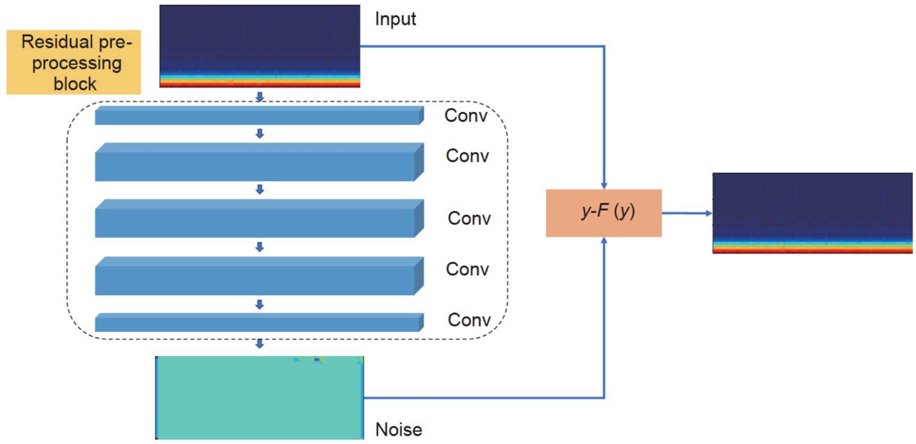

Figure 4 (Color online) The structure of the residual pre-processing block.

training with subsequent networks, we design residual loss to ensure that the module parameters are updated iteratively in the backpropagation process:

$$\text{loss} = \frac{1}{n}\sum \left|\text{output}_i - v_i\right|^2, \tag{3}$$

where v represents the present noise template.

The contributions of the proposed residual pre-processing block are as follows: (1) an end-to-end CNN is proposed, which uses residual learning to remove clean information from noisy information, and (2) the proposed residual pre-processing block can handle ordinary information denoising tasks, also known as "blind denoising".

3.4 Multi-scale convolution block

In combining different time scale features from vibration signals, the multi-scale convolution block (MS-block) skillfully adjusts convolution kernel sizes; hence, the convolution layer can learn the characteristics of different time scales. Specifically, the MS-block utilizes multiple parallel convolution layers with different convolution kernel sizes, learns rich features of different scales simultaneously, and fuses multi-scale features across channels through convolution operation. Figure 5 presents the structure of the MS-block. As shown in Figure 5, we can see that the MS-block has multiple branches to extract multi-scale features from inputs in parallel. The output representation of the convolution layer is shown in eq. (1), in which the convolution kernel sizes of different branches vary. For different convolution kernels, convolution can extract different scales of information from the original signal. Thereafter, the convolution outputs of different branches are fused to obtain complementary information of different time scales y_{out} = MS-block (input). The MS-block extracts deep features from different scales. To comprehensively utilize different levels of features, we fuse different scales features to obtain comprehensive features.

A previous article [38] showed that the low-level features of the shallow learning signal and the high-level features of the DL signal can enrich the features obtained by simply superimposing the convolution layer. For understanding the high-level fault features, a deep multi-scale convolutional network composed of multiple MS-block is constructed in this paper.

3.5 Multi-label classifier

Compared with a signal fault, a compound fault means more than two faults may occur simultaneously, which brings more difficulties. Commonly, the Softmax classifier is used for diagnosing singling faults, which are respectively listed in eqs. (4) and (5). However, because only one classification result can be selected, the Softmax classifier is not suitable for compound fault diagnosis.

$$P_i = \frac{e^{z_i}}{\sum_j e^{z_j}}, \tag{4}$$

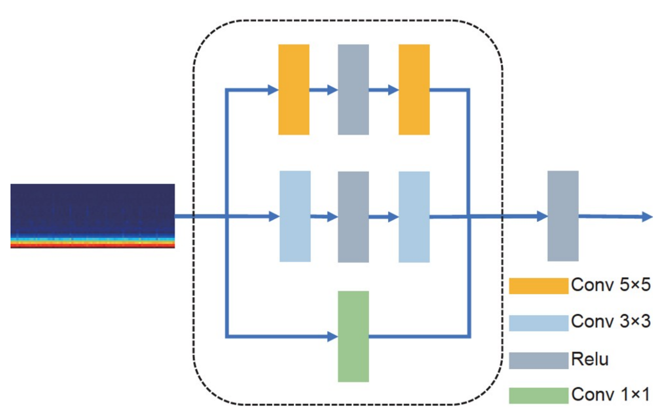

Figure 5 (Color online) The structure of the multi-scale convolution block.

$$y = \underset{i}{\operatorname{argmax}}(P_i). \tag{5}$$

As for compound fault diagnosis, we build a multi-label classifier based on the sigmoid function for calculating each fault type probability. The details of the multi-label classifier are shown in eqs. (6) and (7).

Eq. (6) represents the computing process of the sigmoid function, which is used for calculating each fault probability occurring. Then, each fault probability is restricted by the sigmoid function occurring from 0 to 1. In this paper, we set the threshold at 0.5, which means that when a fault type's probability of occurrence exceeds 0.5, the model considers that there is a corresponding fault in this section of the signal; otherwise, there is no fault. Eq. (7) is used to obtain the output after processing input by the multi-label classifier, which can locate multiple fault types in vibration signals.

$$\tilde{P}_i = \frac{1}{1+e^{-z_i}}, \tag{6}$$

$$\tilde{y}_i = \begin{cases} 1, \text{if } \tilde{P}_i \geq \text{threshold}, \\ 0, \text{if } \tilde{P}_i < \text{threshold}. \end{cases} \tag{7}$$

Thus, to effectively distinguish multiple fault types, the multi-label classifier is utilized to calculate the probability of each fault type and detect the compound fault.

3.6 The architecture

As shown in Figure 6, the architecture of the proposed AM-CNN consists of three main parts: residual pre-processing block, multi-scale convolution block, and the convolutional multi-label classifier. The details of the AM-CNN are listed in Table 2.

As shown in Figure 6, we use a residual pre-processing block for input information cleaning to improve the model's anti-noise performance. The denoised information is fed into the subsequent deep multi-scale convolutional network to extract multi-scale features for representing fault features. These meaningful features are required to classify compound faults. The convolutional multi-label classifier includes convolutional layers, max-pooling layers, and a multi-label classifier. Due to the large feature dimension of features, convolution layers and max-pooling layers are mainly used to extract and filter information until the final judgment information is obtained. The final judgment information is fed into the multi-label classifier for distinguishing compound faults. Algorithm 1 lists the training process for the AM-CNN.

This paper utilizes the Adam optimizer [39] for updating network parameters during the backpropagation process and sets the learning rate as 0.001. Further, we use two NVIDIA GTX 1080 GPUs and an Intel Xeon E5-2620 CPU for training the proposed method. We also set the batch size to 128 during the backpropagation process.

4 Results and verification

We design different experiments to validate the accuracy,

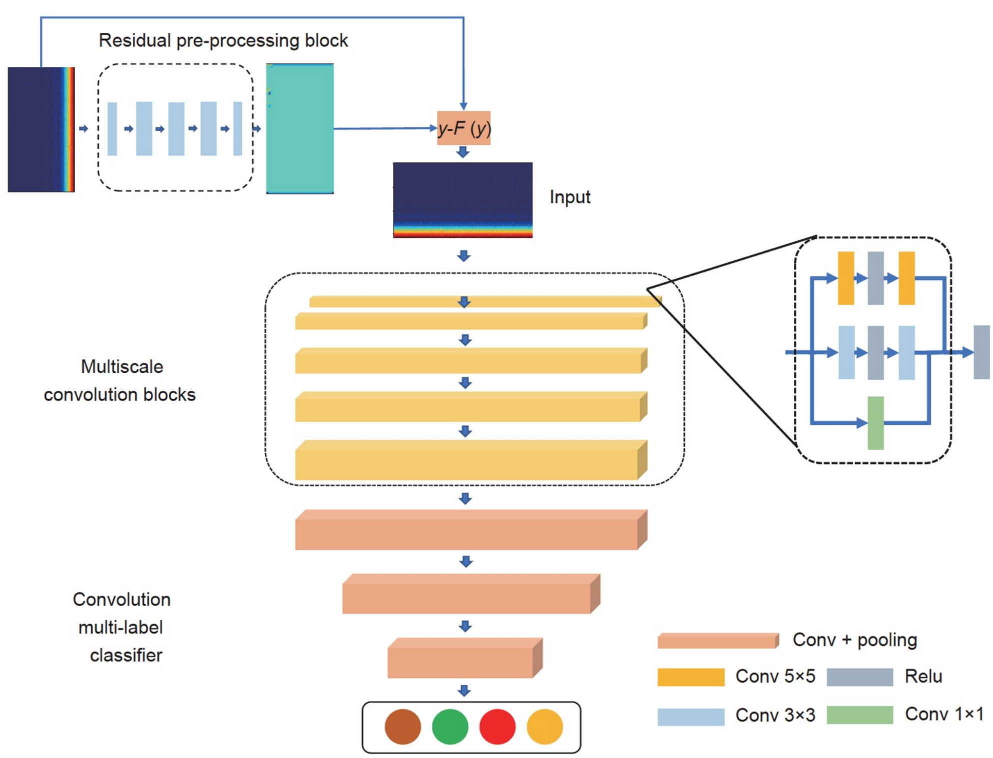

Figure 6 (Color online) The architecture of the proposed AM-CNN.

Table 2 The introduction of feature extractor details of the AM-CNN

Layers	Type	Kernel size	Stride	Output size
0–1	Input	–	–	80 × 32 × 1
1–2	Convolution	3 × 3	1 × 1	80 × 32 × 64
2–3	Convolution	3 × 3	1 × 1	80 × 32 × 64
3–4	Convolution	3 × 3	1 × 1	80 × 32 × 64
4–5	Convolution	3 × 3	1 × 1	80 × 32 × 64
5–6	Convolution	3 × 3	1 × 1	80 × 32 × 1
6–7	MS-block 1	5 × 5; 3 × 3; 1 × 1	1 × 1	80 × 32 × 8
7–8	MS-block 2	5 × 5; 3 × 3; 1 × 1	1 × 1	80 × 32 × 16
8–9	MS-block 3	5 × 5; 3 × 3; 1 × 1	1 × 1	80 × 32 × 32
9–10	MS-block 4	5 ×5; 3 × 3; 1 × 1	1 × 1	80 × 32 × 64
10–11	MS-block 5	5 × 5; 3 × 3; 1 × 1	1 × 1	80 × 32 × 128
11–12	Convolution	3 × 3	1 × 1	80 × 32 × 256
12–13	Max-pooling	2 × 2	2 × 2	40 × 16 × 256
13–14	Convolution	3 × 3	1 × 1	40 × 16 × 256
14–15	Max-pooling	2 × 2	2 × 2	20 × 8 × 256
15–16	Convolution	3 × 3	1 × 1	20 × 8 × 256
16–17	Average-Pooling	20 × 8	1 × 1	1 × 1 × 256

Algorithm 1

Input
X_S: vibration signals from the training set
Y_S: labels from the training set
N: total training epochs
t: number of epoch
s: batch size
Start
Repeat
  For from X_S and X_T, respectively obtain s samples
 Deep features are obtained by the feature extractor of AM-CNN.
 Calculate the probability of each fault occurring according to eq. (6).
 Calculate the multi-label loss by the binary cross-entropy function.
Update the multi-label loss and residual loss based on backpropagation.
  end
Until: t increases to N
Output: Bearing compound fault diagnosis model

effectiveness, and anti-noise performance of the proposed AM-CNN model. First, we introduce the existing methods and compare them with the proposed AM-CNN under no-noise working conditions. Then, we further compare the existing methods with the proposed AM-CNN under different intensity noise working conditions. Furthermore, we train the model with no-noise training data and verify the model under different intensity noise data to validate the model's cross-domain performance.

To conduct a comprehensive evaluation of the model superiority and model anti-noise performance, this paper utilizes accuracy (Acc) and macro F1 score (F1-macro), as shown below:

$$Ppr = \frac{TP}{TP+FP}, \tag{8}$$

$$\mathrm{Sen} = \frac{TP}{TP+FN}, \tag{9}$$

$$\mathrm{F1} = \frac{2\times \mathrm{Sen}\times Ppr}{\mathrm{Sen}+Ppr}, \tag{10}$$

$$\text{F1-macro} = \frac{1}{N}\sum_{i=1}^{N} F1_i, \tag{11}$$

$$\mathrm{Acc} = \frac{TP+TN}{TP+TN+FP+FN}. \tag{12}$$

4.1 Model anti-noise performance

This experiment is designed to validate the anti-noise performance of the proposed AM-CNN. Thus, we add different intensity noises to the original vibration signals based on eq. (13). This paper adds 10, 8, 6, 4, 2, −2, −4, −6, −8, and −10 dB noises to the original vibration signals to generate ten different noise datasets, respectively. For each noise dataset, we select 80% of the whole data as training data to guarantee the model performance during the training process. Additionally, we select 20% of the whole data as validating data in choosing the best model during the training process. Finally, the remainder of the whole dataset is selected as the test data for validating the model performance.

$$\mathrm{SNR} = \log_{10}\left(\frac{P_s}{P_n}\right). \tag{13}$$

For effectively comparing model performance, we choose some existing methods as the baseline models, which are shown in Figure 7.

(1) CNN. CNN uses a convolution layer for replacing the multi-scale convolution block and extracting numerous deep features automatically. Similar to the AM-CNN, we use the convolution multi-label classifier to classify features and detect compound faults.

(2) LSTM. This method uses the long-short term memory (LSTM) layer for replacing multi-scale convolution blocks and extracting numerous deep features automatically. Similar to the AM-CNN, we use the multi-label classifier to classify the features and detect compound faults.

(3) Zou's method [40] with one-dimension input (ANCNN_1d). According to the article, we reproduce the methodology. Then, we use the vibration signal to replace the time-domain input used in the paper as the input information of the network to ensure consistency with the original paper. Finally, we use the multi-label classifier to classify features and detect compound faults.

(4) Zou's method [40] with time-frequency input (ANCNN_tf). According to the article, we reproduce the methodology. Then, we use the time-domain input used in the paper as the input information of the network to evaluate the model performance of Zou's method. Finally, we use the multi-label classifier to classify the features and detect the compound faults.

(5) Jin's method [41] with one-dimension input (ANNN_1d). According to the article, we reproduce the methodology. Then, we also use the vibration signal to replace the time-domain input used in the paper as the input information of the network to ensure consistency with the original paper. Finally, the multi-label classifier is used to classify the features and detect compound faults.

(6) Yuan's method [42] (CNN_1d). According to the article, we reproduce the methodology and retain the proposed CNN structure. Finally, the multi-label classifier is used to classify the features and detect compound faults.

(7) Chen's method [43] (MSCNNLSTM). According to the article, we reproduce the methodology and retain the combination of CNN and LSTM used in this paper. Finally, the multi-label classifier is used to classify the features and detect compound faults.

Figure 8 lists the comparison results of the baseline models and that of the proposed AM-CNN performance under different intensity noises. On average, CNN reaches 99.46% accuracy and 99.78% F1-macro under different intensity noise working conditions. Furthermore, LSTM reaches an average of 99.22% accuracy and 99.82% F1-macro under different intensity noise working conditions. Meanwhile, ANCNN_1d reaches 13.16% accuracy and 33.54% F1-macro under noise working conditions of varying intensity. ANCNN_tf reaches 0% accuracy and 39.15% F1-macro under noise working conditions of varying intensities. Similarly, ANNN_1d reaches 12.59% accuracy and 49.53% F1-macro, also under noise working conditions of varying intensities. CNN_1d reaches 90.56% accuracy and 93.03% F1-macro under noise working conditions of varying intensities. Finally, MSCNNLSTM reaches 65.31% accuracy and 91.08% F1-macro under noise working conditions of varying intensities. Overall, the proposed AM-CNN reaches better model performance, with 100% accuracy and 100%

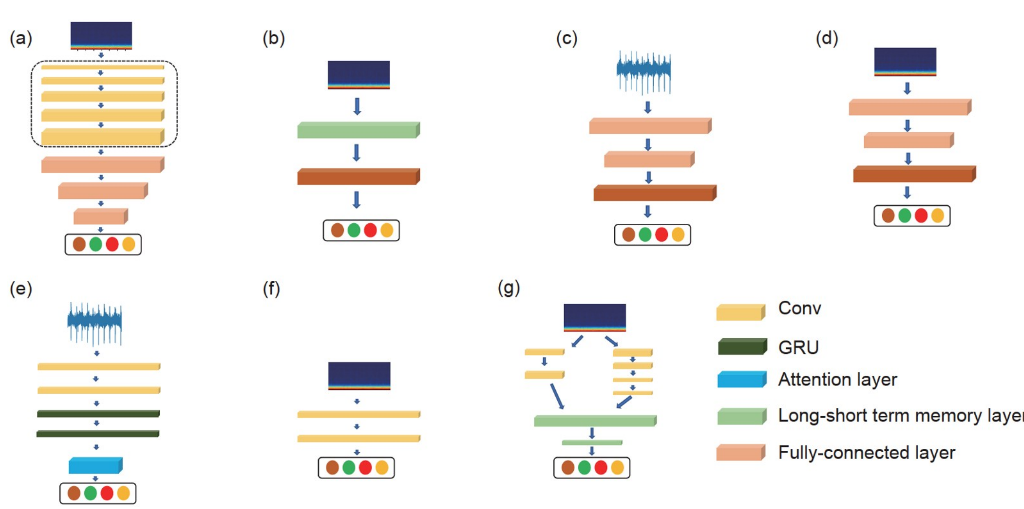

Figure 7 (Color online) The architecture of baseline models. (a) CNN, (b) LSTM, (c) ANCNN_1d, (d) ANCNN_tf, (e) ANNN_1d, (f) CNN_1d, and (g) MSCNNLSTM.

Figure 8 (Color online) The anti-noise performance of Experiment 1.

F1-macro under noise working conditions of varying intensities.

According to the abovementioned experimental results, we can arrive at the following conclusions.

(1) The proposed AM-CNN has better anti-noise ability compared with the baseline models based on CNN, LSTM, and the combination of CNN and LSTM.

(2) Compared with convolution layers, the MS-block can extract multi-scale features, which can better reflect the fault features of compound faults and improve the final diagnosis results.

(3) Due to the limitation of the LSTM layer model structure, the corresponding operations must be completed during feature extraction, which in turn leads to a model that requires a long period of time to complete. Under actual working conditions, the model with shorter time consumption has a more real-time nature. The MS-block can also extract multi-scale features, which can achieve better model performance. Therefore, our method has better advantages in model anti-interference and real-time performances.

(4) Under actual working conditions, data will continue to be generated, and it is necessary to judge whether a fault exists. Therefore, compared with CNN and LSTM, the AM-CNN improves the anti-noise performance of the model and can correctly diagnose more vibration signals in the actual diagnosis process.

(5) Given that the characteristics of compound faults are difficult to locate, the existing methods cannot easily diagnose compound faults, which reflects the advantages of the proposed AM-CNN.

4.2 Model cross-domain performance

This experiment is designed to validate the cross-domain performance of the AM-CNN. First, we add different intensity noises to the remainder of the original vibration signals based on eq. (13). This paper adds 10, 8, 6, 4, 2, −2, −4, −6, −8, and −10 dB noises to the original vibration signals to generate ten different noise datasets, respectively. Table 3 lists the details of this experiment. To guarantee good model performance during the training process, we select 80% of the whole D_i data as the training data. Then, we select 20% of the whole D_i data as validating data in order to choose the best model during the training process. Finally, we select 20% of the whole D_j data as the test data for validating the model performance.

Figure 9 shows the cross-domain performance of the proposed AM-CNN under different intensity noises. As we can see, the proposed AM-CNN reaches better model performance, with 99.94% accuracy and 99.97% F1-macro under different tasks, thus indicating 55.01% accuracy and 35.60% F1-macro. The comparison results reveal that compared with the existing methods, AM-CNN can extract domain-invariant features during the training process, which are helpful in improving cross-domain performance.

Table 3 Details of Experiment 2

Task	Training data/Validating data (dB)	Test data (dB)
1	No_noise	−10/−8/−6/−4/−2/2/4/6/8/10
2	−10	−8/−6/−4/−2/2/4/6/8/10/No_noise
3	−8	−10/−6/−4/−2/2/4/6/8/10/No_noise
4	−6	−10/−8/−4/−2/2/4/6/8/10/No_noise
5	−4	−10/−8/−6/−2/2/4/6/8/10/No_noise
6	−2	−10/−8/−6/−4/2/4/6/8/10/No_noise
7	2	−10/−8/−6/−4/−2/4/6/8/10/No_noise
8	4	−10/−8/−6/−4/−2/2/6/8/10/No_noise
9	6	−10/−8/−6/−4/−2/2/4/8/10/No_noise
10	8	−10/−8/−6/−4/−2/2/4/6/10/No_noise
11	10	−10/−8/−6/−4/−2/2/4/6/8/No_noise

As we can see in Figure 9, the proposed AM-CNN is stable under different intensity noise environments.

(1) Through the parameter training of the network model under one working condition, the trained model can still achieve good performance under other different intensity noise data. Compared with the laboratory environment, there will be a great deal of background noise in the actual working environment. The experimental results reveal that the proposed AM-CNN can be trained by easily obtaining no-noise data, thus achieving good results even in complex noise environments.

(2) Compared with the existing methods, the AM-CNN can reach better and more stable cross-domain performance, which means the proposed model can extract domain-invariant features in any noise environment. The reasonable architecture of the AM-CNN is helpful in applying the proposed method to practical working conditions.

(3) The AM-CNN can achieve good cross-domain performance without using other additional transferring methods, thus greatly simplifying the model structure and saving on model training time. In practical industrial applications, this can vastly improve the timeliness of diagnosis.

Thus, the proposed AM-CNN has a strong representative capability of denoising noise and extracting essential deep features for diagnosing compound faults.

4.3 Model performance comparisons

This experiment is designed to validate the effectiveness of the AM-CNN under the no-noise working condition. To ensure the model performance during the training process, we select 80% of the whole no-noise data as the training data. Additionally, we select 20% of the whole no-noise data as validating data in choosing the best model during the training process. Finally, the remainder of the whole dataset is se-

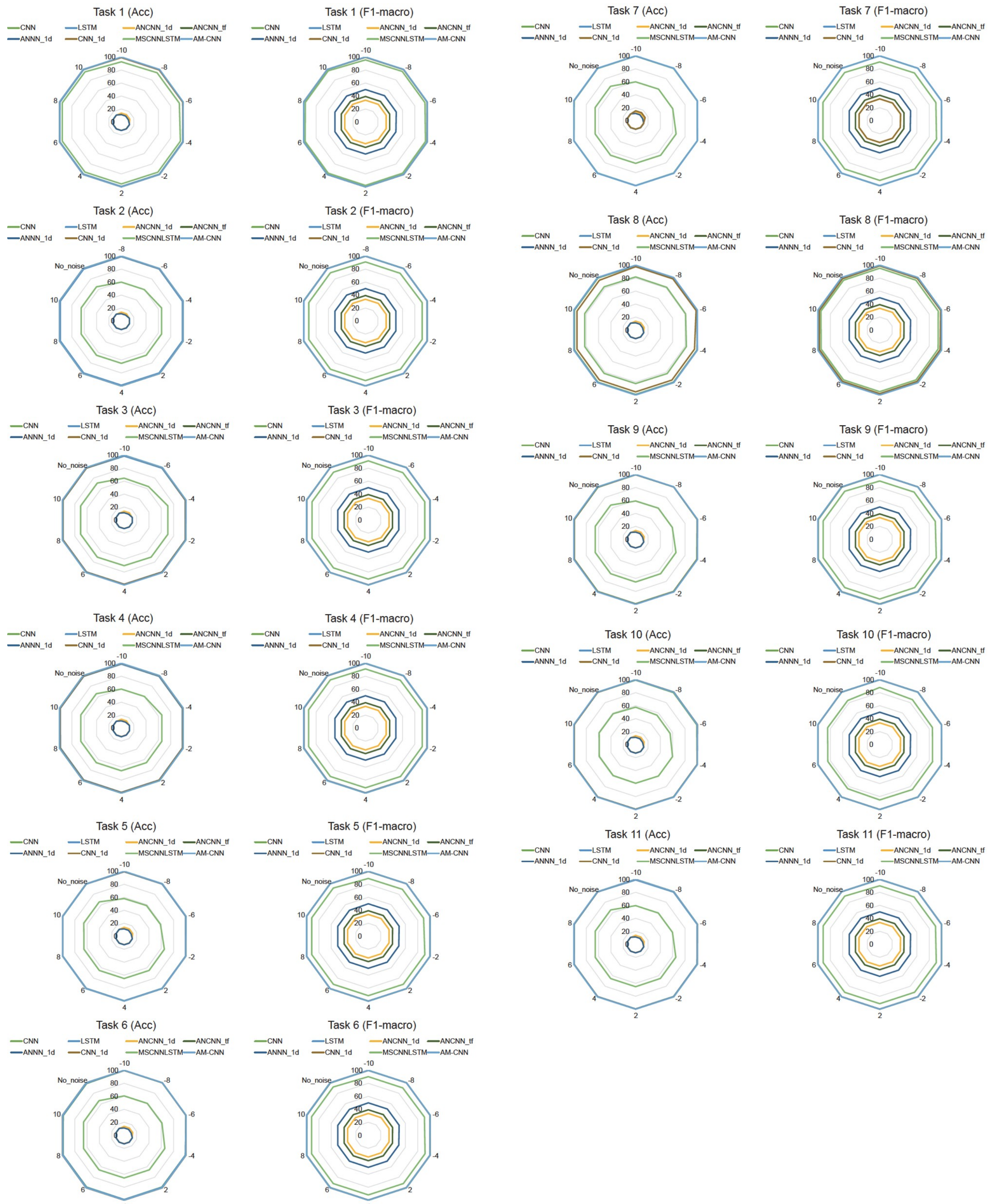

Figure 9 (Color online) The cross-domain performance of Experiment 2.

lected as the test data for validating the model performance.

Table 4 shows the comparison results. As listed in Table 4, under the no-noise working condition, CNN reaches 99.51% accuracy and 99.88% F1-macro, LSTM reaches 99.27% accuracy and 99.83% F1-macro, ANCNN_1d reaches 12.93% accuracy and 33.52% F1-macro, ANCNN_tf reaches 0% accuracy and 39.11% F1-macro, ANNN_1d reaches 13.17% accuracy and 49.35% F1-macro, CNN_1d reaches

Table 4 Results of Experiment 3

Algorithm	Accuracy (%)	F1-macro (%)
CNN	99.51	99.88
LSTM	99.27	99.83
ANCNN_1d	12.93	33.52
ANCNN_tf	0.0	39.11
ANNN_1d	13.17	49.35
CNN_1d	100	100
MSCNNLSTM	95.61	97.44
AM-CNN	100	100

100% accuracy and 100% F1-macro, and MSCNNLSTM reaches 95.61% accuracy and 97.44% F1-macro. Overall, the proposed AM-CNN reaches better model performance, with 100% accuracy and 100% F1-macro under the no-noise working condition.

According to the abovementioned experimental results, we can conclude the following points.

(1) Compared with the existing methods, the AM-CNN can reach better accuracy, which shows that the model has better expression ability.

(2) Compared with the convolution layers and the LSTM layers, the MS-block can extract multi-scale features, which can better reflect the fault features of compound faults to improve the final diagnosis results.

(3) Given that the characteristics of compound faults are difficult to locate, the existing methods cannot easily diagnose compound faults, thus reflecting the advantages of the AM-CNN.

4.4 Practical applications and future works

The experimental results prove that the AM-CNN reaches competitive performance on our collected compound fault dataset. Moreover, compared with the existing methods, the AM-CNN improves model performance under working conditions of varying noise intensities and achieves good cross-domain performance with 100% accuracy and 100% F1-macro. As for practical applications, the AM-CNN can be an accurate and stable fault diagnosis tool in the industrial environment to a certain extent.

As for future work, we will further explore a general structure for fault diagnosis using other structures, such as random convolutional kernels [44] and LSTM. We will also apply the proposed method to other fields, such as industrial equipment maintenance [45–47] and medical diagnosis [48–52].

5 Conclusion

This paper proposes the AM-CNN for addressing the issue of compound fault diagnosis under different intensity noises. First, according to the principle of noise superposition, we propose a residual pre-processing block to process the input information and present the residual loss to construct a new loss function. Additionally, considering the characteristic frequency caused by compound fault changes, we design a multi-scale convolution block to realize multi-scale feature extraction. Multiple convolution layers with different branches and convolution kernel sizes are utilized to extract different time scale features and enhance the robustness of the network model to the learning of compound fault features.

Finally, we use a multi-label classifier to distinguish among multiple bearing faults simultaneously. The proposed AM-CNN is verified under our collected compound fault dataset. The findings show that compared with the existing methods, AM-CNN improves accuracy by 39.93% and 45.67% with 25.84% and 27.72% F1-macro, respectively, under noise working conditions of varying intensities. Finally, the experimental results show that the AM-CNN can achieve good cross-domain performance with 100% accuracy and 100% F1-macro.

This work was supported by the National Key R&D Program of China (Grant No. 2020YFB1709604), the State Key Laboratory of Mechanical System and Vibration (Grant No. MSVZD202103), and the Shanghai Municipal Science and Technology Major Project (Grant No. 2021SHZDZX0102).

1 Di Z Y, Shao H D, Xiang J W. Ensemble deep transfer learning driven by multisensor signals for the fault diagnosis of bevel-gear cross-operation conditions. Sci China Tech Sci, 2021, 64: 481–492
2 Huang H R, Li K, Su W S, et al. An improved empirical wavelet transform method for rolling bearing fault diagnosis. Sci China Tech Sci, 2020, 63: 2231–2240
3 Xu Y G, Wang L, Hu A J, et al. Time-extracting S-transform algorithm and its application in rolling bearing fault diagnosis. Sci China Tech Sci, 2022, 65: 932–942
4 Wang J, Li S, An Z, et al. Batch-normalized deep neural networks for achieving fast intelligent fault diagnosis of machines. Neurocomputing, 2019, 329: 53–65
5 Cui L, Wu N, Ma C, et al. Quantitative fault analysis of roller bearings based on a novel matching pursuit method with a new step-impulse dictionary. Mech Syst Signal Process, 2016, 68-69: 34–43
6 Chen J, Zi Y, He Z, et al. Compound faults detection of rotating machinery using improved adaptive redundant lifting multiwavelet. Mech Syst Signal Process, 2013, 38: 36–54
7 Lyu X, Hu Z, Zhou H, et al. Application of improved MCKD method based on QGA in planetary gear compound fault diagnosis. Measurement, 2019, 139: 236–248
8 Peng Z K, Tse P W, Chu F L. A comparison study of improved Hilbert-Huang transform and wavelet transform: Application to fault diagnosis for rolling bearing. Mech Syst Signal Process, 2005, 19: 974–988
9 Guo J, Zhen D, Li H, et al. Fault feature extraction for rolling element bearing diagnosis based on a multi-stage noise reduction method. Measurement, 2019, 139: 226–235
10 Lei Y, Lin J, He Z, et al. A review on empirical mode decomposition in fault diagnosis of rotating machinery. Mech Syst Signal Process, 2013, 35: 108–126

11 Li C, Tao Y, Ao W, et al. Improving forecasting accuracy of daily enterprise electricity consumption using a random forest based on ensemble empirical mode decomposition. Energy, 2018, 165: 1220–1227
12 Yu X, Dong F, Ding E, et al. Rolling bearing fault diagnosis using modified LFDA and EMD With sensitive feature selection. IEEE Access, 2018, 6: 3715–3730
13 Huang D, Ke L, Mi B, et al. A new incipient fault diagnosis method combining improved RLS and LMD algorithm for rolling bearings with strong background noise. IEEE Access, 2018, 6: 26001–26010
14 Yan R, Gao R X, Chen X. Wavelets for fault diagnosis of rotary machines: A review with applications. Signal Process, 2014, 96: 1–15
15 Tao J, Qin C, Liu C. A synchroextracting-based method for early chatter identification of robotic drilling process. Int J Adv Manuf Technol, 2019, 100: 273–285
16 Tao J, Qin C, Li W, et al. Intelligent fault diagnosis of diesel engines via extreme gradient boosting and high-accuracy time-frequency information of vibration signals. Sensors, 2019, 19: 3280
17 Yang B, Liu R, Chen X. Fault diagnosis for a wind turbine generator bearing via sparse representation and shift-invariant K-SVD. IEEE Trans Ind Inf, 2017, 13: 1321–1331
18 Liu R, Yang B, Zhang X, et al. Time-frequency atoms-driven support vector machine method for bearings incipient fault diagnosis. Mech Syst Signal Process, 2016, 75: 345–370
19 Li W, Zhang S, Rakheja S. Feature denoising and nearest-farthest distance preserving projection for machine fault diagnosis. IEEE Trans Ind Inf, 2016, 12: 393–404
20 Guo L, Gao H, Huang H, et al. Multifeatures fusion and nonlinear dimension reduction for intelligent bearing condition monitoring. Shock Vib, 2016, 2016: 1–10
21 Shao H, Jiang H, Zhao H, et al. A novel deep autoencoder feature learning method for rotating machinery fault diagnosis. Mech Syst Signal Process, 2017, 95: 187–204
22 Chen Z Q, Li C, Sanchez R V. Gearbox fault identification and classification with convolutional neural networks. Shock Vib, 2015, 2015: 1–10
23 Huang W, Cheng J, Yang Y, et al. An improved deep convolutional neural network with multi-scale information for bearing fault diagnosis. Neurocomputing, 2019, 359: 77–92
24 Guo X, Chen L, Shen C. Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis. Measurement, 2016, 93: 490–502
25 Cheng Y, Lin M, Wu J, et al. Intelligent fault diagnosis of rotating machinery based on continuous wavelet transform-local binary convolutional neural network. Knowledge-Based Syst, 2021, 216: 106796
26 Jin Y, Qin C, Huang Y, et al. Actual bearing compound fault diagnosis based on active learning and decoupling attentional residual network. Measurement, 2021, 173: 108500
27 Fukushima K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol Cybernetics, 1980, 36: 193–202
28 Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition. Proc IEEE, 1998, 86: 2278–2324
29 Qin C, Jin Y, Tao J, et al. DTCNNMI: A deep twin convolutional neural networks with multi-domain inputs for strongly noisy diesel engine misfire detection. Measurement, 2021, 180: 109548
30 Jin Y, Qin C, Tao J, et al. An accurate and adaptative cutterhead torque prediction method for shield tunneling machines via adaptative residual long-short term memory network. Mech Syst Signal Process, 2022, 165: 108312
31 Qin C, Shi G, Tao J, et al. An adaptive hierarchical decomposition-based method for multi-step cutterhead torque forecast of shield machine. Mech Syst Signal Process, 2022, 175: 109148
32 Gan J, Wang W, Lu K. A new perspective: Recognizing online handwritten Chinese characters via 1-dimensional CNN. Inf Sci, 2019, 478: 375–390
33 Jin Y, Li Z, Qin C, et al. A novel interpretable method based on attentional deep neural network for actual ECG quality assessment. Biomed Signal Process Control, 2023, 79: 104064
34 Zhang J, Tian J, Cao Y, et al. Deep time-frequency representation and progressive decision fusion for ECG classification. Knowledge-Based Syst, 2020, 190: 105402
35 He Z, Shao H, Zhong X, et al. Ensemble transfer CNNs driven by multi-channel signals for fault diagnosis of rotating machinery cross working conditions. Knowledge-Based Syst, 2020, 207: 106396
36 Wang S, Xiang J, Zhong Y, et al. Convolutional neural network-based hidden Markov models for rolling element bearing fault identification. Knowledge-Based Syst, 2018, 144: 65–76
37 Zhang K, Zuo W, Chen Y, et al. Beyond a gaussian denoiser: Residual learning of deep CNN for image denoising. IEEE Trans Image Process, 2017, 26: 3142–3155
38 Zeiler M D, Fergus R. Visualizing and understanding convolutional networks. In: Proceedings of the 13th European Conference on Computer Vision. Zurich, 2014. 818–833
39 Kingma D P, Ba J. Adam: A method for stochastic optimization. arXiv: 1412.6980
40 Zou F, Zhang H, Sang S, et al. An anti-noise one-dimension convolutional neural network learning model applying on bearing fault diagnosis. Measurement, 2021, 186: 110236
41 Jin G, Zhu T, Akram M W, et al. An adaptive anti-noise neural network for bearing fault diagnosis under noise and varying load conditions. IEEE Access, 2020, 8: 74793–74807
42 Yuan Y, Ma G, Cheng C, et al. A general end-to-end diagnosis framework for manufacturing systems. Natl Sci Rev, 2020, 7: 418–429
43 Chen X, Zhang B, Gao D. Bearing fault diagnosis base on multi-scale CNN and LSTM model. J Intell Manuf, 2021, 32: 971–987
44 Dempster A, Petitjean F, Webb G I. ROCKET: Exceptionally fast and accurate time series classification using random convolutional kernels. Data Min Knowl Disc, 2020, 34: 1454–1495
45 Ma G, Zhang Y, Cheng C, et al. Remaining useful life prediction of lithium-ion batteries based on false nearest neighbors and a hybrid neural network. Appl Energy, 2019, 253: 113626
46 Qin C, Xiao D, Tao J, et al. Concentrated velocity synchronous linear chirplet transform with application to robotic drilling chatter monitoring. Measurement, 2022, 194: 111090
47 Shi G, Qin C, Tao J, et al. A VMD-EWT-LSTM-based multi-step prediction approach for shield tunneling machine cutterhead torque. Knowledge-Based Syst, 2021, 228: 107213
48 Jin Y, Qin C, Huang Y, et al. Multi-domain modeling of atrial fibrillation detection with twin attentional convolutional long short-term memory neural networks. Knowledge-Based Syst, 2020, 193: 105460
49 Jin Y, Qin C, Liu J, et al. A novel domain adaptive residual network for automatic atrial fibrillation detection. Knowledge-Based Syst, 2020, 203: 106122
50 Jin Y, Qin C, Liu J, et al. A novel incremental and interactive method for actual heartbeat classification with limited additional labeled samples. IEEE Trans Instrum Meas, 2021, 70: 1–12
51 Jin Y, Liu J, Liu Y, et al. A novel interpretable method based on dual-level attentional deep neural network for actual multilabel arrhythmia detection. IEEE Trans Instrum Meas, 2022, 71: 1–11
52 Jin Y, Li Z, Liu Y, et al. Multi-class 12-lead ECG automatic diagnosis based on a novel subdomain adaptive deep network. Sci China Tech Sci, 2022, doi: 10.1007/s11431-022-2080-6