IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 17 June 2024, accepted 30 June 2024, date of publication 2 July 2024, date of current version 10 July 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3422234

RESEARCH ARTICLE

Intelligent Fault Diagnosis of Rolling Bearing Based on the Depth Feature Fusion Network

ZIHAN FENG[1,2], HUA DING[1], NING LI[1], GUOSHU PU[3], AND WENBO GONG[3]

[1]Shanxi Key Laboratory of Fully Mechanized Coal Mining Equipment, College of Mechanical and Vehicle Engineering, Taiyuan University of Technology, Taiyuan 030024, China
[2]Shaanxi Shuanglong Coal Industry Development Company Ltd., Yan'an 727300, China
[3]Shaanxi Tongwei Digital Technology Company Ltd., Xi'an 710000, China

Corresponding author: Hua Ding (dinghua2002@163.com)

This work supported in part by the National Natural Science Foundation of China under Grant 52174148, in part by Shanxi Science and Technology Major Special Plan "Unveiling and Leading" Project under Grant 202101010101018, in part by Shanxi Province Science and Technology Cooperation and Exchange Special Project under Grant 202104041101003, in part by Shanxi Science and Technology Innovation Talent Team Project under Grant 202204051001017, and in part by Shanxi Provincial Key Research and Development Project under Grant 202202100401013.

ABSTRACT An intelligent fault diagnosis method of rolling bearings based on a depth feature fusion network was proposed to solve the problems of a single original signal, large noise interference and difficult diagnosis of varying working conditions of rolling bearings. Using various transform domain signals can make the input contain comprehensive information, extract useful fault features, and conduct feature fusion through the attention mechanism feature fusion method to enhance the ability to obtain effective features. In addition, an improved dense network integrates the pooling layer with dense blocks and introduces a multi-scale convolution kernel and squeeze excitation module. It can promote feature reuse, reduce the dimensions of the feature map, and automatically measure the importance of the weight of each feature channel, enhancing the useful features of the current task and suppressing useless features. Hence, the model has an optimum feature extraction capability. The bearing dataset of Case Western Reserve University was selected to verify the fault diagnosis ability of the proposed method. The recognition accuracy rates under ideal, anti-noise and generalization tests were 99.75%, 97.81% and 96.88%, respectively. The recognition accuracy rate was higher than that of the other depth learning models. Experiments and comparative analysis showed that the proposed method has good anti-noise and generalization abilities compared to traditional methods.

INDEX TERMS Rolling bearing, intelligent fault diagnosis, depth feature fusion, transform domain signal, improved dense network, attention mechanism feature fusion method.

I. INTRODUCTION

Rolling bearings are important for supporting the rotating body of machinery and are widely used in automobile gearboxes, aircraft engines and other fields. Its accuracy and reliability significantly impact the performance of a machine. Therefore, it is imperative in contemporary machinery [1]. Deep groove ball bearings, as the most common type of rolling bearings, are mainly composed of outer rings, inner rings, steel balls, and retainers. They have a simple structure, convenient use, small friction coefficient, high limit speed, diverse in size ranges and forms, and are the most widely used type of bearings in the mechanical industry. In a complex operating environment, bearings are easily damaged due to overload, fatigue, wear, and pitting, resulting in the failure of the rotating machinery. This failure induces huge property losses, time wastage and even casualties [2], [3]. Therefore, to ensure the safe and reliable operation of rolling bearings and reduce maintenance costs, realising the fault diagnosis of rolling bearings is urgent and necessary [4].

Traditional fault diagnosis methods mainly include two steps: feature extraction and state classification.

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

Feature extraction adopts transform domain methods, such as time domain and frequency domain [5], [6], [7], while state classification usually adopts the k-nearest neighbour algorithm [8], support vector machine [9], artificial neural network [10] and other machine learning algorithms. However, the above methods have high requirements for signal-processing technology and manual experience judgement, and their universality is poor. Moreover, important features may be lost during signal analysis, affecting the diagnosis results.

In recent years, with the rapid development of artificial intelligence technology, deep learning has made great achievements in computer vision [11], image processing [12], speech recognition [13] and other fields. Its strong learning ability has achieved adaptive feature extraction and fault recognition, reducing its excessive dependence on expert knowledge and signal-processing technology. Therefore, many deep neural networks have been proposed, such as convolutional neural networks (CNN), deep confidence networks, long and short-term memory network, etc. Among them, the CNN, an important network model in the field of deep learning, has achieved remarkable results in the field of image processing. Furthermore, scholars have introduced it to fault diagnosis and they conducted in-depth research on it. Zhang et al. [14] built the first layer of the CNN using a large convolution kernel and combined it with the AdaBN algorithm [15] to achieve fault diagnosis of rolling bearings. Li et al. [16] proposed a bearing fault diagnosis model based on an integrated deep neural network and CNN to complete bearing fault diagnosis. In addition, Gao et al. [17] proposed an adaptive CNN based on Nesterov momentum for the fault diagnosis of rolling bearings, which improved the accuracy of bearing fault classification. Xie et al. [18] proposed a hybrid model based on CNN and individual classifiers, which maximised the feature extraction capability of CNN by selecting classifiers, thus completing bearing fault diagnosis.

CNN has achieved a good diagnostic performance in fault diagnosis. However, a deep CNN improves diagnostic accuracy by increasing the number of network layers, resulting in the disappearing, and exploding gradient problem. To overcome the problem associated with deep CNNs, many researchers have added jump connections between convolutional layers. In particular, a dense network (DN) uses a dense connexion mode to ensure the maximisation of information between layers in the network, alleviating the learning problem and introducing feature reuse. Therefore, DN's deep learning method has been widely applied. Wu et al. [19] used two parallel dense blocks in an improved one-dimensional DN to improve the ability of the network to extract the deep features of the automatic link establishment signal. Zhai et al. [20] proposed an improved single-shot multi-box detector (SSD) target detection algorithm based on DN and feature fusion. This solved the lack of feature complementarity between feature layers of single-point multi-box detectors and the weak ability of SSD to detect small targets. Yang et al. [21] introduced the DN model into transfer learning and realised intelligent grading of liver steatosis.

DN has compelling advantages, but it also has obvious disadvantages. First, each layer of the network obtains the feature map of the current layer by combining the feature maps of all previous layers, it without considering the interdependence between different channels [22]. Second, rolling bearings are susceptible to noise interference during operation, and noise is easily introduced using a deep network, which affects feature extraction [23]. Third, DN has numerous network layers, a complex structure and slow calculation speed, so it cannot quickly identify faults and timely feedback fault information. This results in the occurrence of mechanical shutdowns and property loss. Fourth, the dense blocks and the transition layer of DN are composed of alternating connections. Dense blocks are responsible only for feature map extraction, and the size of the feature map between layers remains unchanged. Reducing the dimension of the feature graph in the transition layer is necessary, resulting in redundancy in most dense block calculations [24]. Therefore, this paper proposes an improved dense network (IDN) to reduce noise interference and improve computing speed. This network uses a dense block structure to maximise information between layers. Furthermore, the pooling layer is introduced into the hop connexion of dense blocks to reduce the dimension of the feature map and simplify the model structure. Finally, a multi-scale convolution kernel and squeeze excitation module (SEM) are integrated into the dense block to enhance the network feature extraction ability.

Currently, the fault diagnosis method based on deep learning determines the fault type by mining and collecting the depth information of the signal. The bearing vibration signal collected using the sensor only contains time domain information, which is not sensitive to non-stationary signals and has unstable characteristic performance. Hence, the difference between faults cannot be highlighted. To obtain comprehensive and multi-dimensional fault information of rolling bearings, the original signal is analysed from multiple angles and multiple domain transformations, which is conducive to improving fault identification accuracy. Many scholars have conducted numerous studies on multiple transform domain signals. Wang et al. [25] used short-time Fourier transform and wavelet transform to process the original signal to achieve improved bearing fault diagnosis. Long et al. [26] conducted Hilbert transform and Fourier transform on different signals to obtain the corresponding frequency domain feature information and achieved motor fault diagnosis using an improved AdaBoost multi-classifier and dynamic weight distribution matrix. Furthermore, Chen et al. [27] conducted wavelet packet decomposition of the original vibration signal for time and frequency domain statistical analyses. Furthermore, they conducted intelligent fault diagnosis based on a multi-core correlation vector machine model.

Based on the above analysis, this paper proposes a rolling bearing intelligent fault diagnosis method based on depth

feature fusion to solve the problems of variable working conditions, single original signal, and strong noise in rolling bearing fault diagnosis. To fully extract the fault feature information of the original signal and maximally reduce the loss of information, the proposed method uses the Fourier transform [28] and Hilbert envelope spectrum [29] methods for the original signal to make the input signal contain time domain and multiple frequency domain information. Next, IDN is used to extract fault information from various transform domain signals for improved mining of the fault characteristics of different transform domain signals. Finally, the attention mechanism feature fusion method (AMFFM) is used to fuse the depth features of different signals extracted from the IDN, and the SoftMax function is used to achieve fault classification. The bearing dataset of Case Western Reserve University was selected for a comparative test to verify the effectiveness of the proposed method, the test results show that the proposed method has good fault diagnosis performance compared to traditional fault diagnosis methods under ideal working conditions, noise environment and variable load conditions.

The rest of this paper is organised as follows: Section II introduces the theoretical foundation, Section III introduces the depth feature fusion network (DFFN) model in detail, Section II introduces the experiment and comparative analysis and, finally, Section II provides the study conclusions.

II. THEORETICAL FOUNDATION

A. MULTIPLE TRANSFORM DOMAIN SIGNAL-PROCESSING

Non-linear and non-stationary vibration signals of rolling bearings contain much useful information. The mode of mining the internal information to accurately identify bearing faults under variable working conditions is crucial. The original vibration signals of rolling bearings contain only time domain information. However, when local damage occurs during operation, it produces vibrations at a specific frequency. Therefore, frequency spectrum analysis has always been an important part of vibration signal analysis. To fully extract the fault feature information of the original signal, this paper proposes taking the time domain, frequency domain and envelope spectrum signals as the input of the model to maximally reduce information loss.

1) FOURIER TRANSFORM

Fourier transform can transform time domain signals into frequency domain signals and can modulate the low-frequency component of the signal into a high-frequency to improve the anti-noise performance of the signal. Unlike time domain signals, frequency domain signals can characterise the global spectrum change characteristics of a non-stationary signal. For discrete vibration signal $x(n)$, its discrete Fourier transform formula is shown in Equation (1):

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi}{N}kn} \tag{1}$$

where, $k = 0, 1, \cdots, N-1, N$ is the length of the discrete sequence.

2) ENVELOPE SPECTRUM METHOD

Compared to spectrum analysis, the envelope spectrum method is more sensitive to shock signals and can provide fault location information through peak frequency. Moreover, it can demodulate high-frequency resonance to low-frequency fault frequency, eliminate unnecessary frequency interference, highlight fault characteristic frequency, and provide good frequency resolution for bearing fault diagnosis. The envelope spectrum calculation method in this paper is based on the Hilbert transform, which is divided into two steps.

Let $x(t)$ be a time domain signal, and its definition of Hilbert transform is represented as follows.

$$h(t) = H[x(t)] = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{x(\tau)}{t-\pi} d\tau \tag{2}$$

The original signal $x(t)$ and its Hilbert transform signal $h(t)$ construct a new analytic signal $z(t)$, as shown in Equation (3).

$$z(t) = x(t) + ih(t) \tag{3}$$

where, i is the imaginary unit, $h(t)$ is the imaginary coefficient of the analytic signal $z(t)$. The modulus $a(t)$ of $z(t)$ analytic signal is represented as follows.

$$a(t) = |z(t)| = |x(t) + ih(t)| = \sqrt{x^2(t) + h^2(t)} \tag{4}$$

Fourier transform is used to obtain the envelope spectrum $E(t)$, as shown in Equation (5).

$$E(t) = |FT[a(t)]| \tag{5}$$

B. DENSE NETWORK

DN was first proposed by Gao et al. [30]. It mainly includes dense blocks and transition layers. Each layer of dense blocks takes all previous layers as inputs, and its output is input to all subsequent layers. Hence, the DN of the L layer has $L(L+1)/2$ connections, ensuring a maximum information flow between layers in the network, eliminating the problem of gradient disappearance and explosion and promoting feature reuse. The transition layer includes the convolution layer (CL) and pooling layer, which are mainly used to reduce the dimension of the feature map. The structure of the original dense block in DN is shown in Figure 1, and the output x_l of the L-th layer is represented as follows.

$$x_l = H_l([x_0, x_1, \cdots, x_{l-1}]) \tag{6}$$

where, $[x_0, x_1, \cdots, x_{l-1}]$ represents the series connexion of the characteristic graphs of Layers 0 to l–1, the process of combining multiple tensors into one, and $H_l(\cdot)$ represents the non-linear transformation function, including a series of operations, such as convolution (Conv), batch normalisation (BN), and rectified linear unit (ReLU).

With the wide application of DN, remarkable achievements have been made in data classification, image processing and

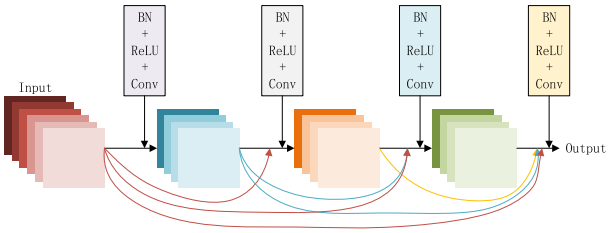


FIGURE 1. Original dense block.

other fields. Tong et al. [31] proposed a method of applying DN to image super-resolution and achieved good results. Zhang et al. [32] proposed a multi-stream network based on residual dense blocks for single-image denoising. Furthermore, the network was applied to mechanical fault diagnosis. Lin et al. [33] proposed a variational mode decomposition DN for bearing fault diagnosis. The network extracts feature from each image block of the image to obtain advanced, accurate diagnosis results, which provides a more effective method for bearing fault diagnosis.

C. SQUEEZE EXCITATION MODULE

The SEM acquires the weight of each feature channel through automatic learning, which makes effective features weigh more and ineffective ones weigh less. This enhances useful features and suppresses unimportant features, increasing the selection and capture ability of the whole network for features. Figure 2 shows the SEM.

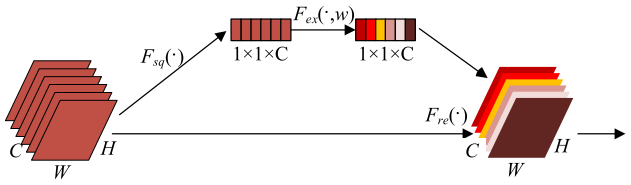


FIGURE 2. Squeeze excitation module.

The number, length, and width of feature channels for a given input feature map are C, W and H, respectively. The characteristics with the number of feature channels are obtained through squeezing, excitation, and reweighting. The detailed steps are as follows:

The squeeze turns each two-dimensional feature channel into a real number with the global receptive field, representing the global distribution of the response on the feature channel as follows.

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} u_c(i,j) \tag{7}$$

where $u_c(i,j)$ is the pixel value of Point (i,j) on the c-th channel feature map, z_c represents the c-th element of z, F_{sp} is the squeeze algorithm.

Excitation introduces two full connexion layers to automatically adjust the correlation between channels and generate the weight for each feature channel, as shown in Equation (8).

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2\delta(W_1 z)) \tag{8}$$

where s is the output of the excitation, w_1 and w_2 are mapping matrices, δ and g are mapping functions,z is the channel information value of the characteristic map and F_{ex} is the excitation algorithm.

The reweight multiplies the weight after excitation with the original feature diagram channel by channel to complete the recalibration of different channel features.

$$\tilde{x}_c = F_{re}(u_c, s_c) = s_c \cdot u_c \tag{9}$$

$\tilde{x}_c$ is the significant feature map obtained using c-th channel, F_{re} is the reweight algorithm.

III. DEPTH FEATURE FUSION NETWORK

In the fault diagnosis of rolling bearings, different transform domains of the original signal have different value information and fault features. The existing depth learning models mainly focus on single-dimension signal feature extraction, which easily induces information loss. To improve the accuracy of fault diagnosis and optimally use the fault information of the signal, this paper proposes a deep feature fusion network based on IDN and AMFFM (Figure 3). First, the original signal is pre-processed to obtain the frequency domain signal and envelope spectrum signal, which form the input of the model with the time domain signal. Second, features are simultaneously extracted from each transform domain through the CL and IDN, and feature fusion is achieved using the AMFFM. Subsequently, features are classified through the flatten layer, full connexion layer and SoftMax. Finally, the diagnosis results are output.

A. IMPROVED DENSE NETWORK

DN has been introduced into fault diagnosis with its unique advantages. However, it has many shortcomings, such as poor dependence between channels, weak anti-noise ability, slow computing speed and redundant data. To overcome the above shortcomings, many variants of DN have recently been proposed and applied to various fields. Li et al. [23] proposed an automatic DN sparse enhancement method to reduce feature redundancy, which improved computing speed. Zhang et al. [24] proposed a new multi-feature reweighting DN architecture, which adaptively recalibrated channel features and adjusted the relationship between the features of different CLs, thus reducing the error rate. Although the above variants of DN improve some shortcomings, they cannot avoid the problem of noise interference and low calculation efficiency. Therefore, it is unsuitable for fault diagnosis under variable working conditions.

Based on the above analysis, this paper proposes an IDN structure (Figure 4). First, the input of each CL of DN concatenates the characteristic graphs of all previous layers without considering the interdependence between different channels. Therefore, SEM is added behind each CL, the weight of each feature channel is obtained through automatic

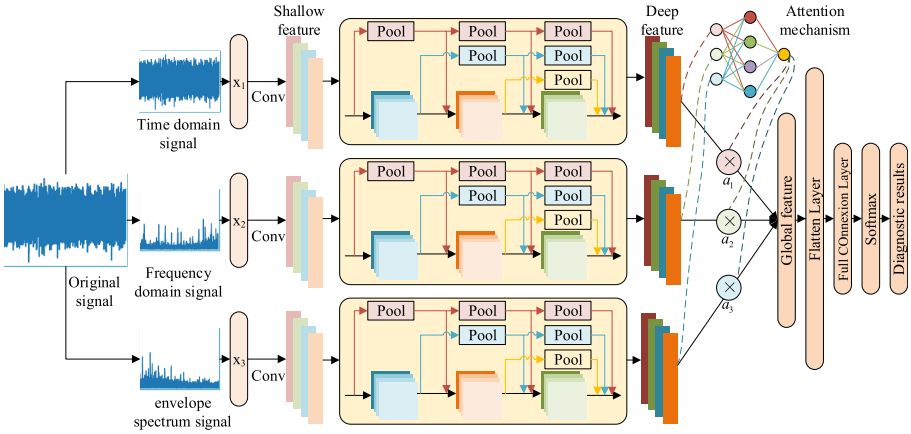


FIGURE 3. Depth feature fusion network structure.

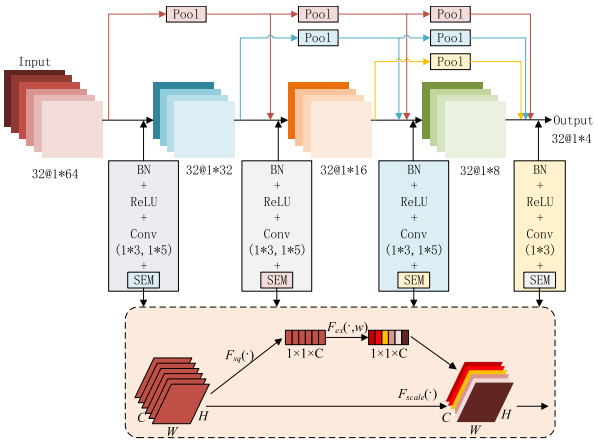


FIGURE 4. Improved dense network.

learning, the useful features are enhanced according to the weight and the useless features are suppressed to establish the relationship between channels. Second, to reduce noise interference and data redundancy, the IDN adopts a dense block structure and introduces its jumping connexion part into the pooling layer to ensure that the information between layers in the network is maximised while reducing the dimension of the feature map. This structure combines the transition layer in DN with dense blocks, inherits the advantages of DN, reduces the complexity of the model and enhances the ability of the model to diagnose faults under variable working conditions. Finally, a multi-scale convolution kernel is used in the CL of dense blocks to increase the network's ability to extract features.

The four-layer CL in Figure 4 further illustrates the IDN structure. The structure comprises four CLs and six pooling layers. The first three CLs have convolution kernel sizes of 1×3 and 1×5, step sizes of 2, and convolution kernel numbers of 16. The fourth CL has convolution kernel sizes of 1×3, step sizes of 2, and convolution kernel numbers of 32. Table 1 shows parameters, such as kernel size, number of input channels, number of output channels, output feature map size and step size in the CL. The CLs are numbered according to the sequence in the figure, and the pre-activation method (BN-ReLU-Conv) is used to reduce the complexity of the model. The size of the six pooling layers is 1×2; hence, the feature map of the jump connexion matches the feature map obtained by the CL, and the dimension of the feature map is reduced while the feature extraction is achieved. Thus, the generation of redundant features is reduced and the convergence speed of the network is increased. IDN was used to learn and extract the time domain, frequency domain and envelope spectrum signals, respectively, to obtain optimal effective depth features for rolling bearing fault diagnosis.

TABLE 1. IDN volume layer parameters.

CL	Kernel size	Input	Output	Feature map	Step
CL1	1×3	32	16	1×32	2
	1×5	32	16		
CL2	1×3	64	16	1×16	2
	1×5	64	16		
CL3	1×3	96	16	1×8	2
	1×5	96	16		
CL4	1×3	128	32	1×4	2

B. ATTENTION MECHANISM FEATURE FUSION METHOD

After the depth features of the time domain, frequency domain and envelope spectrum signals are extracted using IDN, these depth features must be fused to obtain advanced comprehensive bearing fault information. Currently, most scholars use the serial connexion method to fuse multiple features (Figure 5). However, this method ignores the factors with different contributions of different features to fault diagnosis.

Regarding the shortcomings of the fusion method of traditional serial connexion features, this paper proposes an AMFFM (Figure 6). This stems from the different degrees of human attention to various parts of the target object. Particularly, prominent positions always attract attention first, thus enhancing the ability to obtain effective features.

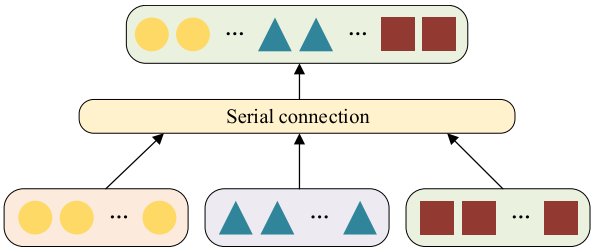


FIGURE 5. Serial connection.

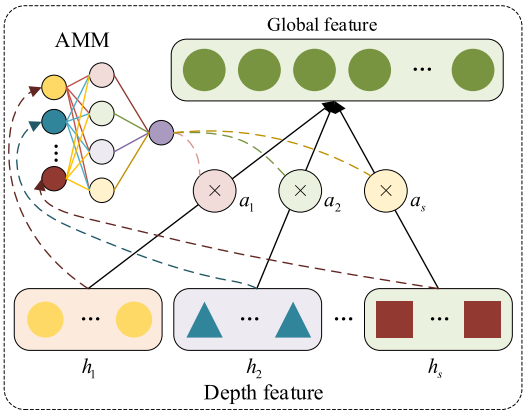


FIGURE 6. Attention mechanism feature fusion method.

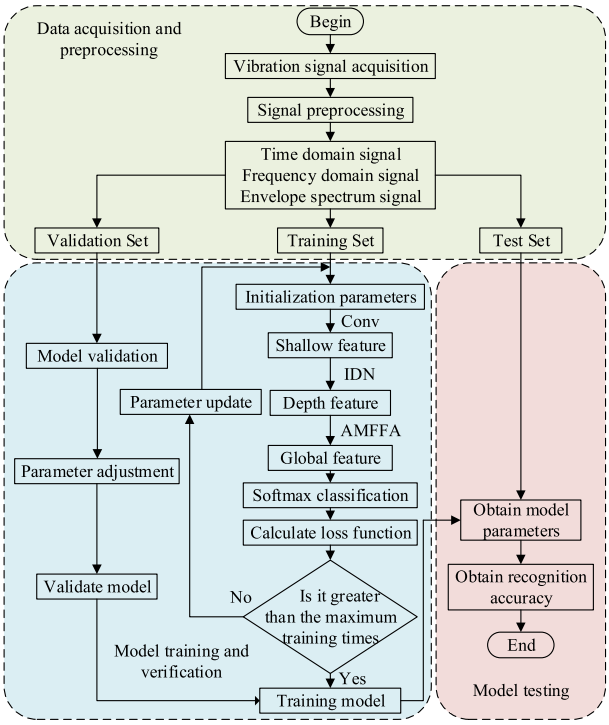


FIGURE 7. Fault diagnosis process of the rolling bearing based on DFFN.

The AMFFM is used to fuse the depth features of multiple signals extracted by IDN to form the global features of bearing fault diagnosis. The specific calculation is as follows.

The weight u_i of each feature in the depth feature map is calculated using dimensionality reduction (equation (10)).

$$u_i = \sigma(w_j h_i + b_j) \tag{10}$$

where, h_i is the depth feature corresponding to the time domain, frequency domain and envelope spectrum signals, w_j and b_j, are the weights and offsets learned, σ is the non-linear activation function.

Subsequently, calculate the importance weight a_i of the feature in the feature map, as shown in Equation (11).

$$a_i = \frac{\exp(u_i)}{\sum \exp(u_s)} \tag{11}$$

where s is the number of features in the feature map. Finally, the global feature f is obtained using weighted fusion, as shown in Equation (12).

$$f = \sum_{i=1}^{n} a_i h_i \tag{12}$$

C. FAULT DIAGNOSIS PROCESS

Figure 7 shows the fault diagnosis process of the rolling bearing based on DFFN, including three stages: data acquisition and pre-processing, model training and verification and model testing. The specific steps are as follows:

Step 1: Sensors were used to collect rolling bearing vibration signals, and the vibration signals were pre-processed to obtain the time domain, frequency domain and envelope spectrum signals. The input sample dimensions of the time domain, frequency domain and envelope spectrum signals are 1×1024, 1×512 and 1×512, respectively.

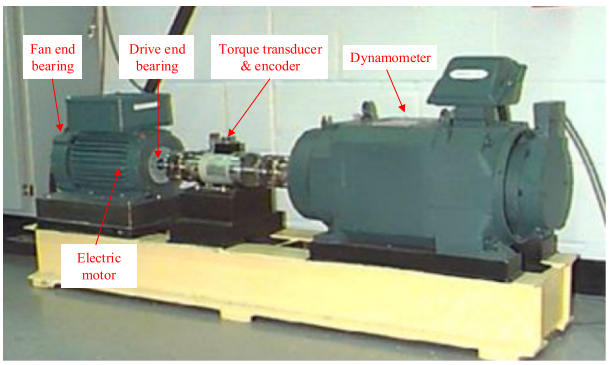


FIGURE 8. CWRU dataset test bench.

Step 2: The three signals were divided into the training, verification and test sets following a certain proportion, and the training set of the three signals was input into the initialisation network.

Step 3: The CL of 32 large-scale convolution kernels with sizes of 64 and steps of 16, 8 and 8 were used to extract the shallow features of the signals, and the depth features of the signals were extracted using IDN. Next, the AMFFM was used to fuse the depth features of the three signals to obtain the global features, and fault classification was conducted using SoftMax. Finally, error back-propagation and network parameter update were achieved by calculating the loss function until the maximum training times were reached. The model was verified by combining the verification set to obtain the training model.

TABLE 2. Specific sample information of the time domain signal.

Sample type	Sample length	Sample number	Sample label
Normal state	1024	200	0
Inner ring damage (7 mils)	1024	200	1
Outer ring damage (7 mils)	1024	200	2
Rolling body damage (7 mils)	1024	200	3
Inner ring damage (14 mils)	1024	200	4
Outer ring damage (14 mils)	1024	200	5
Rolling body damage (14 mils)	1024	200	6
Inner ring damage (21 mils)	1024	200	7
Outer ring damage (21 mils)	1024	200	8
Rolling body damage (21 mils)	1024	200	9

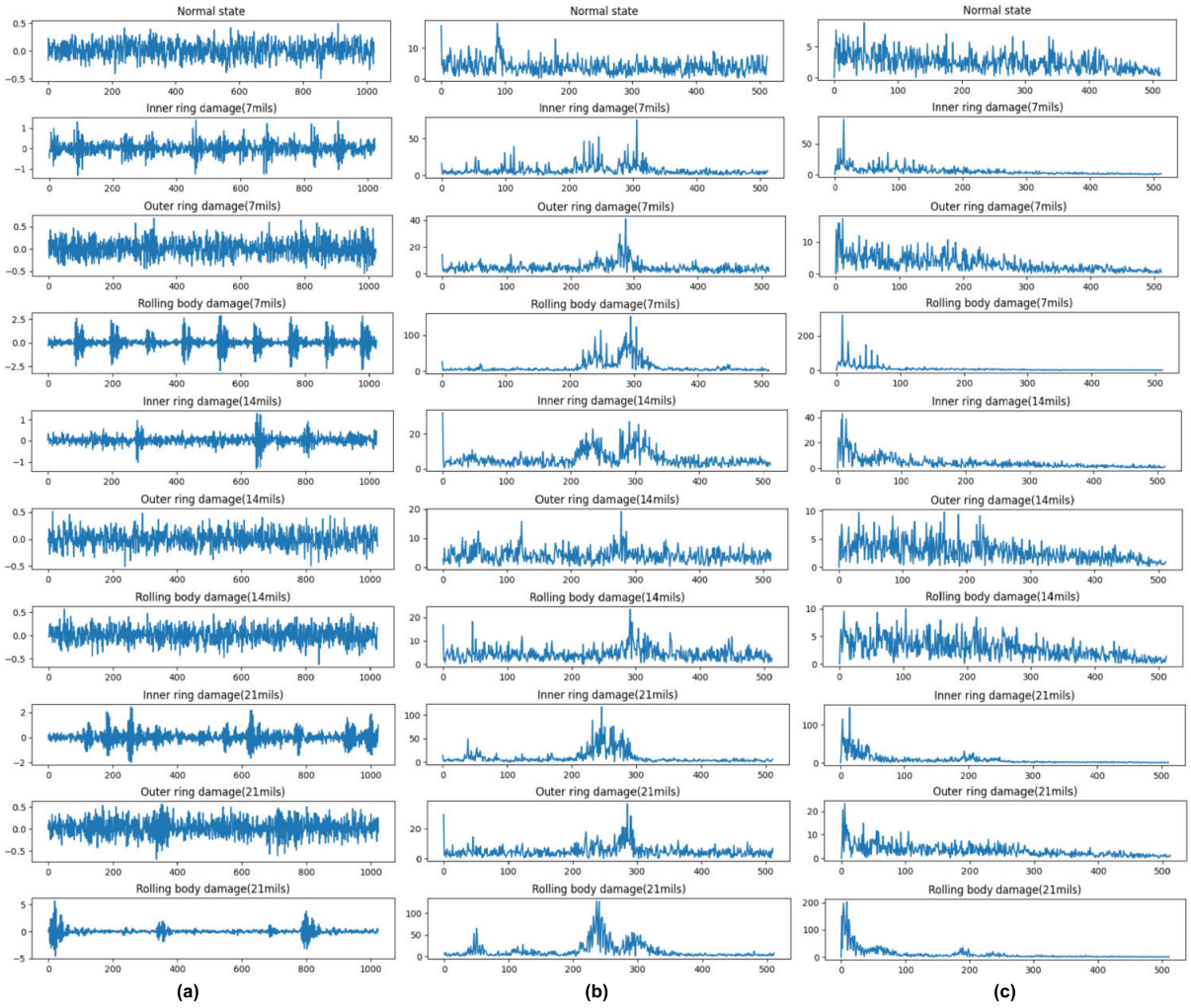


FIGURE 9. Three signal samples of 10 bearing states with a load of 0 hp: (a) time domain sample; (b) frequency domain sample; and (c) envelope spectrum sample.

Step 4: The accuracy of the fault identification of the model was tested using the test set.

IV. TEST AND COMPARATIVE ANALYSIS

A. TEST VERIFICATION

To verify the effectiveness of the proposed method, this paper uses the dataset of the Rolling Bearing Data Centre [34] of Case Western Reserve University (CWRU). The bearing data acquisition equipment includes an electric motor, fan end bearing, drive end bearing, torque transducer and encoder, dynamometer, and other equipment. Figure 8 shows the CWRU dataset test bench, where the data acquisition frequency at the drive end is 12 kHz. All datasets include o hp, 1 hp, 2 hp and 3 hp load types, and each load simulates 10 bearing states, including normal state, inner ring damage, outer ring damage and rolling body damage. The damage

diameters of the inner ring, outer ring and rolling body damage are 7, 14 and 21 mils, respectively. Each bearing state contains 100,000 data points, and 1024 data points are randomly selected as samples. Table 2 shows the specific sample information of the time domain signal. Hence, the original one-dimensional signal is pre-processed to obtain the frequency domain signal and envelope spectrum signal. Due to the symmetry of the sampling frequency, the samples of the frequency domain signal and envelope spectrum signal have only 512 data points. Consider one sample from each of the 10 bearing states with a load of 0 hp to obtain the images of the three signals (Figure 9). Before training, 2000 subsamples are formed for each signal through random sampling, wherein 60% of the samples are selected as the training set, 20% as the verification set and 20% as the test set. The tests in this paper are run on Windows 10, Intel Core i7 CPU, 32 GB processor, and the Keras deep learning framework under Python.

Fig. 9 shows that compared to the time domain samples, the frequency domain and envelope spectrum samples show their characteristics. The signals differ in different states, and compared with single domain signal input signal, using multiple transform domain signals is more conducive to feature classification. Moreover, the amplitude of the frequency domain and envelope spectrum samples in a normal state is relatively small and has increased stability.

To verify the effectiveness of the proposed method, this paper uses the above four load data for the comparative analysis test. The influence of different input modes on the fault identification effect was studied. Four load data were used for ten tests for each input mode, and the average value was taken to obtain the identification accuracy of different input modes under four loads (Table 3).

TABLE 3. Identification accuracy of different input modes/%.

Input method	0 hp	1 hp	2 hp	3 hp	Mean
Time domain signal	98.43	98.08	98.33	98.40	98.31
Frequency domain signal	98.48	98.25	98.48	98.38	98.40
Envelope spectrum signal	98.35	99.38	98.45	98.23	98.35
Proposed method	99.75	99.80	99.73	99.70	99.75

Table 3 shows that the same feature extraction network is used for data of different load types. Compared to the single input method of three signals, the diagnostic accuracy of this method is higher, and the average recognition accuracy can reach 99.75%, indicating that this method has improved recognition and classification capabilities.

In this paper, the IDN is developed from DN, the pooling layer is added to the jumping connexion part and the multi-scale module and SEM are introduced; thus, this improves the accuracy of bearing fault identification. To verify the feature extraction ability of IDN, it is compared with residual network (RN), DN and multi-scale dense network (MDN). The RN includes a large-scale convolution kernel, four residual blocks, a global average pooling layer and an output layer. DN includes a large-scale convolution kernel, three dense blocks, two transition layers, a global average pooling layer and an output layer. Moreover, based on DN, MDN sets the convolution kernel size of dense blocks to 1×3 and 1×5. Table 4 shows the statistical results of the average accuracy, standard deviation, and training time of the four networks.

Table 4 shows that the classification and recognition accuracy of these four networks reached 99%, indicating that the deep learning method has good fault diagnosis ability. Compared to other networks, the proposed network has the highest recognition accuracy among the test results, the minimum standard deviation among multiple tests, improved stability, and the least time consumption, which improves the fault recognition accuracy and diagnostic efficiency. Therefore, the IDN in this paper is superior to the above networks in accuracy and training efficiency. Figure 10 shows the recognition accuracy of the above networks under the four-load data.

TABLE 4. Test results of different networks.

Network	Average recognition rate/%	Standard deviation/%	Training time/s
RN	99.085	0.1335	1.62
DN	99.153	0.1575	1.40
MDN	99.460	0.1611	1.74
IDN	99.745	0.0364	0.51

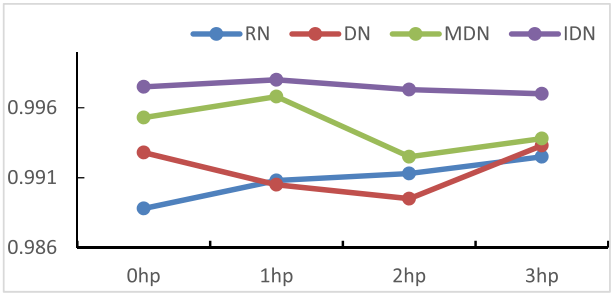


FIGURE 10. Identification accuracy of four networks under four loads.

Finally, to verify the effectiveness of the AMFFM, it was compared to the serial connexion method. The above four load data are used for multiple tests to obtain the identification accuracy and average value of the five serial connexion modes (Figure 11). The recognition accuracy of the serial connexion method was 99.20%. Compared to the recognition accuracy of the proposed method, the recognition accuracy of the serial connexion method is significantly lower, indicating that the AMFFM has improved feature fusion capability.

B. ANTI-NOISE PERFORMANCE ANALYSIS

During the production process, because of the complex working environment of the rolling bearing, the vibration signals collected are vulnerable to noise pollution, which presents higher requirements for the noise resistance of the model. To verify the anti-noise diagnostic capability of the IDN in this paper, Gaussian white noise with a signal-to-noise ratio of 0, 3 dB, 6 dB and 9 dB is added to the test set by using

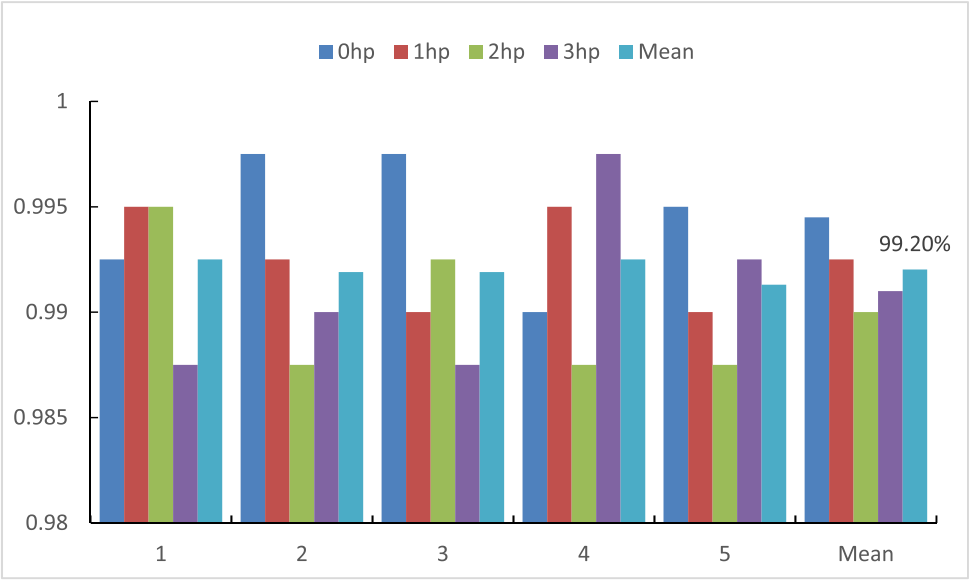


FIGURE 11. Identification accuracy of serial connexion mode in four loads.

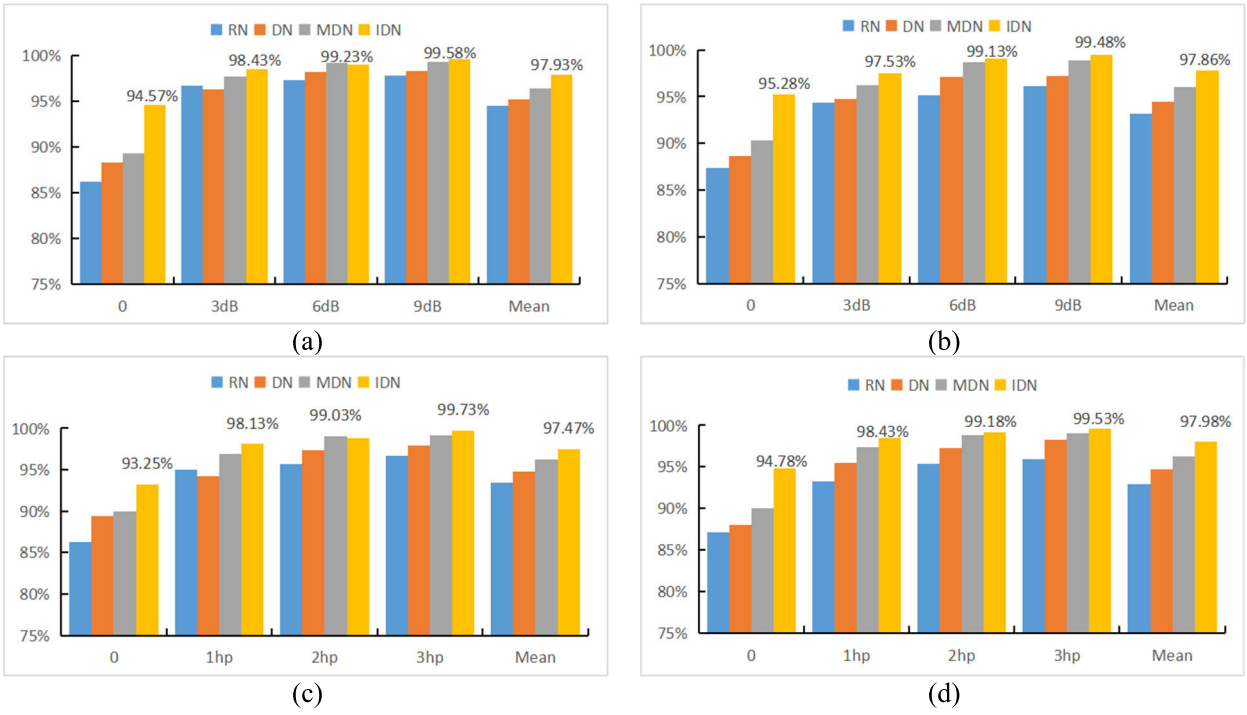


FIGURE 12. Recognition accuracy of the four methods under different load types and noises: recognition accuracy with (a) 0 hp load; (b) 1 hp load; (c) 2 hp load; and (d) 3 hp load.

the four kinds of load data in Section III-A. Multiple tests were conducted on RN, DN, MDN and IDN in different noise environments; the average values were collected to obtain the recognition accuracy of four kinds of loads in noisy environments (Figure 12).

For data of different load types, the fault diagnosis recognition rate of IDN at 0, 3 dB and 9 dB is higher than that of other networks. This is because RN, DN and MDN have considerable network depth and an excellent diagnostic effect on clean signals. However, their diagnostic performance under noise interference was significantly lower than that of IDN. MDN can extract the multi-scale features of signals, and the anti-noise performance is partially improved compared to DN. This indicates that the multi-scale convolution kernel has partial anti-noise capability. The IDN adopts a dense block structure with a reduced number of network layers and adds a pooling layer to the jump connexion part. This structure inherits the advantages of DN, simplifies the DN structure,

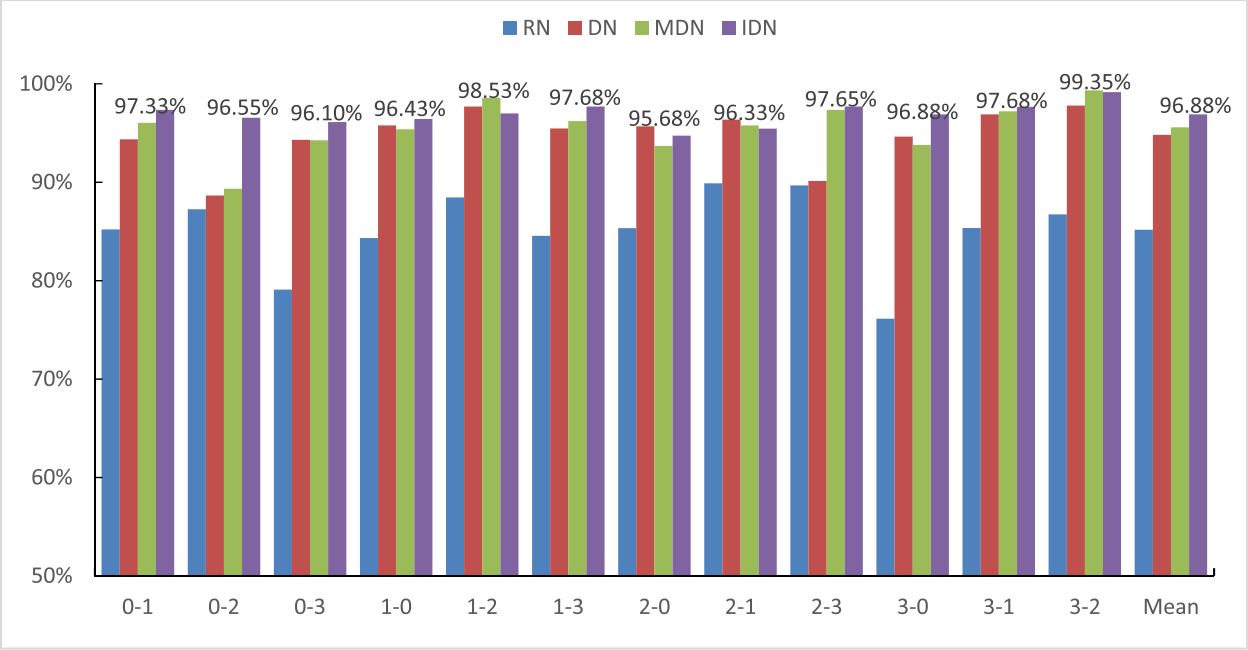


FIGURE 13. Comparison of recognition accuracy of the four methods under different variable loads.

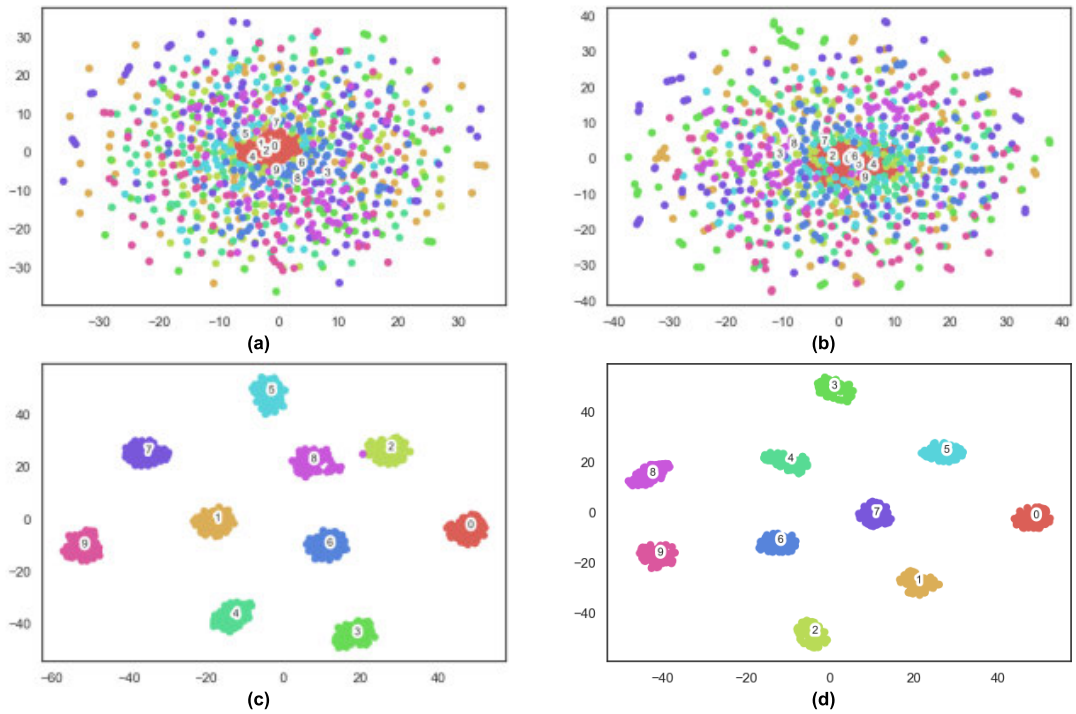


FIGURE 14. Visualisation results of the t-sNE method: (a) input layer; (b) CL; (c) IDN layer; and (d) output layer.

enhances the anti-noise ability of the model and introduces a multi-scale convolution kernel and SEM, which greatly enhances the anti-noise ability of the network. This maintains a high diagnostic performance in 3 dB, 6 dB and 9 dB noise environments and achieves an average recognition accuracy of 94.47% under noise interference with a signal-to-noise ratio of 0. The above tests show that IDN has excellent fault diagnosis ability in a noisy environment.

C. GENERALISATION PERFORMANCE ANALYSIS

During the production process, due to the complex working conditions, the equipment may work under various loads,

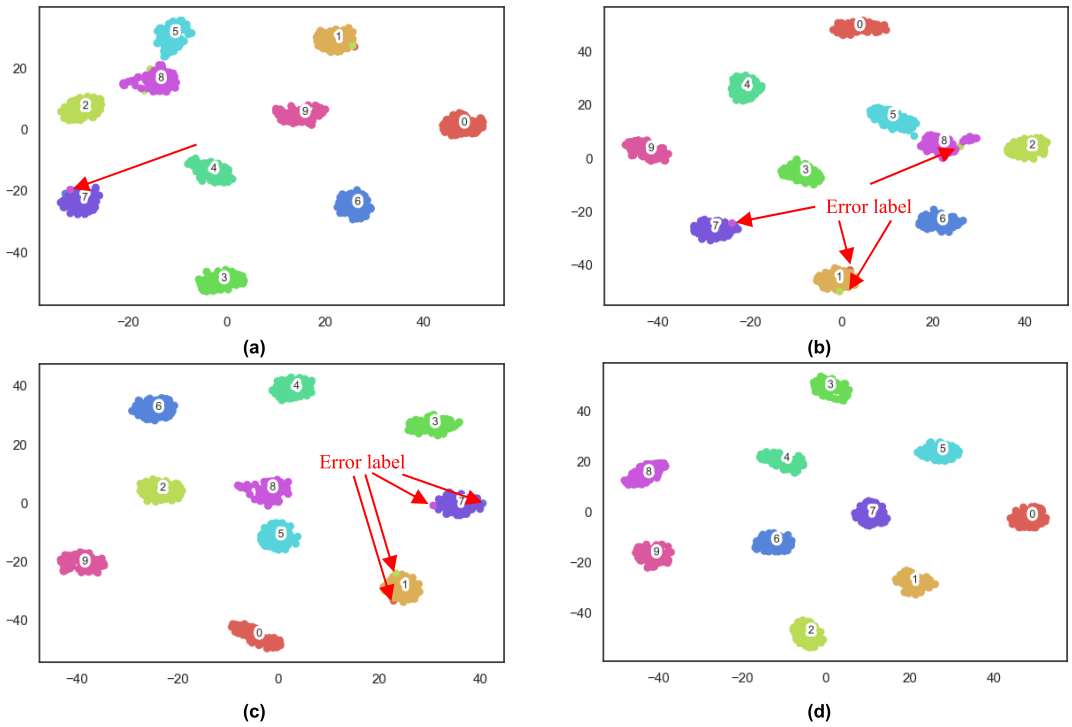


FIGURE 15. Visualisation results of four kinds of t-SNE networks: (a) RN; (b) DN; (c) MDN; and (d) IDN.

which requires the fault diagnosis model to have good generalisation performance. Therefore, to verify the fault diagnosis generalisation capability of the IDN, the load data of the above four loads are used as the training set, and the remaining three load data are used as the test set. For example, the 0 hp load data is used as the training set, and 1 hp, 2 hp and 3 hp load data are used as the test set, represented by 0–1, 0–2 and 0–3.

Figure 13 shows the recognition rates of the four networks under different variable loads. The recognition accuracy of RN decreases significantly when the load changes and the maximum and minimum recognition accuracy rates are 89.67% and 76.12%, respectively, with a maximum difference of 13.55% and an average recognition accuracy rate of 85.16%. This result indicates that the RN method has a poor diagnostic ability under variable load conditions. The maximum and minimum recognition rates of DN under variable loads were 97.78% and 88.64%, respectively, with a maximum difference of 9.14% and an average recognition rate of 94.80%. Compared to the RN method, the recognition accuracy of DN under variable loads has greatly improved, indicating that increasing the information flow between network layers can enhance the network's generalisation ability. Although the recognition accuracy of MDN has decreased under some variable loads compared to single-scale DN, the average recognition accuracy has reached 95.57%, which is higher than that of single-scale DN. This is because the feature information extracted using multi-scale convolution is more comprehensive, further improving generalisation ability. The maximum and minimum recognition accuracy rates of the IDN were 99.14% and 94.72%, respectively, with a maximum difference of 4.42%, and the average recognition accuracy rate reached 96.88%. This is because the network is improved based on DN, which not only inherits most of the advantages of the network but also compensates for its shortcomings and simplifies the network structure. Combining a multi-scale convolution kernel and SEM, the generalisation ability of the proposed method is enhanced. In conclusion, compared to the other three networks, the IDN has better recognition accuracy, stability, and generalisation performance.

D. T-SNE VISUAL ANALYSIS

To illustrate the feature extraction capability of each network layer of the model, t-SNE [35] technology was used in this paper for visual analysis of its input layer, CL, IDN layer and output layer to judge the classification and recognition capability of the model. Figure 14 shows that different colours and numbers represent various fault states of bearings, and the horizontal and vertical coordinates represent different dimensions.

Figure 14 shows that various features of the input layer are mixed without regularity. After the CL, the tags of each feature begin to show a dispersion trend. However, due to its limited feature extraction ability, the tags of each feature are still messy. After passing through the IDN layer, all kinds of features were separated, and tags of the same

kind were gathered, reflecting the powerful feature extraction capability of the IDN. Finally, various features complete separation and convergence through the fusion layer of the attention mechanism feature and the whole full connexion layer, and the prediction tag of various feature in the output layer is obtained.

Furthermore, the t-SNE visualisation analysis of IDN, RN, DN and MDN was conducted (Figure 15). Compared with the other three networks, the proposed network has fewer label errors and higher recognition accuracy, which can optimally reflect the powerful feature extraction capability of the proposed network. And the fundamental reason for the misclassification of the IDN method is that the collected data is interfered by various uncertain factors such as noise, which leads to the same characteristics among different categories of data, resulting in a misjudgment of the types of collected data, and ultimately misclassification of the IDN.

V. CONCLUSION

To solve the difficulty of diagnosing the fault of a rolling bearing under different working conditions, an intelligent fault diagnosis method of rolling bearing based on a DFFN is proposed. the proposed method is validated by using the data set of CWRU rolling bearing data center in the United States, and the following conclusions are drawn:

1) Compared to traditional dense network, the IDN combines the dense block with the pooled layer, which not only realizes the connection of all the front layers with the back layers, promotes the flow of features in the network, but also effectively reduces the data dimension by layer-by-layer pooling, improves the computing power of the network, and reduces the generation of redundant data. Furthermore, the feature extraction method of multi-transform domain signals is more sensitive to vibration signals, and more abundant fault features can be extracted.
2) The AMFFM is proposed for fusing the depth features of various transform domain signals to form the global features of bearing fault diagnosis, which can compile abundant fault information and solve the problem of low accuracy of single-feature fault diagnosis.
3) The experimental results show that compared with ResNet, DenseNet and multi-scale DenseNet networks, IDN has the highest average recognition accuracy of 99.87%, 97.83% and 96.88% respectively under ideal, noisy, and variable load conditions. In addition, under ideal conditions, the stability of multiple experiments is good, with the lowest standard deviation, only 4.63%, and the training time is only 0.692 seconds, indicating the highest computational efficiency. Therefore, the proposed method achieves bearing fault diagnosis under variable operating conditions and strong noise conditions.

REFERENCES

[1] M. M. Manjurul Islam and J.-M. Kim, "Reliable multiple combined fault diagnosis of bearings using heterogeneous feature models and multiclass support vector machines," *Rel. Eng. Syst. Saf.*, vol. 184, pp. 55–66, Apr. 2019.

[2] F. Zou, H. Zhang, S. Sang, X. Li, W. He, and X. Li, "Bearing fault diagnosis based on combined multi-scale weighted entropy morphological filtering and bi-LSTM," *Appl. Intell.*, vol. 51, no. 10, pp. 1–18, 2021.

[3] F. Deng, S. Yang, G. Tang, R. Hao, and M. Zhang, "Self adaptive multi-scale morphology AVG-hat filter and its application to fault feature extraction for wheel bearing," *Meas. Sci. Technol.*, vol. 28, no. 4, Apr. 2017, Art. no. 045011.

[4] Y. Li, X. Liang, and M. J. Zuo, "Diagonal slice spectrum assisted optimal scale morphological filter for rolling element bearing fault diagnosis," *Mech. Syst. Signal Process.*, vol. 85, pp. 146–161, Feb. 2017.

[5] M. Zhang, J. Yin, and W. Chen, "Rolling bearing fault diagnosis based on time-frequency feature extraction and IBA-SVM," *IEEE Access*, vol. 10, pp. 85641–85654, 2022.

[6] M. J. Hasan, M. M. M. Islam, and J.-M. Kim, "Bearing fault diagnosis using multidomain fusion-based vibration imaging and multitask learning," *Sensors*, vol. 22, no. 1, p. 56, Dec. 2021.

[7] C. Li, K. Yang, H. Tang, P. Wang, J. Li, and Q. He, "Fault diagnosis for rolling bearings of a freight train under limited fault data: Few-shot learning method," *J. Transp. Eng. A, Syst.*, vol. 147, no. 8, Aug. 2021, Art. no. 04021041.

[8] I. Aldekoa, A. del Olmo, L. Sastoque-Pinilla, S. Sendino-Mouliet, U. Lopez-Novoa, and L. N. L. de Lacalle, "Early detection of tool wear in electromechanical broaching machines by monitoring main stroke servomotors," *Mech. Syst. Signal Process.*, vol. 204, Dec. 2023, Art. no. 110773.

[9] X. Zhang, P. Han, L. Xu, F. Zhang, Y. Wang, and L. Gao, "Research on bearing fault diagnosis of wind turbine gearbox based on 1DCNN-PSO-SVM," *IEEE Access*, vol. 8, pp. 192248–192258, 2020.

[10] D. Checa, G. Urbikain, A. Beranoagirre, A. Bustillo, and L. N. L. De Lacalle, "Using machine-learning techniques and virtual reality to design cutting tools for energy optimization in milling operations," *Int. J. Comput. Integr. Manuf.*, vol. 35, no. 9, pp. 951–971, Sep. 2022.

[11] A. Ilioudi, A. Dabiri, B. J. Wolf, and B. De Schutter, "Deep learning for object detection and segmentation in videos: Toward an integration with domain knowledge," *IEEE Access*, vol. 10, pp. 34562–34576, 2022.

[12] X. Zhang, "Deep learning-based multi-focus image fusion: A survey and a comparative study," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 9, pp. 4819–4838, Sep. 2022.

[13] X. Cui, W. Zhang, U. Finkler, G. Saon, M. Picheny, and D. Kung, "Distributed training of deep neural network acoustic models for automatic speech recognition: A comparison of current training strategies," *IEEE Signal Process. Mag.*, vol. 37, no. 3, pp. 39–49, May 2020.

[14] W. Zhang, G. Peng, C. Li, Y. Chen, and Z. Zhang, "A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals," *Sensors*, vol. 17, no. 2, p. 425, Feb. 2017.

[15] Y. Li, N. Wang, J. Shi, X. Hou, and J. Liu, "Adaptive batch normalization for practical domain adaptation," *Pattern Recognit.*, vol. 80, pp. 109–117, Aug. 2018.

[16] H. Li, J. Huang, and S. Ji, "Bearing fault diagnosis with a feature fusion method based on an ensemble convolutional neural network and deep neural network," *Sensors*, vol. 19, no. 9, p. 2034, Apr. 2019.

[17] S. Gao, Z. Pei, Y. Zhang, and T. Li, "Bearing fault diagnosis based on adaptive convolutional neural network with Nesterov momentum," *IEEE Sensors J.*, vol. 21, no. 7, pp. 9268–9276, Apr. 2021.

[18] W. Xie, Z. Li, Y. Xu, P. Gardoni, and W. Li, "Evaluation of different bearing fault classifiers in utilizing CNN feature extraction ability," *Sensors*, vol. 22, no. 9, p. 3314, Apr. 2022.

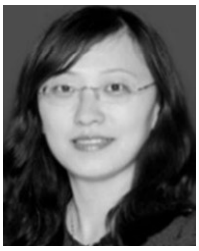
[19] Z. Wu, H. Chen, Y. Lei, and H. Xiong, "Recognizing automatic link establishment behaviors of a short-wave radio station by an improved unidimensional DenseNet," *IEEE Access*, vol. 8, pp. 96055–96064, 2020.

[20] S. Zhai, D. Shang, S. Wang, and S. Dong, "DF-SSD: An improved SSD object detection algorithm based on DenseNet and feature fusion," *IEEE Access*, vol. 8, pp. 24344–24357, 2020.

[21] R. Yang, Y. Zhou, W. Liu, and H. Shang, "Study on the grading model of hepatic steatosis based on improved DenseNet," *J. Healthcare Eng.*, vol. 2022, Mar. 2022, Art. no. 9601470.

[22] K. Zhang, Y. Guo, X. Wang, J. Yuan, and Q. Ding, "Multiple feature reweight DenseNet for image classification," *IEEE Access*, vol. 7, pp. 9872–9880, 2019.
[23] J. Wang, H. Xu, H. Wang, and Z. Wang, "Infrared and visible image fusion based on residual dense block and self coding network," *J. Beijing Univ. Technol. Natural Ed.*, vol. 41, no. 10, pp. 1077–1083, 2021.
[24] T. Li, W. Jiao, L.-N. Wang, and G. Zhong, "Automatic DenseNet sparsification," *IEEE Access*, vol. 8, pp. 62561–62571, 2020.
[25] Y. Wang, M. Yang, Y. Li, Z. Xu, J. Wang, and X. Fang, "A multi-input and multi-task convolutional neural network for fault diagnosis based on bearing vibration signal," *IEEE Sensors J.*, vol. 21, no. 9, pp. 10946–10956, May 2021.
[26] Z. Long, X. Zhang, L. Zhang, G. Qin, S. Huang, D. Song, H. Shao, and G. Wu, "Motor fault diagnosis using attention mechanism and improved AdaBoost driven by multi-sensor information," *Measurement*, vol. 170, Jan. 2021, Art. no. 108718.
[27] F. Chen, M. Cheng, B. Tang, W. Xiao, B. Chen, and X. Shi, "A novel optimized multi-kernel relevance vector machine with selected sensitive features and its application in early fault diagnosis for rolling bearings," *Measurement*, vol. 156, May 2020, Art. no. 107583.
[28] S. Weimann, A. Perez-Leija, M. Lebugle, R. Keil, M. Tichy, M. Gräfe, R. Heilmann, S. Nolte, H. Moya-Cessa, G. Weihs, D. N. Christodoulides, and A. Szameit, "Implementation of quantum and classical discrete fractional Fourier transforms," *Nature Commun.*, vol. 7, no. 1, pp. 1–8, Mar. 2016.
[29] J. Wang, L. Qiao, Y. Ye, and Y. Chen, "Fractional envelope analysis for rolling element bearing weak fault feature extraction," *IEEE/CAA J. Autom. Sinica*, vol. 4, no. 2, pp. 353–360, Apr. 2017.
[30] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 2261–2269.
[31] T. Tong, G. Li, X. Liu, and Q. Gao, "Image super-resolution using dense skip connections," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 4809–4817.
[32] H. Zhang and V. M. Patel, "Density-aware single image de-raining using a multi-stream dense network," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 695–704.
[33] S.-L. Lin, "Intelligent fault diagnosis and forecast of time-varying bearing based on deep learning VMD-DenseNet," *Sensors*, vol. 21, no. 22, p. 7467, Nov. 2021.
[34] W. A. Smith and R. B. Randall, "Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study," *Mech. Syst. Signal Process.*, vols. 64–65, pp. 100–131, Dec. 2015.
[35] L. Van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, no. 11, pp. 2579–2605, 2008.

ZIHAN FENG received the bachelor's degree in mining engineering from Henan Polytechnic University, in 2012. Since 2012, he has been with Shaanxi Shuanglong Coal Industry Development Company Ltd., successively serving as a Mining Technician, a Leader of the Coal Mining Team, the Director of the Production Technology Department, and the Deputy Chief Engineer. He is responsible for the work of one mine and three prevention, mining technology, and scientific and technological innovation management.



HUA DING received the Ph.D. degree in mechanical design and theory (major) from Taiyuan University of Technology, Shanxi, China, in 2011. She is currently a Professor and the Doctoral Supervisor with the Department of Mechanical Design, College of Mechanical and Vehicle Engineering. She was with the Postdoctoral Workstation, Taiyuan Heavy Coal Machine Company Ltd., from 2013 to 2016; and U.S. Pacic University, as a Visiting Scholar, for one year. Her research interests include intelligent design, deep learning, and virtual reality.

NING LI received the B.S. degree in mechanical design, manufacturing and automation from Xi'an Technological University, and the M.S. degree in mechanical engineering from Taiyuan University of Technology, where he is currently pursuing the Ph.D. degree in mechanical engineering. His current research interest includes fault monitoring and diagnosis using data-driven and digital twin methods.

GUOSHU PU received the Graduate degree. He was a former Researcher with the State Coal Industry Bureau. He has been engaged in the design and research of coal mining equipment for a long time. He has presided the technical research and development of multiple national special projects and coal industry research projects. He has won two first and two second prizes for scientific and technological progress at the provincial and ministerial levels, three national invention patents, and 15 national utility model patents. He has published more than 20 professional papers both domestically and internationally. His research field mainly focuses on the intelligent construction of coal mine fully mechanized mining face equipment.

WENBO GONG received the master's degree in surveying and mapping engineering from Anhui University of Science and Technology, in 2015. He was with Shanghai Trony Puyu Automation Engineering Company Ltd., from 2015 to 2020. He was with the Fenxi Huayi Intelligence Equipment Research Institute, from 2020 to 2022. He is currently the Deputy General Manager with Shaanxi Tongwei Digital Technology Company Ltd., and the President of Intelligent Mine Design and Research Institute. His research interests include intelligent fully mechanized mining, design of complete sets of equipment, and intelligent control system for fast digging face.