actuators

MDPI

Article

Wind Turbine Bearing Failure Diagnosis Using Multi-Scale Feature Extraction and Residual Neural Networks with Block Attention

Yuanqing Luo [1], Yuhang Yang [1], Shuang Kang [2,*], Xueyong Tian [1], Shiyue Liu [1] and Feng Sun [3]

[1] School of Environmental and Chemical Engineering, Shenyang University of Technology, Shenyang 110870, China; yqluo1091@sut.edu.cn (Y.L.); yangyuhang@smail.sut.edu.cn (Y.Y.); tianxueyong@sut.edu.cn (X.T.); liushiyue@sut.edu.cn (S.L.)
[2] School of Mechanical and Control Engineering, Baicheng Normal University, Baicheng 137000, China
[3] School of Mechanical Engineering, Shenyang University of Technology, Shenyang 110870, China; sunfeng@sut.edu.cn
* Correspondence: kangshuang2008@126.com; Tel.: +86-155-6793-0678

Abstract: Wind turbine rolling bearings are crucial components for ensuring the reliability and stability of wind power systems. Their failure can lead to significant economic losses and equipment downtime. Therefore, the accurate diagnosis of bearing faults is of great importance. Although existing deep learning fault diagnosis methods have achieved certain results, they still face limitations such as inadequate feature extraction capabilities, insufficient generalization to complex working conditions, and ineffective multi-scale feature capture. To address these issues, this paper proposes an advanced fault diagnosis method named the two-stream feature fusion convolutional neural network (TSFFResNet-Net). Firstly, the proposed method combines the advantages of one-dimensional convolutional neural networks (1D-ResNet) and two-dimensional convolutional neural networks (2D-ResNet). It transforms one-dimensional vibration signals into two-dimensional images through the empirical wavelet transform (EWT) method. Then, parallel convolutional kernels in 1D-ResNet and 2D-ResNet are used to extract multi-scale features, respectively. Next, the Convolutional Block Attention Module (CBAM) is introduced to enhance the network's ability to capture key features by focusing on important features in specific channels or spatial areas. After feature fusion, CBAM is introduced again to further enhance the effect of feature fusion, ensuring that the features extracted by different network branches can be effectively integrated, ultimately providing more accurate input features for the classification task of the fully connected layer. The experimental results demonstrate that the proposed method outperforms other traditional methods and advanced convolutional neural network models on different datasets. Compared with convolutional neural network models such as LeNet-5, AlexNet, and ResNet, the proposed method achieves a significantly higher accuracy on the test set, with a stable accuracy of over 99%. Compared with other models, it shows better generalization and stability, effectively improving the overall performance of rolling bearing vibration signal fault diagnosis. The method provides an effective solution for the intelligent fault diagnosis of wind turbine rolling bearings.

Keywords: wind turbine; rolling bearings; feature extraction; feature fusion; fault diagnosis

check for updates

Citation: Luo, Y.; Yang, Y.; Kang, S.; Tian, X.; Liu, S.; Sun, F. Wind Turbine Bearing Failure Diagnosis Using Multi-Scale Feature Extraction and Residual Neural Networks with Block Attention. *Actuators* **2024**, *13*, 401. https://doi.org/10.3390/act13100401

Academic Editor: Giorgio Olmi

Received: 23 August 2024
Revised: 23 September 2024
Accepted: 3 October 2024
Published: 5 October 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Rolling bearings are crucial mechanical components in wind turbines, and their stable operational status directly impacts the overall safety of the equipment. In practical operation, they are often subjected to high temperatures and complex, variable conditions, making them highly susceptible to failure. According to statistical data, approximately 45% to 55% of failures in rotating machinery are caused by bearing failures. Once a failure occurs, it can cause significant economic losses by affecting the operational efficiency of

the machinery and equipment. In severe cases, it can lead to equipment damage and even fatalities [1–4]. Therefore, conducting research on the technology for diagnosing rolling bearing faults in wind turbines is of great importance to ensure the normal and reliable operation of mechanical equipment.

With the rapid development of machine learning, a series of fault diagnosis methods based on machine learning have continuously emerged. Hu et al. [5] proposed a multi-branch SKNet-based method and an enhanced Inception-ResNet-v2 combined bearing fault diagnosis method, named SIR-CNN, which effectively improves the deep learning model's ability to extract features and the issue of parameter tuning in bearing fault diagnosis. Liang et al. [6] integrated the small-wavelet transform method with advanced domain adaptive networks, proposing a semi-supervised domain adaptive fault diagnosis method that effectively addresses the problem of the unknown fault diagnosis of rolling bearings under different working conditions. Ding et al. [7] proposed a deep unbalanced adaptive architecture (DIDA) by integrating deep sensitivity learning and classification boundary constraints to solve the problem of the fault diagnosis of rolling bearings under multiple unbalanced label transitions. Tan et al. [8] introduced a novel multi-scale sample entropy method (NBMSE) for the fatigue analysis and fault diagnosis of rolling bearings. Xu et al. [9] proposed a deep belief network-based method combined with clustering analysis to extract the bearing vibration signal, followed by dimensionality reduction using principal component analysis to realize the diagnosis of rolling bearing faults. Although traditional intelligent diagnosis methods have achieved good results to some extent, they suffer from shortcomings in industrial big data, such as difficulties in extracting specialized features and weak generalization, making them unreliable and gradually losing their dominance in fault diagnosis.

In recent years, the rapid development of deep learning has drawn widespread attention, ever since Professor Hinton proposed it in 2006. Its powerful feature extraction capabilities and strong adaptability to complex and variable data have made it a popular research topic both domestically and internationally [10,11]. As a new approach, deep learning has gained broad applications in the field of mechanical fault diagnosis. Zhou et al. [12] proposed a new method for generating adversarial networks, avoiding the problem of unbalanced fault data, and designed a new generation and adversarial network-based generator and discriminator to improve the effectiveness of the proposed method. Liu et al. [13] proposed an advanced feature extraction method for rolling bearing fault diagnosis: a two-step sparse stacked autoencoder (STETSAE), which effectively addressed the issue of fault diagnosis under time-varying conditions. Liu et al. [14] introduced an advanced rolling bearing fault diagnosis method based on deep learning and the small-wavelet transform method, which effectively diagnosed rolling bearing faults by analyzing health status indicators under severe heating and abnormal data during long-term operation. Han et al. [15] proposed a rolling bearing fault diagnosis method based on modified sample entropy and unbalanced fault classification, which effectively addressed the problem of insufficient fault samples in practical applications. Deng et al. [16] proposed a multi-scale network (MSNet) combined with auxiliary signals and multiple data sources for fault diagnosis, which demonstrated good generalization and stability across various data types and complex conditions. Wang et al. [17] combined variational mode decomposition (VMD) and one-dimensional convolutional neural networks (1D CNNs) to propose a novel hybrid fault diagnosis method for signal denoising and fault classification. The random noise in the original signal is removed by VMD, and the modal component reconstructed signal after the optimized VMD decomposition is used as the input of the 1D CNN to obtain the fault diagnosis model. Wang et al. [18] proposed a method for improving feature extraction by combining dual-frequency domain features with a deep belief network for fault diagnosis. Zhang et al. [19] proposed a method that integrates multi-dimensional feature extraction with local and global domain-adaptive networks to diagnose rolling bearing faults under varying conditions. Han et al. [20] proposed an advanced method combining residual networks with adaptive noise reduction for multi-task

fault diagnosis. Dubey et al. [21] proposed an automated time–frequency analysis method for fault diagnosis using a self-adaptive wavelet, achieving the automatic initialization of wavelet basis functions and realizing accurate classification under noisy and varying conditions. Zhang et al. [22] proposed a method based on a CNN combined with EEMD and STFT for fault diagnosis, which effectively extracted features and performed fault classification under variable conditions. Liu et al. [23] proposed a method combining a CNN and DBN to enhance feature extraction capabilities. Wang et al. [24] proposed a method based on CapsNet to address the unbalanced fault data, designing a deep network structure to extract valuable and sensitive features, developing a FED-CapsNet model for fault diagnosis under varying conditions. Xu et al. [25] adopted a multi-scale feature learning strategy to extract rich fault information from complex vibration signals. They also introduced a lightweight dynamic separable convolution to reduce the model size, decrease computational loss, and enhance the model's adaptive learning ability for different inputs. Cheng et al. [26] proposed a rolling bearing fault diagnosis method based on a transfer deep learning network, which addresses the limitations of traditional deep learning algorithms, such as large training parameters, long training times, and insufficient training samples, through feature transfer and network transfer. Chen et al. [27] proposed a rolling bearing fault diagnosis method based on cyclic spectral coherence (CSCoh) 2D feature map denoising and a multi-scale convolutional neural network (MSCNN), which effectively improves feature representation and diagnostic accuracy. Fu et al. [28] combined a multi-scale feature extraction method with an improved 1D-ConvNeXt architecture, fusing low-level and high-level features into multi-scale fault features. They also introduced a channel attention mechanism to adaptively assign weights to the fused multi-scale features. Yin et al. [29] adopted a multi-scale feature extraction module and a graph iteration module to extract vibration signal features with different receptive fields twice. By combining these with a Bayesian method-based fusion module, they achieved the in-depth mining of both global and local information in the vibration signal features.

The methods described above have achieved good results in their respective fault diagnosis approaches. However, fault diagnosis methods based on deep learning are also limited by the following two factors: (1) These methods are all single-scale simple convolutional layer feature extraction methods, in which it is easy to ignore the tiny fault features in the single-scale convolutional layer network for rolling bearings under multi-working conditions. In order to solve this problem, a multi-scale feature extraction module is introduced in this paper to obtain those easily ignored features from vibration signals. (2) Since the one-dimensional signal and two-dimensional image signal each have the function of representing different dimensional features [30,31], in order to retain their respective advantages in extracting features, this paper adopts a parallel multi-channel structure of a one-dimensional signal and two-dimensional signal to improve data reliability.

In general, by combining the advantages of 1D-ResNet and 2D-ResNet, using multi-scale feature extraction and fusion strategies, and introducing an attention mechanism (CBAM module), this paper effectively improves the accuracy and robustness of rolling bearing vibration signal fault diagnosis. In summary, the main contributions of this paper are as follows:

1. A reconstruction method is proposed to convert one-dimensional vibration signals into two-dimensional images.
2. A multi-scale feature extraction module combining 1D-ResNet and 2D-ResNet was designed and implemented, using parallel convolutional kernels to extract features.
3. After feature extraction, a CBAM module is added before fusion to enhance the network's ability to capture key features by focusing on important features in specific channels or spatial areas. After feature fusion, CBAM is introduced again to further enhance the effect of feature fusion, ensuring that the features extracted by different network branches can be effectively integrated, ultimately providing more accurate input features for the classification task of the fully connected layer.

The structure of this paper is as follows: Section 2 provides a comprehensive explanation of the theoretical methods. In Section 3, the experimental validation of the proposed approach is discussed. Finally, Section 4 offers the concluding remarks of the paper.

2. Materials and Methods

2.1. Residual Network

ResNet was proposed to address issues such as gradient vanishing, gradient explosion, and training difficulties that arise as the depth of the network increases. Its core idea is to introduce residual connections, allowing the network to learn the residuals between the input and output, thereby simplifying the training of the model. Due to its depth and complexity, ResNet can effectively extract high-level features from images [32].

The typical structure of ResNet, as shown in Figure 1, mainly consists of residual blocks, convolutional layers, batch normalization layers, activation functions, and more. It addresses the gradient vanishing problem in deep neural networks by introducing shortcuts (skip connections).

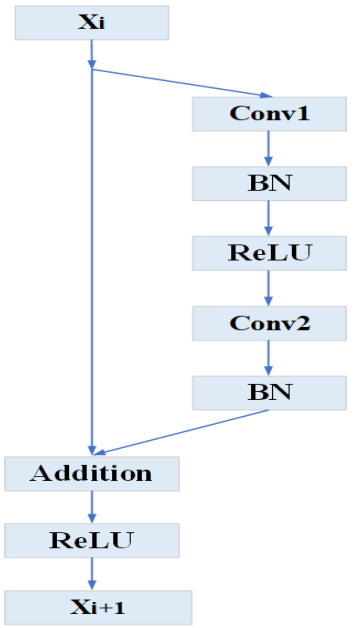


Figure 1. Convolutional neural network structure diagram.

Let X_i represent the input feature map, and X_{i+1} represent the output feature map after passing through the residual block. The core formula of the ResNet residual block is as follows:

$$X_{i+1} = \mathrm{Re}LU(X_i + F(X_i)) \tag{1}$$

$F(X_i)$ is the output after two convolution operations (including BN and ReLU), representing the feature transformation of the input residual. $X_i + F(X_i)$ is the addition of the skip connection and the convolution result. ReLU is a nonlinear activation function used to introduce nonlinearity.

2.2. CBAM Module

In the learning process of neural networks, to address issues such as excessive information overload due to too many model parameters, attention mechanisms are introduced to filter out key information from a large amount of data, reducing attention to irrelevant information, thereby improving the efficiency and accuracy of the model [33]. The Convolutional Block Attention Module (CBAM) applies the fine-grained characteristics of attention to two different modules, namely, the channel attention and spatial attention [34]. The flowchart of the CBAM module is shown in Figure 2.

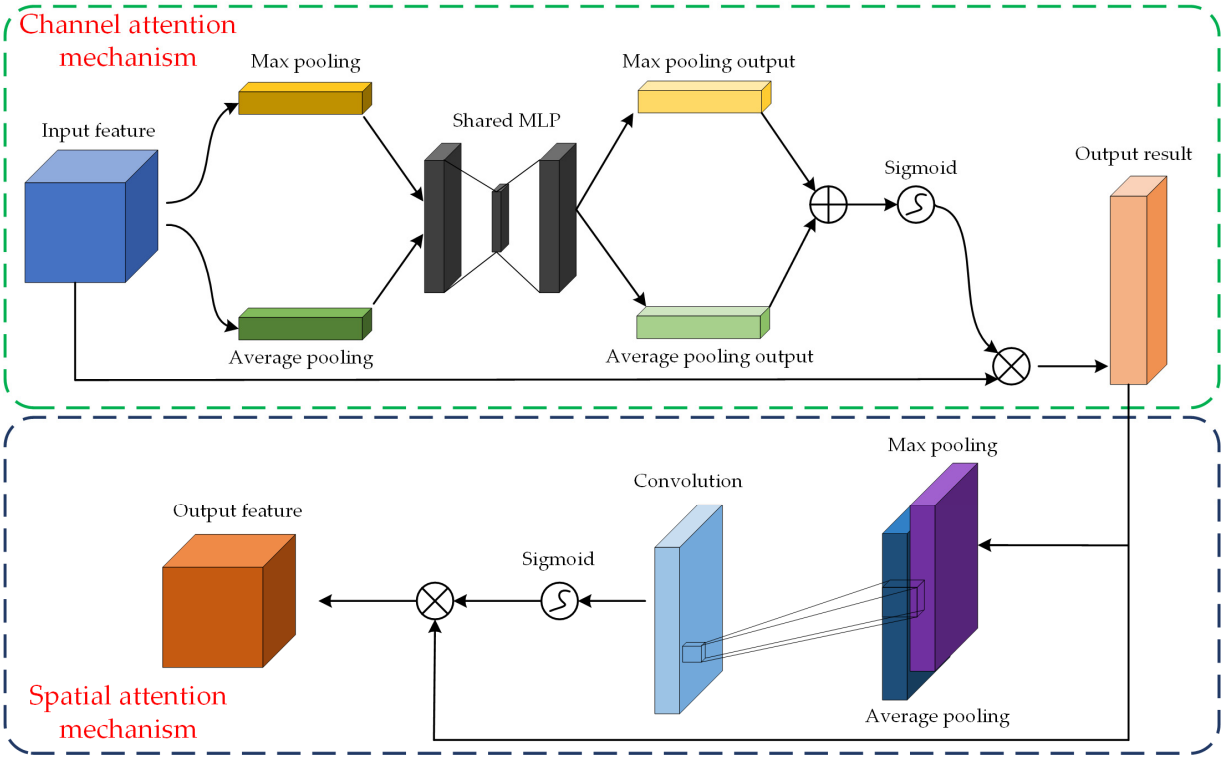


Figure 2. CBAM module flow.

The channel attention module first compresses the feature map globally, then performs feature learning. The results are sent to a multi-layer perceptron (MLP) for learning, added based on the corresponding relationships of the elements, and then processed through the Sigmoid function to obtain the final result, which serves as the input to the spatial attention module. The calculation formula is as follows:

$$M_{ch} = \partial(MLP(AvgPool(F)) + MLP(MaxPool(F))) \tag{2}$$

where ∂ is the Sigmoid activation function, and *MLP* stands for multi-layer perceptron. *AvgPool* and *MaxPool* refer to average pooling and maximum pooling, respectively.

The spatial attention module first performs max pooling and average pooling on the input features, concatenates the results channel-wise, and then processes them through a convolution operation and the Sigmoid activation function to finally obtain the output features. The calculation formula is as follows:

$$M_{sp} = \partial(Conv_{3*3}([MaxPool(F); AvgPool(F)])) \tag{3}$$

where $Conv_{3*3}$ is a convolution operation with a 3×3 kernel size.

2.3. Continuous Wavelet Transform

The continuous wavelet transform (CWT) method has the advantages of multi-resolution time–frequency analysis, adaptability, and time–frequency localization, providing more accurate analysis results in signal and image processing. It is defined as follows: Let $\psi(t)$ be square-integrable in the real number space $L^2(R)$ and satisfy the condition $\int_{-\infty}^{+\infty} \psi(t)\mathrm{d}t = 0$. $\psi(t)$ is called the mother wavelet or basic wavelet. A family of functions is generated by scaling and translating $\psi(t)$.

$$\psi_{a,b}(t) = a^{-\frac{1}{2}}\psi(\frac{t-b}{a}), a \neq 0,\ a, b \in R \tag{4}$$

where a is the scaling factor and b is the translation factor. For any continuous signal $x(t) \in L^2(R)$, the continuous wavelet transform is expressed as follows:

$$WT_x(a,b) =< x(t), \psi_{a,b} >= a^{-\frac{1}{2}} \int_{-\infty}^{+\infty} x(t)\psi(\frac{t-b}{a})\mathrm{d}t \quad (5)$$

where $\langle x(t), \psi_{a,b} \rangle$ denotes the inner product of the two functions.

2.4. Multi-Scale Feature Extraction Module

Due to the long-term operation of bearings in high-temperature and high-pressure working environments, the collected vibration signals are not only affected by the rotating signals of the bearings but also by other noise interferences in the environment. These factors make it easy to overlook some weak feature signals when using simple single-scale feature extraction modules. Therefore, multi-scale feature extraction modules are used to improve the model's feature extraction capabilities [35]. The structure is shown in Figure 3.

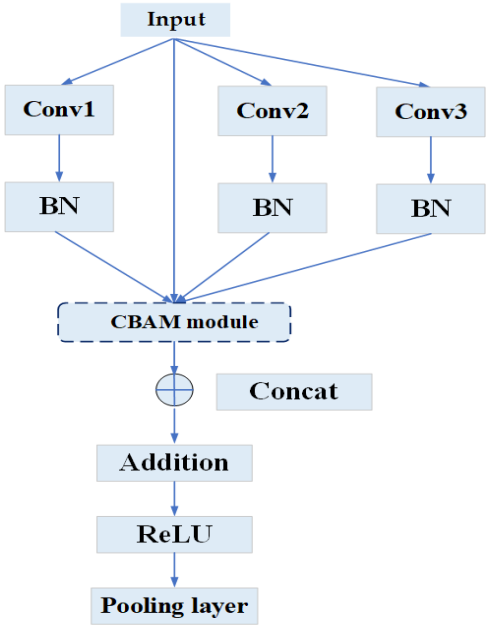


Figure 3. Multi-scale feature extraction module.

Convolutional layers with kernel sizes of 1, 3, and 5 are used to extract multi-scale features from the input data, with a batch normalization (BN) layer following each convolutional layer to stabilize the learning process. The CBAM module effectively assists convolutional networks in better understanding the spatial and channel information of the input data through attention mechanisms, thereby enhancing the quality of features extracted from three different scales. Then, a Concat operation is used for merging. The merged features are then added to the original input data (addition), forming a residual connection that aids in information flow and gradient propagation. The ReLU activation function is used to further enhance nonlinearity. Finally, a pooling layer is employed to downsample the features, reducing the dimension while extracting important information.

2.5. TSFFResNet-Net

This paper proposed a convolutional neural network model based on dual-stream convolutional feature fusion, as shown in Figure 4. From the figure, it can be seen that the network model consists of two parallel main structures: 1D-ResNet and 2D-ResNet. The standardized one-dimensional signal is used as the input for the 1D-ResNet. The one-dimensional signal undergoes continuous wavelet transform to obtain the wavelet

time–frequency map as the input for the 2D-ResNet. First, to extract effective information, a multi-scale feature extraction module is directly introduced, using parallel convolutional kernels of different sizes to capture multi-scale features from the input data. In the 1D-ResNet part, 1×1, 3×1, and 5×1 1D convolutional kernels are used to process the input signal. In the 2D-ResNet part, a similar strategy is adopted, using 1×1, 3×3, and 5×5 convolutional kernels to perform 2D convolution, capturing multi-scale features in the spatial dimension to provide richer spatial information.

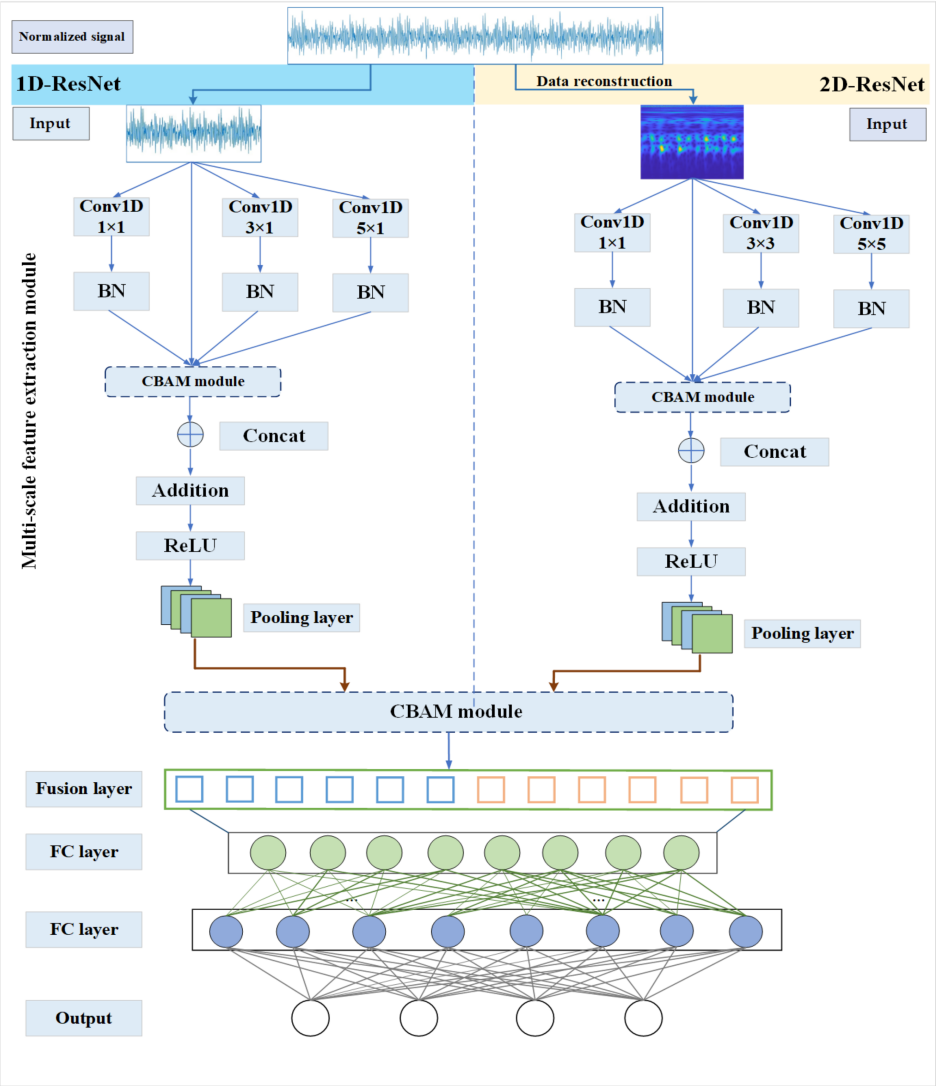


Figure 4. TSFFResNet-Net structure model.

The outputs of multiple convolutional kernels are subjected to batch normalization (BN); then, the CBAM module is introduced, combining the channel attention mechanism and spatial attention mechanism, to enhance the network's ability to capture key features, further optimize the extracted features, and make the network more focused on features that are helpful for the diagnostic task. Then, a Concat operation is used for connection to aggregate features from various scales. This ensures that elements extracted at different spatial scales or ranges can share information and avoid losing important elements. The concatenated features are added to the original input features through a residual connection, introducing skip connections to prevent the vanishing gradient problem and facilitate the training of deeper networks. The addition result is passed through the ReLU activation function, introducing nonlinearity to enhance the model's representation capability. Next, a pooling layer is applied to downsample the features, thereby reducing the size

of the feature maps and lowering computational complexity while retaining the most important information.

After multi-scale feature extraction, the network reintroduces CBAM, which is used to fuse the features extracted by 1D-ResNet and 2D-ResNet, ensuring that the features extracted by different network branches can be effectively integrated, ultimately providing more accurate input features for the classification task of the fully connected layer. The fused features are fed into fully connected layers for processing, further extracting high-level information and progressively compressing the feature dimensions, ultimately being used for decision-making in the output layer. Finally, the output layer achieves result prediction and classification for the specific task.

3. Results

In order to evaluate and verify the effectiveness and accuracy of the proposed method, this paper takes rolling bearings as the research object and uses the open bearing dataset (CWRU) of Case Western Reserve University in the United States and the bearing data collected from a wind turbine test bench built by a company for experimental verification.

3.1. Case 1

The experimental setup at Case Western Reserve University in the United States includes a 1.5 KW motor, a torque sensor, and a power testing device. The model of the driven end bearing is SKF6205. In the experiment, electrical discharge machining (EDM) is used to create single-point defects on the inner ring, outer ring, and rolling elements of the bearing to simulate bearing faults (the above equipments are all from SKF, Cleveland, Ohio, USA). The test bench is shown in Figure 5.

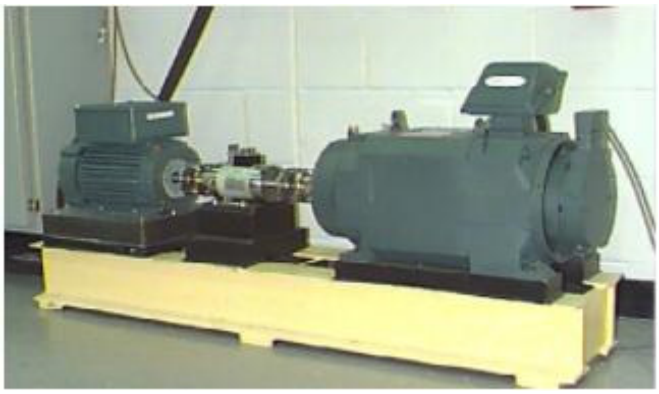


Figure 5. CWRU bearing fault diagnosis testbench.

The defect diameters include 0.1778 mm, 0.3556 mm, and 0.5334 mm. Under four different load conditions of 0~3 hp, the corresponding rotational speeds are 1797 r/min, 1772 r/min, 1750 r/min, and 1730 r/min, respectively. The data were sampled at a frequency of 12 kHz using an accelerometer. Considering that the same type of fault may appear under different conditions and various damage levels, the dataset should include all possible scenarios. Therefore, the data collection includes normal bearing conditions and different fault types across 10 categories. With 1024 data points per sample, a total of 2000 experimental samples were collected, as shown in Table 1.

In the data preprocessing stage, we standardized the data to ensure consistency in magnitude and facilitate model training, thus speeding up the training process. The mathematical expression is as follows:

$$X^* = \frac{X - X(\min)}{X(\max) - X(\min)} \tag{6}$$

As shown in Figure 6a, in the 1D-ResNet channel, the standardized one-dimensional signal is used as the input for this channel. Meanwhile, in the 2D-ResNet channel, as shown in Figure 6, the wavelet time–frequency map obtained by performing CWT on the standardized signal is used as the input for this channel.

Table 1. The CWRU experimental dataset.

Category	Bearing Condition	Fault Diameter/mm	Length	Sample Size	Rotation/(r·min^{-1})
1	Normal	0	1024	400	1730~1797
2	Inner ring fault 1	0.1778	1024	400	1730~1797
3	Inner ring fault 2	0.3556	1024	400	1730~1797
4	Inner ring fault 3	0.5334	1024	400	1730~1797
5	Rolling element fault 1	0.1778	1024	400	1730~1797
6	Rolling element fault 2	0.3556	1024	400	1730~1797
7	Rolling element fault 3	0.5334	1024	400	1730~1797
8	Outer ring fault 1	0.1778	1024	400	1730~1797
9	Outer ring fault 2	0.3556	1024	400	1730~1797
10	Outer ring fault 3	0.5334	1024	400	1730~1797

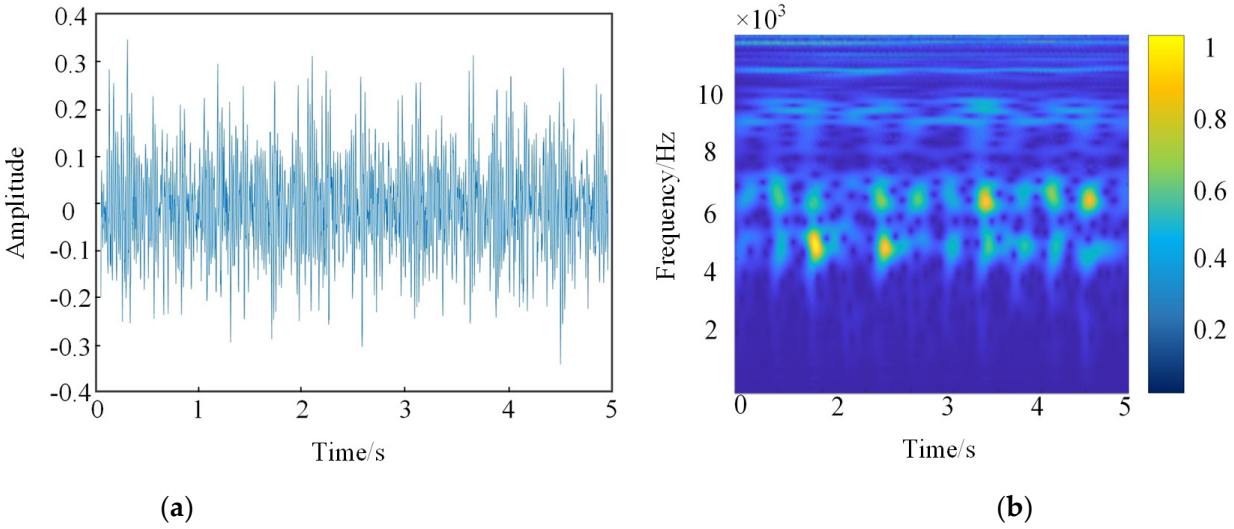


Figure 6. Data input for different channels. (**a**) One-dimensional signal. (**b**) Two-dimensional signal.

The experiment was conducted on a computer with a Windows 11 system, an Intel i9–12900H processor, and an RTX 3060 graphics card (NVIDIA, Santa Clara, CA, USA), using the PyCharm platform and the Python 3.10.6 programming language. The deep learning framework used was PyTorch. The dataset is divided into a training set and a test set in an 8:2 ratio. To prevent gradient explosion and degradation due to the increase in network layers, the Adam optimizer was used to continuously update the network parameters during the experiment, with a batch size of 64, a learning rate of 0.001, and 100 epochs. The experimental network model parameters are shown in Tables 2 and 3.

The training and testing process of the proposed method is shown in Figure 7. From Figure 7, it can be clearly observed that as the number of steps increases, the method demonstrates rapid convergence, finally stabilizing at 99.62%. With the increase in the number of steps, the loss function also converges rapidly. The results indicate that the proposed method has a relatively high fault diagnosis accuracy.

To verify that the proposed method indeed learns discriminative features from the raw shaft vibration signals, the t-SNE algorithm is employed for dimensionality reduction and visualization. The visualization results are shown in Figure 8. Specifically, the graph input signal represents the input data of 10 classes, and it can be observed that they are unorganized within these classes. The graph 1D-Resnet displays the clustering results

of the 1D-Resnet channel. These figures indicate that the clustering results are not ideal, and these features do not effectively separate the fault categories. The graph 2D-Resnet shows the clustering results of the 2D-Resnet channel. Similarly, these figures indicate that effective clustering cannot be achieved due to the mixing of two or more categories. The graph output result displays the clustering results after feature fusion through 1D and 2D channels. The results show that the features obtained by this model are well-clustered, with clearer boundaries and greater distances between different categories, further confirming the effectiveness of the proposed method.

Table 2. Design of 1D-ResNet network structure parameters.

Layer	Name	Kernel Parameter
	Input	1024 × 1
1-1	Convolution layer	24 & 1 × 1
	Convolution layer	24 & 3 × 1
	Convolution layer	24 & 5 × 1
1-2	CBAM module	-
1-3	Pooling layer	1 × 64
1-4	CBAM module	-
1-5	Fully connected layer 1	256 × 64
1-6	Fully connected layer 2	64 × 10

Table 3. Design of 2D-ResNet network structure parameters.

Layer	Name	Kernel Parameter
	Input	64 × 64
2-1	Convolution layer	24 & 1 × 1
	Convolution layer	24 & 3 × 3
	Convolution layer	24 & 5 × 5
2-2	CBAM module	-
2-3	Pooling layer	64 × 64
2-4	CBAM module	-
2-5	Fully connected layer 1	256 × 64
2-6	Fully connected layer 2	64 × 10

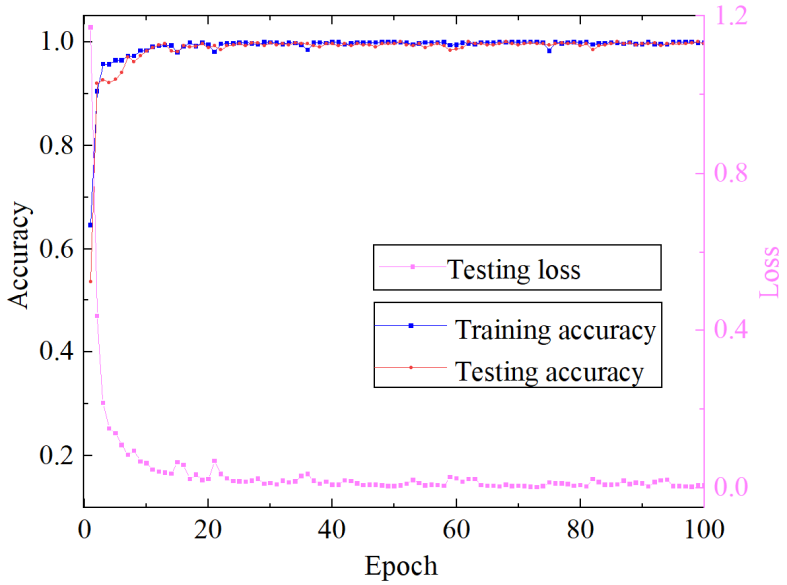


Figure 7. Accuracy and loss during training.

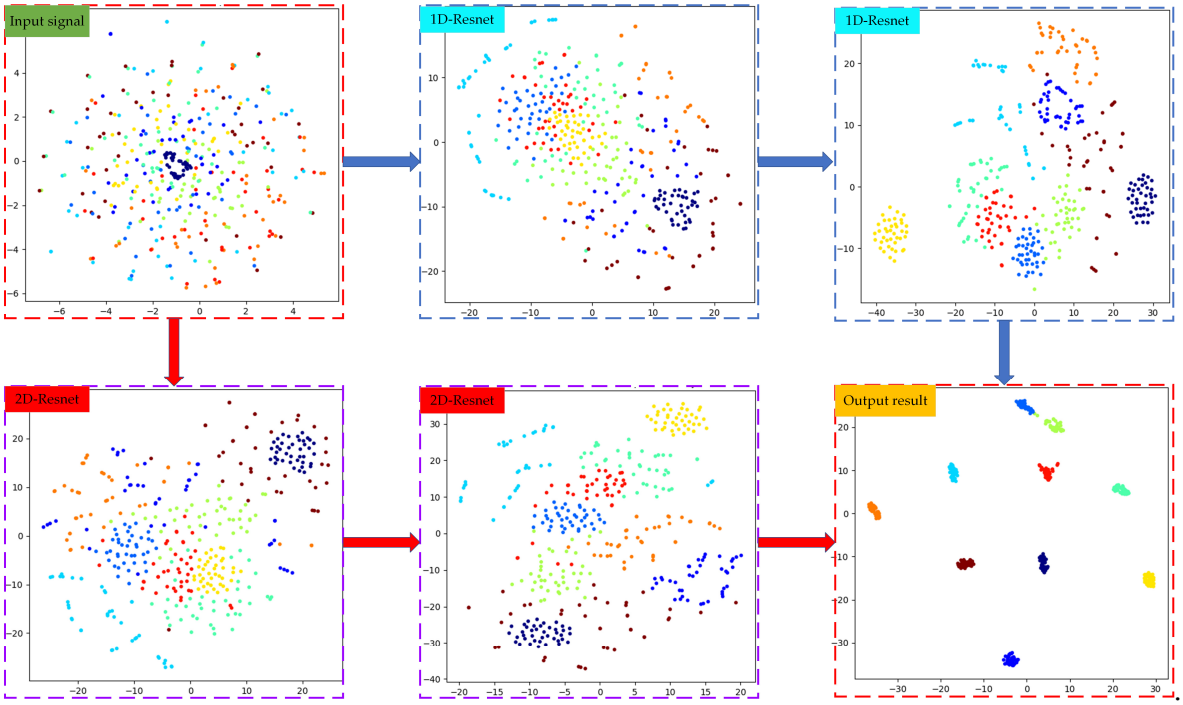


Figure 8. Data visualization results.

In order to verify the reliability of the proposed method, noise variation experiments need to be incorporated into the process. Gaussian white noise with signal-to-noise ratios (SNRs) of 3, 6, 9, and 12 dB is added to the samples in the CWRU rolling bearing dataset. Figure 9 shows a comparison of the fault diagnosis accuracy of the proposed method at different SNRs. It can be seen from the figure that even in a strong noise environment of 3 dB, the diagnostic accuracy of the test data can reach over 96%.

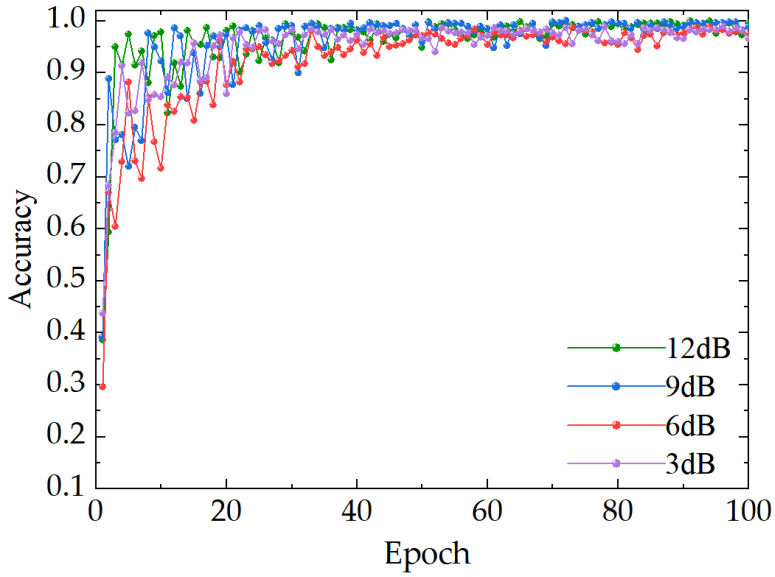


Figure 9. Precision comparison of the proposed method under different SNR.

The obtained results were compared with Alexnet, ResNet, and LetNet-5, and the experimental results are shown in Table 4. It can be clearly seen from Table 4 that the proposed method has significant advantages compared with other methods. When the SNR is 12 dB, the test accuracy can reach 99.62%, which is 4~10% higher than other methods.

Because the multi-scale feature extraction module and attention mechanism module are adopted in this method, when the SNR is 6 dB, the test accuracy is 97.50%, while the test accuracy of ResNet is 90.59%. By comparison, the proposed method still has a higher fault diagnosis rate and better anti-noise ability on different datasets.

Table 4. Fault diagnosis accuracy of different methods and different SNRs.

Method	Diagnostic Accuray			
	3 dB	**6 dB**	**9 dB**	**12 dB**
Alexnet	83.04	85.86	88.50	90.50
ResNet	90.59	91.38	94.80	95.25
LetNet-5	85.81	86.35	88.00	89.25
Our method	96.55	97.50	98.85	99.62

To further verify the superiority of the proposed method, this paper compares the test results of AlexNet, ResNet, and LeNet-5 under the experimental condition of a signal-to-noise ratio of 3 dB through confusion matrix experiments. It is evident from Figure 10 that the proposed method demonstrates the best performance in low signal-to-noise ratio environments, not only with high classification accuracy and low misclassification rates but also with more stable and balanced performance across different categories. This indicates that our proposed method significantly outperforms other models in terms of noise resistance, generalization ability, and classification accuracy.

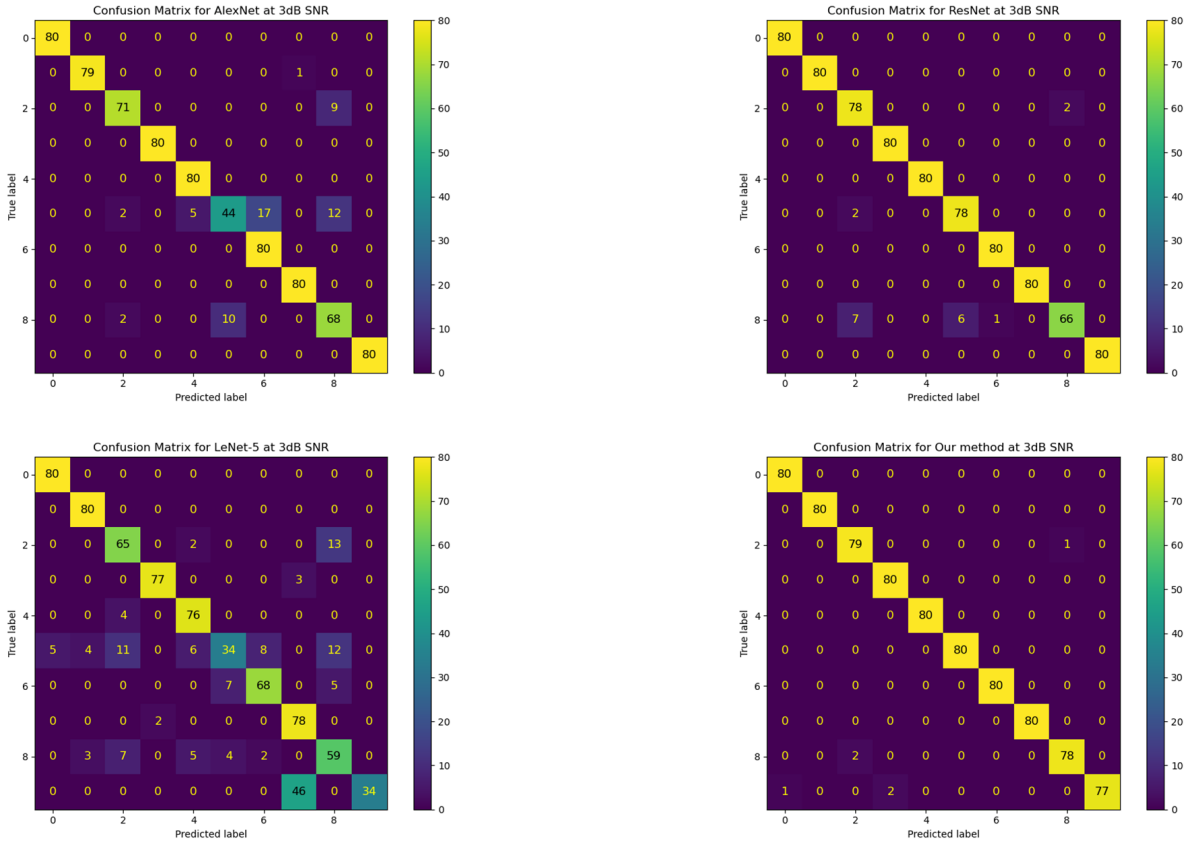


Figure 10. Comparison results of confusion matrix with SNR of 3dB.

To further validate the superiority of the proposed method, this paper compares it with ResNet, LeNet-5, and AlexNet. The trend of fault diagnosis accuracy of the four algorithms is shown in Figure 11. It can be seen from the figure that our method has a significantly faster convergence rate in the initial stage compared to other models, while LeNet-5 and

AlexNet exhibit larger fluctuations in the early stage and slower convergence rates. As the number of epochs increases, it can be observed that the accuracy of our method eventually stabilizes at nearly 97%, outperforming other methods. Although AlexNet and ResNet show similar performance, their accuracy rates are still lower than that of our method, and they have more fluctuations in the middle training stages, indicating their instability during training. In summary, the performance of our proposed method is significantly better than other methods in the figure, with faster convergence speed, higher final accuracy, and better stability in the training process.

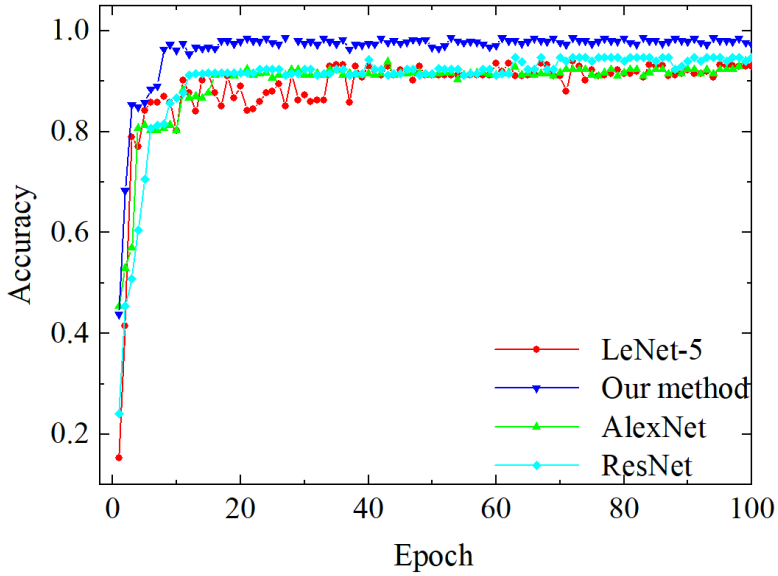


Figure 11. Accuracy curves of four algorithms.

3.2. Case 2

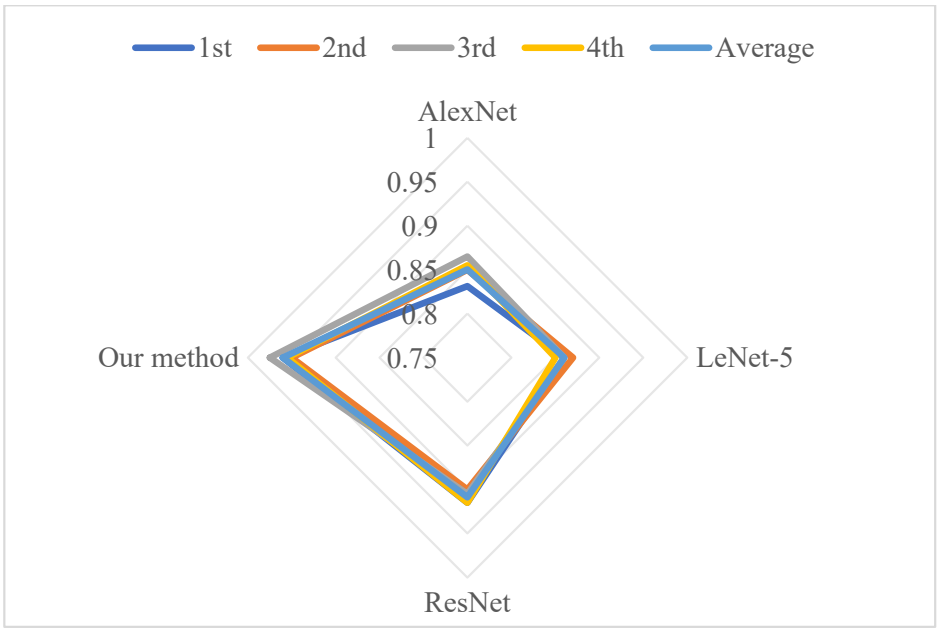
To further verify the superiority of the proposed method, data were collected on the axis of a wind power generator motor test rig built by a certain company. The bearing test platform is shown in Figure 12. In this experiment, two B & K Vibro AS-020 accelerometers (B & K Vibro, Gardnerville, NV, USA) were used to collect data in the vertical and horizontal directions of the bearing under test. The sampling frequency was set to 20 kHz, and each data sample length was 4096. The dataset contains a total of 10 types of faults with different levels of damage, including normal conditions, inner ring damage, outer ring damage, and cage damage. Each type has 40 faulty samples, making a total of 400 faulty samples. The dataset is divided into a training set and a test set in an 8:2 ratio, resulting in 320 training samples and 80 test samples. The specific classification is shown in Table 5.



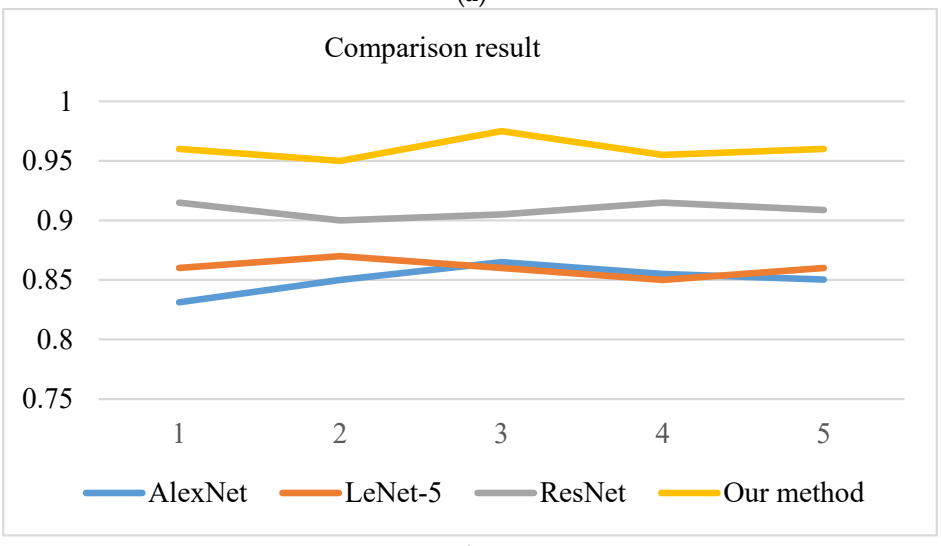
Figure 12. Bearing test platform.

In Case 2, the experimental network model structure is consistent with that of Case 1, with specific parameters shown in Tables 2 and 3. To further verify the superiority of

the proposed method, experiments were conducted on AlexNet, ResNet, and LeNet-5, with comparative results shown in the Figure 13. From Figure 13a, it can be seen that the accuracy of the method proposed in this paper can reach over 95%, while other methods are lower than that of this paper. From Figure 13b, it can be more clearly seen that the diagnostic accuracy obtained by the method proposed in this paper is higher than the other methods in the four experiments.



(**a**)



(**b**)

Figure 13. Experimental comparison results:(**a**) Comparative results of a single experiment. (**b**) The average performance change results of five experiments.

The comparison results of fault diagnosis accuracy under the same load using different methods are shown in Figure 14. From the figure, it can be seen that the training accuracy of our method exceeds 99% (99.04%), indicating that the model has learned almost all patterns during the training process and can fit the training data very well. Other models have lower training accuracy, especially LeNet-5 and AlexNet, whose training accuracy is 91.47% and 93.50%, respectively, showing a significant gap compared to our model. This suggests that they have an insufficient ability to fit the data during training. In terms of testing

accuracy, our method still performs excellently, achieving 98.85%, indicating that the model has a strong generalization ability when dealing with new data. In contrast, AlexNet and LeNet-5 have testing accuracy values of 90.75% and 88.75%, respectively, which are much lower than our method, suggesting weaker performance and an insufficient generalization ability on the test set. Although ResNet also performs well in both training and testing, its testing accuracy is 95.50%, which is still lower than our method, indicating that our approach is more stable in complex tasks.

Table 5. The experimental dataset.

Category	Bearing Condition	Training Set Size	Test Set Size
0	Normal	320	80
1	Inner ring fault 1	320	80
2	Inner ring fault 2	320	80
3	Inner ring fault 3	320	80
4	Rolling element fault 1	320	80
5	Rolling element fault 2	320	80
6	Rolling element fault 3	320	80
7	Outer ring fault 1	320	80
8	Outer ring fault 2	320	80
9	Outer ring fault 3	320	80

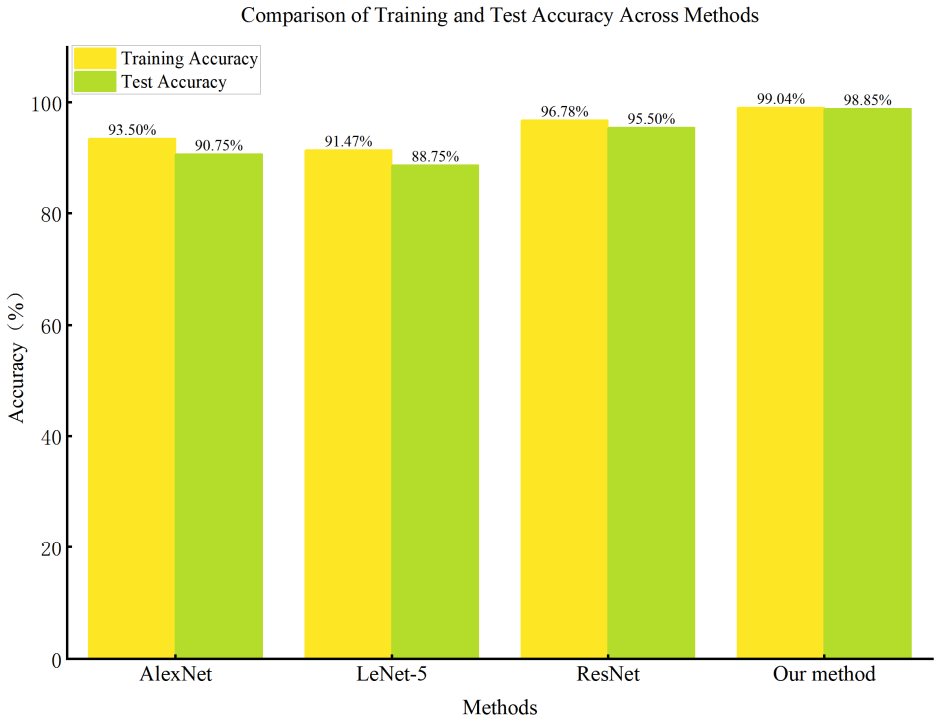


Figure 14. Diagnostic accuracy of different methods.

To further verify the superiority of the proposed method, confusion matrix experiments were conducted, and the test results of AlexNet, ResNet, and LeNet-5 were compared. As shown in Figure 15, our method demonstrates strong robustness and accuracy, while the performance of other classic models significantly declines. AlexNet, LeNet-5, and ResNet exhibit a large number of misclassifications in classes 5, 8, and 9. In contrast, our method has almost no misclassifications, with the majority of samples being correctly classified, further illustrating the stability of the method. Especially in complex noisy environments, our method shows a significant advantage, including higher classification accuracy, lower misclassification rates, and better generalization capabilities.

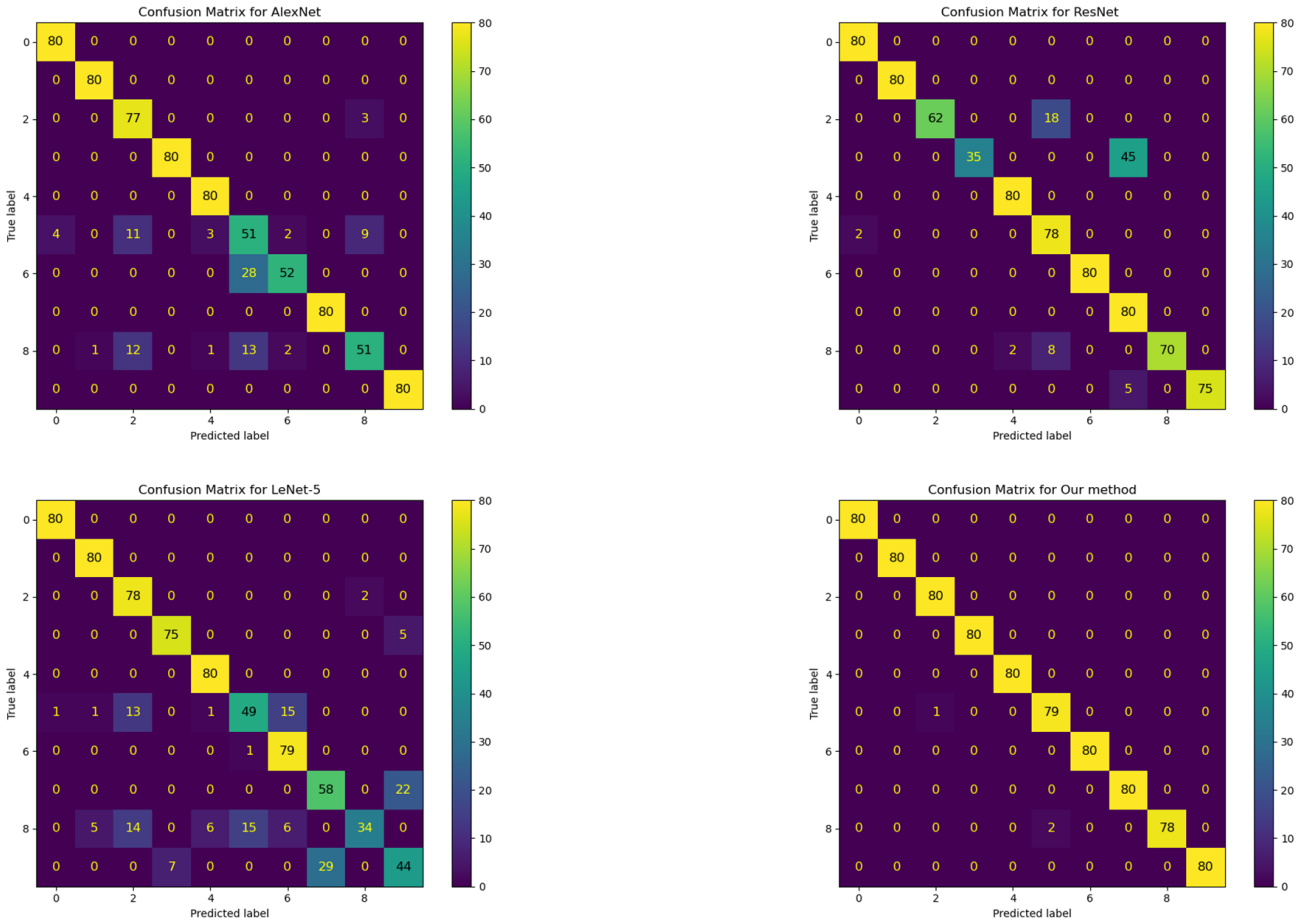


Figure 15. Comparison results of confusion matrix of different methods.

4. Discussion and Conclusions

This paper provides an innovative solution for the fault diagnosis of wind turbine rolling bearings, laying a solid foundation for practical applications by combining multi-scale feature extraction, feature fusion, and attention mechanisms to achieve more efficient and accurate fault identification. The specific conclusions are as follows:

1. The proposed wind turbine rolling bearing fault diagnosis method, based on one-dimensional and two-dimensional convolutional neural networks combined with the empirical wavelet transform method, demonstrated excellent performance in various signal-to-noise ratios and different datasets, showing higher diagnostic accuracy and better robustness. Compared to other classic convolutional neural network methods (such as LeNet-5, AlexNet, and ResNet), this method achieved significantly higher diagnostic accuracy under various signal-to-noise ratio conditions, especially in low signal-to-noise ratio situations, highlighting its superiority in handling noise and complex working conditions.
2. Through comprehensive comparison with different methods, the proposed method achieved an accuracy of over 99% on the test set, significantly outperforming other comparison methods. Additionally, in the confusion matrix analysis, it showed high recognition rates for various fault modes, further demonstrating its outstanding performance in multi-class fault diagnosis tasks.

The exceptional performance of this method in diagnosing rolling bearing vibration signal faults proves its potential for application in actual wind power systems. This method provides a reliable and effective solution for the intelligent fault diagnosis of wind turbine bearings, offering new technical support for the preventive maintenance and fault prediction of industrial equipment.

Author Contributions: Conceptualization, Y.L. and S.K.; methodology, Y.L.; software, Y.L.; validation, Y.L. and Y.Y.; formal analysis, Y.L. and S.K.; investigation, Y.L.; resources, X.T. and S.L.; data curation, Y.L. and S.L.; writing—original draft preparation, Y.L.; writing—review and editing, Y.L.; visualization, Y.Y.; supervision, F.S.; project administration, F.S. and X.T.; funding acquisition, F.S. and X.T. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the National Natural Science Foundation of China (No. 52375258, No. 52005345, No. 52005344).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The detailed data supporting the results of this study are available from the corresponding authors upon request.

Acknowledgments: Thanks to Bai of Shenyang Jianzhu University for providing the experimental bench.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Niu, G.X.; Liu, E.H.; Wang, X.; Zhang, B. A Hybrid Bearing Prognostic Method with Fault Diagnosis and Model Fusion. *IEEE Trans. Ind. Inform.* **2024**, *20*, 864–872. [CrossRef]
2. Ma, Z.P.; Zhao, M.; Dai, X.B.; Bi, H.N. Compound fault diagnosis of wind turbine bearing under ultra-low speed operations using generalized sparse spectral coherence. *Mech. Syst. Signal Process.* **2024**, *208*, 111027. [CrossRef]
3. Damou, A.; Ratni, A.; Benazzouz, D. Intelligent multi-fault identification and classification of defective bearings in gearbox. *Adv. Mech. Eng.* **2024**, *16*, 1–6. [CrossRef]
4. Li, X.M.; Wang, Y.X.; Yao, J.C.; Li, M.; Gao, Z.K. Multi-sensor fusion fault diagnosis method of wind turbine bearing based on viewable neural networks. *Reliab. Eng. Syst. Saf.* **2024**, *245*, 109980. [CrossRef]
5. Hu, B.Q.; Liu, J.; Xu, Y.; Huo, T.L. An Integrated Bearing Fault Diagnosis Method Based on Multibranch SKNet and Enhanced Inception-ResNet-v2. *Shock. Vib.* **2024**, *2024*, 9071328. [CrossRef]
6. Liang, P.F.; Tian, J.Y.; Wang, S.Y.; Yuan, X.M. Multi-source information joint transfer diagnosis for rolling bearing with unknown faults via wavelet transform and an improved domain adaptation network. *Reliab. Eng. Syst. Saf.* **2024**, *242*, 109788. [CrossRef]
7. Ding, Y.F.; Jia, M.P.; Zhuang, J.C.; Cao, Y.D.; Zhao, X.L.; Lee, C.G. Deep imbalanced domain adaptation for transfer learning fault diagnosis of bearings under multiple working conditions. *Reliab. Eng. Syst. Saf.* **2023**, *230*, 108890. [CrossRef]
8. Tan, H.C.; Xie, S.C.; Liu, R.D.; Cheng, J.Q.; Jing, K.K. Fatigue condition diagnosis of rolling bearing based on normalized balanced multiscale sample entropy. *Int. J. Fatigue* **2023**, *172*, 107642. [CrossRef]
9. Xu, F.; Fang, Y.J.; Wang, D.; Liang, J.Q.; Tsui, K.L. Combining DBN and FCM for Fault Diagnosis of Roller Element Bearings without Using Data Labels. *Shock. Vib.* **2018**, *2018*, 3059230. [CrossRef]
10. Wang, C.Y.; Zhu, G.Y.; Liu, T.Y.; Xie, Y.H.; Zhang, D. A Sub-Domain Adaptive Transfer Learning Base on Residual Network for Bearing Fault Diagnosis. *JVC/J. Vib. Control.* **2023**, *29*, 105–117. [CrossRef]
11. Liu, C.Y.; Ban, Y.X.; Li, H.Y.; Guo, N.; Ma, X.Q.; Yang, F.; Sui, X.; Huang, Y. Accurate Recognition Method for Rolling Bearing Failure of Mine Hoist in Strong Noise Environment. *Machines* **2023**, *11*, 632. [CrossRef]
12. Zhou, F.N.; Yang, S.; Fujita, H.; Chen, D.M.; Wen, C.L. Deep learning fault diagnosis method based on global optimization GAN for unbalanced data. *Knowl.-Based Syst.* **2020**, *187*, 104837. [CrossRef]
13. Liu, Y.; Xiang, H.; Jiang, Z.S.; Xiang, J.W. Second-order transient-extracting S transform for fault feature extraction in rolling bearings. *Reliab. Eng. Syst. Saf.* **2023**, *230*, 108955. [CrossRef]
14. Liu, X.Y.; Chen, G.; Wei, X.K.; Liu, Y.B.; Wang, H. A rolling bearing fault evolution state indicator based on deep learning and its application. *J. Mech. Sci. Technol.* **2023**, *37*, 2755–2769. [CrossRef]
15. Han, Y.F.; Li, B.Z.; Huang, Y.K.; Li, L.; Yan, K. Imbalanced fault classification of rolling bearing based on an improved oversampling method. *J. Braz. Soc. Mech. Sci. Eng.* **2023**, *45*, 223. [CrossRef]
16. Deng, J.; Liu, H.; Fang, H.R.; Shao, S.Y.; Wang, D.; Hou, Y.M.; Chen, D.S.; Tang, M.C. MgNet: A fault diagnosis approach for multi-bearing system based on auxiliary bearing and multi-granularity information fusion. *Mech. Syst. Signal Process.* **2023**, *193*, 110253. [CrossRef]
17. Wang, Q.H.; Yang, C.G.; Wan, H.Q.; Deng, D.H.; Nandi, A.K. Bearing fault diagnosis based on optimized variational mode decomposition and 1D convolutional neural networks. *Meas. Sci. Technol.* **2022**, *32*, 104007. [CrossRef]
18. Niu, G.X.; Liu, E.H.; Wang, X.; Ziehl, P.; Zhang, B. Enhanced Discriminate Feature Learning Deep Residual CNN for Multitask Bearing Fault Diagnosis with Information Fusion. *IEEE Trans. Ind. Inform.* **2023**, *19*, 762–770. [CrossRef]
19. Bai, R.X.; Meng, Z.; Xu, Q.S.; Fan, F.J. Fractional Fourier and time domain recurrence plot fusion combining convolutional neural network for bearing fault diagnosis under variable working conditions. *Reliab. Eng. Syst. Saf.* **2023**, *232*, 109076. [CrossRef]
20. Han, C.K.; Lu, W.; Wang, H.Q.; Song, L.Y.; Cui, L.L. Multistate fault diagnosis strategy for bearings based on an improved convolutional sparse coding with priori periodic filter group. *Mech. Syst. Signal Process.* **2023**, *188*, 109995. [CrossRef]

21. Dubey, R.; Sharma, R.R.; Upadhyay, A.; Pachori, R.B. Automated Variational Nonlinear Chirp Mode Decomposition for Bearing Fault Diagnosis. *IEEE Trans. Ind. Inform.* **2023**, *19*, 10873–10882. [CrossRef]
22. Zhang, Y.C.; Ji, J.C.; Ren, Z.H.; Ni, Q.; Gu, F.S.; Feng, K.; Yu, K.; Ge, J.; Lei, Z.H.; Liu, Z. Digital twin-driven partial domain adaptation network for intelligent fault diagnosis of rolling bearing. *Reliab. Eng. Syst. Saf.* **2023**, *234*, 109186. [CrossRef]
23. Chen, J.Y.; Lin, C.Y.; Yao, B.Q.; Yang, L.C.; Ge, H.J. Intelligent fault diagnosis of rolling bearings with low-quality data: A feature significance and diversity learning method. *Reliab. Eng. Syst. Saf.* **2023**, *237*, 109343. [CrossRef]
24. Liu, J.; Zhang, C.H.; Jiang, X.X. Imbalanced fault diagnosis of rolling bearing using improved MsR-GAN and feature enhancement-driven CapsNet. *Mech. Syst. Signal Process.* **2022**, *168*, 108664. [CrossRef]
25. Xu, M.; Shi, Y.M.; Deng, M.Q.; Liu, Y.; Ding, X.; Deng, A.D. An improved multi-scale branching convolutional neural network for rolling bearing fault diagnosis. *PLoS ONE* **2023**, *18*, e0291353. [CrossRef]
26. Cheng, Z.W.; Li, X.J.; Guo, P.; Deng, Y.; Xie, Z.H.; Lian, L. Transfer Deep Learning Network for Rolling Bearing Fault Diagnosis of Wind Turbines. *J. Phys. Conf. Ser.* **2023**, *2503*, 012095. [CrossRef]
27. Chen, X.H.; Lou, W.; Zhao, W.H.; Yang, G.X.; Ding, K.; Zhang, J.W. A Rolling Bearing Fault Diagnosis Method Via 2D Feature Map of CSCoh after Denoising and MSCNN under Different Conditions. *J. Vib. Control.* **2023**, *30*, 1241–1253. [CrossRef]
28. Fu, H.Y.; Yu, D.; Zhan, C.S.; Zhu, X.Z.; Xie, Z.J. Unsupervised Rolling Bearing Fault Diagnosis Method across Working Conditions Based on Multiscale Convolutional Neural Network. *Meas. Sci. Technol.* **2024**, *35*, 035018. [CrossRef]
29. Yin, P.Z.; Nie, J.; Liang, X.Y.; Yu, S.S.; Wang, C.L.; Nie, W.Z.; Ding, X.Q. A Multiscale Graph Convolutional Neural Network Framework for Fault Diagnosis of Rolling Bearing. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 2520713. [CrossRef]
30. Hou, X.L.; Guo, W.C.; Ren, S.J.; Li, Y.; Si, Y.; Su, L.Z. Bolt-Loosening Detection Using 1D and 2D Input Data Based on Two-Stream Convolutional Neural Networks. *Materials* **2022**, *15*, 6757. [CrossRef]
31. Xue, F.; Zhang, W.M.; Xue, F.; Li, D.D.; Xie, S.L.; Fleischer, J. A novel intelligent fault diagnosis method of rolling bearing based on two-stream feature fusion convolutional neural network. *Measurement* **2021**, *176*, 109226. [CrossRef]
32. Li, B.; Lima, D. Facial Expression Recognition via ResNet-50. *Int. J. Cogn. Comput. Eng.* **2021**, *2*, 57–64. [CrossRef]
33. Woo, S.; Park, J.; Lee, J.Y.; Kweon, I.S. CBAM: Convolutional Block Attention Module. In Proceedings of the Computer Vision—ECCV 2018, Munich, Germany, 8–14 September 2018; Volume 11211, pp. 3–19.
34. Chen, Y.T.; Zhang, X.Z.; Chen, W.N.; Li, Y.Y.; Wang, J.S. Research on Recognition of Fly Species Based on Improved RetinaNet and CBAM. *IEEE Access* **2020**, *8*, 102907–102919. [CrossRef]
35. Zhang, Y.L.; Lu, Y.; Zhu, W.Q.; Wei, X.; Wei, Z. Traffic Sign Detection Based on Multi-Scale Feature Extraction and Cascade Feature Fusion. *J. Supercomput.* **2023**, *79*, 2137–2152. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.