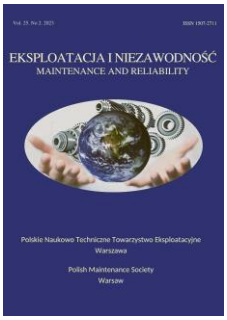


Eksploatacja i Niezawodnosc – Maintenance and Reliability

Volume 28 (2026), Issue 3

journal homepage: http://www.ein.org.pl

Article citation info:
Liu Z, Wang G, Wang B, Shi N, Li Y, Yang X, A bearing fault diagnosis method using a double branch lightweight network under noise interference, Eksploatacja i Niezawodnosc – Maintenance and Reliability 2026: 28(3) http://doi.org/10.17531/ein/218334

A bearing fault diagnosis method using a double branch lightweight network under noise interference

Indexed by: Web of Science Group

Zhen Liu[a], Guoqiang Wang[b,c,*], Bo Wang[a], Nianfeng Shi[b,c], Yingying Li[b], Xianglan Yang[b]

[a] School of Electronic Information, Luoyang Institute of Science and Technology, China
[b] School of Computer Science, Luoyang Institute of Science and Technology, China
[c] Henan Key Laboratory of Green Building Materials Manufacturing and Intelligent Equipment, China

Highlights

- Adaptability in low SNR environments.
- Fusing different fault features extracted by the dual-branch structure.
- Lightweight architecture with integrated residual and depthwise separable convolutions.
- The complex preprocessing is avoided and the original signal is processed directly.

Abstract

For practical industrial production scenarios where collected vibration signals are easily interfered by environmental noise and bearing operating conditions are complex and variable, this paper proposes a Double Branch Lightweight Convolutional Neural Network (DBLCNN). The model adopts a dual-branch architecture: the one-dimensional branch enhances feature extraction capability under low signal-to-noise ratio conditions, while the two-dimensional branch improves feature representation while significantly reducing the number of parameters. The complementary fault features extracted by the dual branches effectively enhance the accuracy of fault diagnosis. Under varying operating conditions, the model achieves an average accuracy of 95.58%; with the addition of 0 dB Gaussian white noise, its average accuracy under varying conditions remains at 90.17%. This study demonstrates that, even based on raw vibration signals without cumbersome preprocessing, the model can achieve excellent diagnostic performance in noisy environments and under variable operating conditions.

Keywords
bearing fault diagnosis, noise interference, residual block, depthwise separable convolution, lightweight

1. Introduction

Rolling-element bearings serve as critical components in rotating machinery, directly influencing system reliability, operational safety, and service longevity. When operating under demanding conditions such as high rotational speeds, variable loading, and corrosive surroundings, these bearings are especially susceptible to failures including surface cracking, spalling, and pitting corrosion. These defects may lead to unscheduled downtime or catastrophic breakdowns. Therefore, the development of bearing fault diagnosis methods remains a central yet challenging focus in the fields of condition monitoring and intelligent maintenance.

Early approaches to bearing fault diagnosis primarily relied on signal processing techniques and shallow machine learning models. Among these, vibration signal analysis served as the most common method, typically requiring manual feature engineering to extract discriminative characteristics from raw sensor data. These handcrafted features often included attributes from the time domain, frequency domain, and time-frequency

(*) Corresponding author.
E-mail addresses: Z. Liu (ORCID:0009-0003-2272-4536) 200900501791@lit.edu.cn, G. Wang (ORCID: 0000-0001-6277-5467) wgq@lit.edu.cn, B. Wang, bowang@lit.edu.cn, N. Shi (ORCID:0000-0002-0346-6071) shinf@lit.edu.cn, Y. Li (ORCID: 0000-0002-2096-8313) liyingying@lit.edu.cn, X. Yang, yxl11162023@163.com

representations [1–3]. While these methods have achieved some success, their performance heavily depends on expert knowledge and manual feature engineering, which is both labor-intensive and subjective. With the progress in machine learning, shallow models such as Support Vector Machines (SVM) [4] and Artificial Neural Networks (ANN) [5] were introduced, partially reducing the need for manual feature design. However, in complex industrial environments, bearing vibration signals often exhibit strong noise, non-stationarity, and non-linearity. Traditional methods struggle to adaptively extract representative fault features under these conditions, leading to limited generalization capability and diagnostic accuracy [6,7]. The field of fault diagnosis has been greatly transformed by recent advances in deep learning, which enable models to learn features directly from raw or minimally processed sensor data using an end-to-end methodology [8], thereby avoiding the reliance on labor-intensive and expert-dependent manual feature extraction processes. This ability to autonomously acquire hierarchical and abstract representations has given fresh impetus to fault diagnosis research. Within the suite of deep learning architectures, convolutional neural networks (CNNs) have demonstrated notable achievements in the realm of fault detection [9–13]. Nevertheless, the practical deployment of these models in industrial settings faces a significant challenge: vibration data collected from bearings are frequently corrupted by various noise sources. These disturbances often originate from the operational dynamics of the machinery itself as well as external environmental conditions. Furthermore, complex vibration superposition and coupling effects make it difficult for acceleration sensors to capture fault signals accurately—especially when the fault characteristics are weak. Under such conditions, fault features can become indistinguishable and may even be entirely masked by noise.

The reliable and efficient extraction of subtle fault characteristics from bearing vibration signals remains a fundamental problem in intelligent fault diagnosis, particularly in industrial settings characterized by intense noise and low signal-to-noise ratios. To address this, Eduard Muratbakeev and colleagues introduced a hybrid architecture integrating 1D-CNN and LSTM network. This model uses the 1D-CNN to learn local temporal patterns, and the LSTM to model temporal dynamics across sequences, leading to enhanced accuracy in fault recognition. Experimental results indicate that the proposed method achieves effective identification of various fault categories, delivers reliable classification results, and maintains robust performance under noisy conditions [14]. Wang et al. developed an end-to-end denoising model based on convolutional neural networks, which incorporates a Discrete Wavelet Attention Layer (DAL) alongside standard convolutional operations to capture features within the wavelet domain. By utilizing the inherent periodic self-similarity property of vibration signals, a Gramian Noise Reduction (GNR) technique is applied to accentuate fault-related information. The integration of both GNR and DAL modules within a cohesive framework enables the model to leverage multi-source features derived from original vibration data as well as enhanced signals after noise suppression. This unified approach leads to a substantial improvement in fault diagnosis accuracy under high-noise conditions [15]. Ni et al. proposed an end-to-end model termed the Discrete Wavelet Ensemble Convolutional Residual Network (DWCResNet). In this architecture, a layer based on the Discrete Wavelet Transform (DWT) is used in place of conventional downsampling to break down input signals into high- and low-frequency subbands. This process helps mitigate high-frequency noise interference and strengthens discriminative feature learning. As a result, the proposed approach significantly enhances both the precision and reliability of fault diagnosis in bearings [16].

Although existing bearing fault diagnosis models demonstrate effective performance under noisy conditions, they often rely on combining multiple networks or deepening model architectures to extract useful features from highly noisy signals. This approach inevitably increases structural complexity, making such models difficult to deploy on embedded edge devices with limited storage and insufficient computational resources. Consequently, their practicality in real-time industrial monitoring scenarios remains constrained.

Therefore, developing a fault diagnosis model that is lightweight, low-complexity, and highly accurate to meet edge-side deployment requirements has become a critical challenge. Yuan et al. introduced an integrated intelligent diagnostic system utilizing empirical mode decomposition and multi-modal feature fusion. The approach employs an enhanced multi-

scale convolutional neural network to capture multi-band characteristics derived from vibration signals. By leveraging adaptive signal processing and collaborative feature learning, the framework demonstrates improved performance under challenging operational environments such as variable rotational speeds and high levels of noise interference. By replacing the standard multi-head attention in Transformer with a lightweight alternative, the model significantly reduces computational complexity while maintaining accuracy. Experimental results demonstrate superior diagnostic performance and robustness, particularly under complex operating conditions [17]. To enhance feature extraction for fault diagnosis, Zhang and colleagues developed a parallel deep separable residual network integrated with an attention mechanism, named PDSResNet-AM. The model employs a dual-input architecture, whereby vibration signals are converted into time-frequency representations via both continuous wavelet transform (CWT) and short-time Fourier transform (STFT), thereby supplying a more comprehensive set of input features. By integrating depthwise separable convolutions (DSConv) and a convolutional block attention module (CBAM), the network strengthens the emphasis on informative features, decreases parameter volume, and helps prevent overfitting. Comprehensive experimental evaluations across diverse operational scenarios, including varying noise levels, load magnitudes, and rotational speeds, demonstrate that the proposed approach surpasses existing deep learning methods in diagnostic precision, generalization capability, and robustness to noise [18]. Cai et al. developed an end-to-end vision Transformer model that integrates time-frequency features with spatial and channel dual-attention mechanisms. The dual-branch architecture includes a temporal branch capturing local and global features via spatial and channel attention, and a frequency branch using FFT to extract spectral information. Multi-scale feature fusion is applied to combine time and frequency domains. Results show that the model exceeds existing approaches in noise immunity and diagnostic performance [19]. While these models achieve a promising balance between lightweight design and noise robustness, they typically require vibration signals to undergo preprocessing steps such as wavelet or Fourier transforms before being input into the diagnostic model. These operations still demand significant computational resources on embedded devices, leading to increased latency and thereby limiting real-time performance and practical deployment on resource-constrained edge platforms.

To tackle these issues, this study introduces a Dual-Branch Lightweight Convolutional Neural Network (DBLCNN), designed for embedded fault diagnosis with limited computational resources. The network features a two-path parallel architecture. One branch processes 1D inputs with residual blocks to enhance the detection of subtle fault-related patterns under low-SNR conditions. The other branch employs 2D depthwise separable convolutional layers to reduce the number of trainable parameters while refining discriminative feature learning. Operating in an end-to-end manner, the model bypasses intricate signal transformation and preprocessing steps, thereby lowering both computational and memory requirements.

2. Bearing fault diagnosis method based on DBLCNN

2.1. Residual block

The residual module is shown in Figure 1, where $H(x) = F(x) + x$, If the network structure is optimal when$F(x) = 0$, then $H(x) = x$,meaning $H(x)$ is an identity map. However, when $H(x) = x$ is not the optimal solution for the neural network, the powerful fitting capability of the neural network allows $F(x)$ to approximate the optimal value as closely as possible. Moreover, fitting the identity mapping$H(x) = x$ is more difficult than fitting the residual mapping $F(x) = H(x) - x$ During the training process of the neural network, in addition to the gradient of the objective function, the residual gradient is also utilized. As a result, residual neural networks can effectively increase network depth, thereby enhancing the feature learning capacity of the neural network.

When the number of input and output channels of the previous and later layers is the same, the residual connection is the identity mapping connection, and the output of the residual block is:

$$y = F(x, W) + x \tag{1}$$

If the number of channels in the input and output of the current posterior layer is inconsistent, it is necessary to use a 1×1 convolution to align the number of channels, and the output of the residual block is:

$$y = F(x, W) + W_s x \tag{2}$$

Where W_x is a 1×1 convolution.

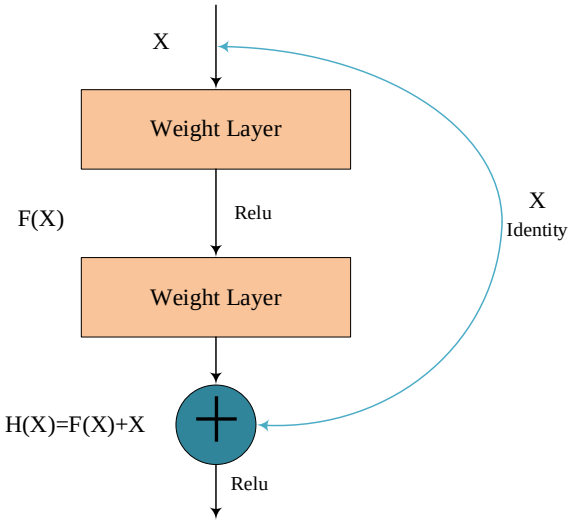


Figure 1. Schematic diagram of a residual block.

2.2. Depthwise separable convolutions

Depthwise Separable Convolution [20], proposed by Laurent Sifre in 2013, forms the core of the MobileNet [21] and Xception [22] architectures. Unlike standard convolution, which correlates all input and output channels together, this operation separately processes each input channel through depthwise convolution and then combines outputs via 1×1 pointwise convolution. This decomposition significantly reduces parameter count while enhancing feature representation capacity. Mathematically, standard convolution can be regarded as a closed-form special case of depthwise separable convolution [23].

In depthwise separable convolution, the process of standard convolution is divided into two steps. First, each input channel is convolved separately using depthwise convolution, and then the output of different depthwise convolution is combined using pointwise convolution. The decomposition process is shown in Figure 2. Assuming that the input feature map is of size$D_F \times D_F \times M$, the output feature map is of size$D_F \times D_F \times N$, and the convolution kernel is of size $D_K \times D_K \times M$, the computation of the standard convolution is as follows:

$$D_K \times D_K \times M \times N \times D_F \times D_F \quad (3)$$

In depthwise separable convolution, the convolution is integrated into two steps. Firstly, Mconvolution kernels of size $D_K \times D_K \times 1$ are used to perform independent depth convolution for each channel to obtain M feature maps, and then N convolution kernels of size $1 \times 1 \times M$ are used to convolve M feature maps to obtain the feature map of $D_F \times D_F \times N$. The computational cost of depthwise separable convolution is as follows:

$$D_K \times D_K \times M \times D_F \times D_F + M \times N \times D_F \times D_F \quad (4)$$

Therefore, the ratio of computation of depthwise separable convolution to standard convolution can be obtained as:

$$\frac{D_K \times D_K \times M \times D_F \times D_F + M \times N \times D_F \times D_F}{D_K \times D_K \times M \times D_F \times D_F} = \frac{1}{N} + \frac{1}{D_K^2} \quad (5)$$

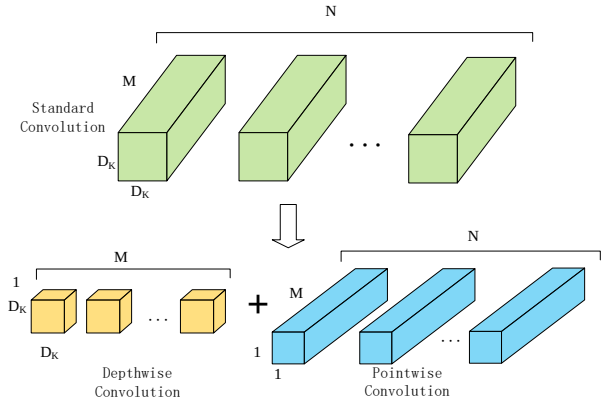


Figure 2. Diagram of depthwise separable convolution.

2.3. Overall structure of DBLCNN model

The proposed DBLCNN-based bearing fault diagnosis method is designed to suppress environmental noise interference during vibration signal acquisition while achieving high fault recognition accuracy. As illustrated in Figure 3, the proposed model utilizes a dual-path structural design: the left branch operates on one-dimensional vibration signals of size 1024×1 through residual building blocks, improving the ability to learn representative features under noisy conditions and mitigating performance degradation in deep networks. Meanwhile, the right branch converts the original one-dimensional input into a two-dimensional 32×32 format and applies depthwise separable convolution to capture spatially distributed characteristics of faults, simultaneously decreasing parameter complexity. To enhance training stability and generalization under noisy settings, batch normalization is applied within each branch of the network. Furthermore, the initial convolutional layer in both pathways employs wider kernel sizes to strengthen robustness against noise and elevate classification performance [24]. Prior to merging the features, global average pooling is utilized to compress the feature maps from both branches, reducing dimensionality while promoting the integration of

multi-source feature information. This process improves data consistency and enriches feature diversity [25].

The architectural configuration and hyperparameter settings of the proposed DBLCNN are outlined in Table 1 for the 1D branch and Table 2 for the 2D branch. Training was conducted using an initial learning rate of 0.001, with the Adam optimizer and the ReLU activation function. The loss function was defined as cross-entropy, and the model was trained for 30 epochs with a batch size of 128.

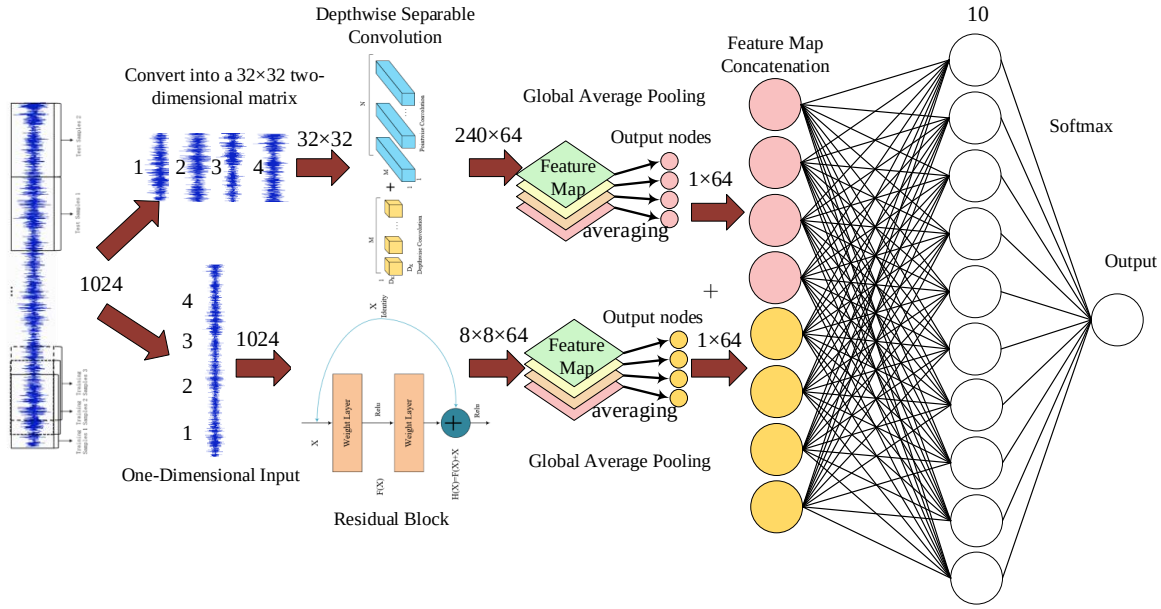


Figure 3. Basic structure of DBLCNN bearing fault diagnosis model.

Table 1. Parameters of left branch of DBLCNN network model.

Network Layer		Number of Kernels	Kernel Size/Stride
Convolutional Layer		16	64×1/1×1
Max Pooling Layer		16	2×1/2×1
	Convolutional Layer	32	1×1/1×1
Residual Block	Convolutional Layer	32	3×1/1×1
	Convolutional Layer	32	1×1/1×1
Feature Stacking Layer		—	—
Max Pooling Layer		32	2×1/2×1
	Convolutional Layer	64	1×1/1×1
Residual Block	Convolutional Layer	64	3×1/1×1
	Convolutional Layer	64	1×1/1×1
Feature Stacking Layer		—	—
GAP Layer		—	—

Table 2. Parameters of the right branch of DBLCNN network model.

Network Layer	Number of Kernels	Kernel Size/Stride
Depthwise Separable Convolutional Layer	32	7×7/1×1
Depthwise Separable Convolutional Layer	32	5×5/1×1
Max Pooling Layer	32	2×2/2×2
Depthwise Separable Convolutional Layer	64	3×3/1×1
Max Pooling Layer	64	2×2/2×2
Depthwise Separable Convolutional Layer	64	3×3/1×1
GAP Layer	—	—
Feature Stacking Layer	—	—
Output Layer	—	—

2.4. The fault diagnosis process based on DBLCNN

The overall workflow for bearing fault diagnosis using the DBLCNN model is illustrated in Figure 4. The process consists of three main stages: data preparation, development and training of the network, and finally the deployment and evaluation of the

fault diagnosis model.

1. Data collection and sample division: In order to prevent over-fitting phenomenon in noisy environment due to too few training samples, a sliding window with a length of 1024 is used to shift and slide the whole original vibration data with a length of 100 steps, and the test samples are not overlapped, that is, the step length is equal to the sample length.
2. During the training phase, samples from the training set are fed directly into the DBLCNN architecture. The model achieving the highest validation accuracy is retained as the optimal diagnostic system.
3. To evaluate the trained diagnostic model, the test set data is input and the resulting predictions are systematically analyzed. This evaluation employs visualization techniques including confusion matrices and feature mapping to interpret the diagnostic outcomes from multiple perspectives.

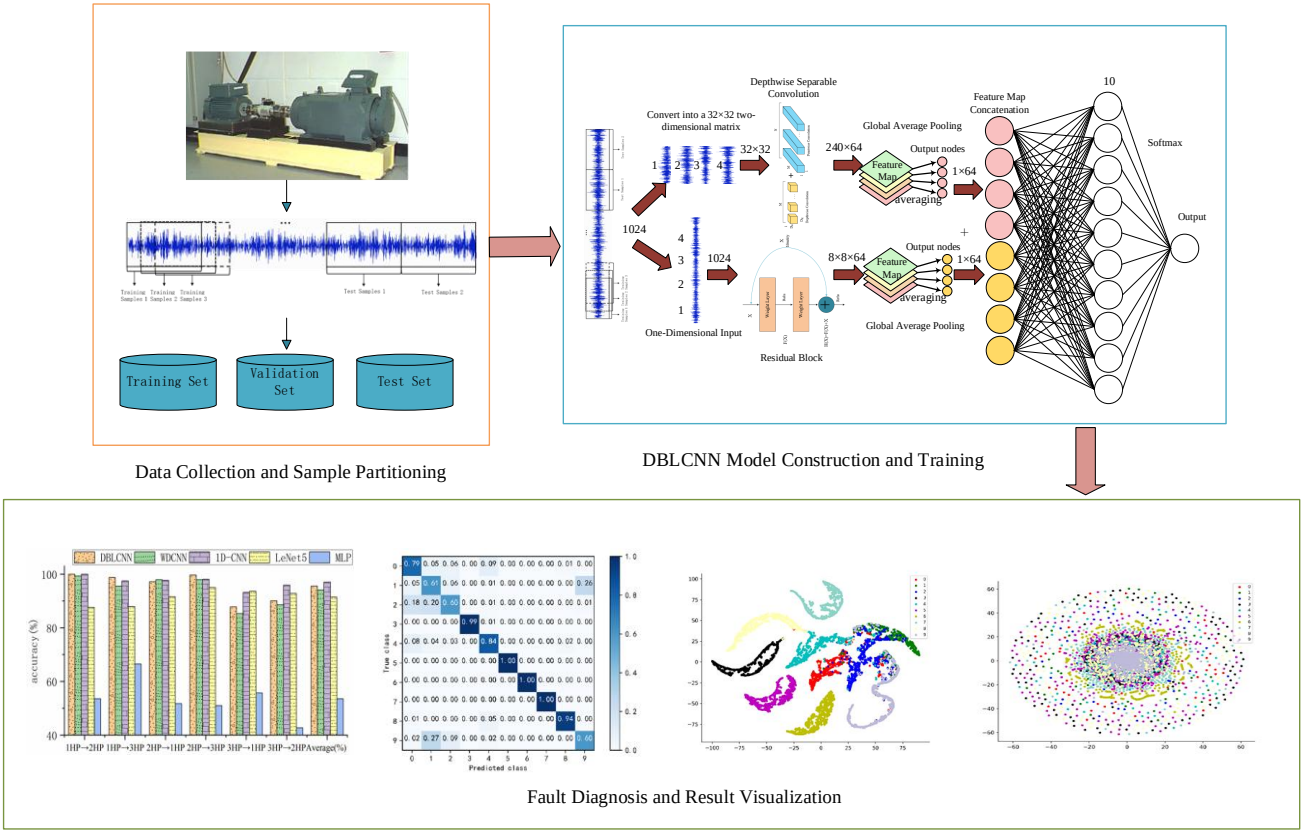


Figure 4. Fault diagnosis process based on DBLCNN.

3. Experimental results and analysis

3.1. Introduction to the dataset

In this paper, the bearing datasets from Case Western Reserve University (CWRU) [26] and Paderborn University (PU) [27] are employed to validate and assess the performance of the proposed DBLCNN model. Next, the CWRU dataset will be introduced first. Its experimental data consist of vibration signals acquired by an accelerometer installed at the drive end, sampled at a frequency of 12 kHz. The dataset encompasses a range of fault conditions, including healthy bearing states, inner race and ball defects with fault diameters of 0.007, 0.014, and 0.021 inches, as well as outer race faults with a damage size of 0.007 inches located at the 3, 6, and 12 o'clock positions. The dataset comprises a total of 10 conditions, including normal operation and multiple fault types.

The data preprocessing method adopted in this paper is shown in Figure 5. By using a sliding window of length 1024 to translate and slide on the whole original vibration data with a step of length 100, the test sample is not overlapped and segmented, that is, the step length is equal to the sample length. To assess the robustness of the proposed model under conditions resembling real industrial noise, Gaussian white noise was added to the original bearing vibration signals [28]. The detailed dataset split is provided in Table 3. The added noise level is quantified by the Signal-to-Noise Ratio (SNR), expressed as:

$$SNR_{db} = 10\,log_{10}\left(\frac{P_{signal}}{P_{noise}}\right) \quad (6)$$

Where P_{signal} and P_{noise} represent the power of the signal and noise, respectively. When the value of SNR is zero, it means that the energy of the signal and noise are basically equal, and a smaller value of SNR means that the signal contains more noise.

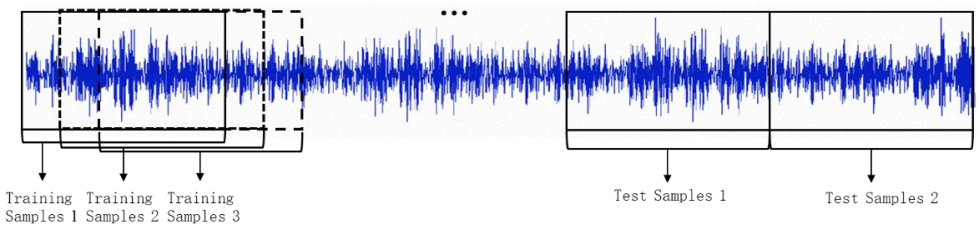


Figure 5. Schematic diagram of bearing vibration data segmentation.

Table 3. CWRU Dataset Partition.

Fault Category / Label	Fault Diameter (inch)	Fault Orientation	Number of Samples			
			Training Set	Validation Set	Test Set	Total
Ball (0)	0.007	—	800	100	100	1000
Ball (1)	0.014	—	800	100	100	1000
Ball (2)	0.021	—	800	100	100	1000
Inner Race (3)	0.007	—	800	100	100	1000
Inner Race (4)	0.014	—	800	100	100	1000
Inner Race (5)	0.021	—	800	100	100	1000
Outer Race (6)	0.007	@3:00	800	100	100	1000
Outer Race (7)	0.007	@6:00	800	100	100	1000
Outer Race (8)	0.007	@12:00	800	100	100	1000
Normal (9)	0	—	800	100	100	1000

Next, we introduce the Paderborn University (PU) bearing dataset. This dataset collects vibration signals from a model 6203 bearing during operation, obtained via a piezoelectric acceleration sensor with a sampling frequency of 64 kHz. This study selects naturally damaged vibration data acquired under experimental conditions including a drive speed of 1500 r/min, a load torque of 0.1 N·m, and a radial load of 1000 N. The damage types primarily include fatigue pitting and plastic indentations. The vibration signals are segmented non-overlappingly into samples of 2500 data points each.

To more effectively evaluate the generalization performance of the model in noisy environments, impulse noise is also introduced into the original bearing vibration signal. This type of noise, characterized by its burst nature, non-continuity, high amplitude, and short duration, can accurately simulate the transient interference commonly encountered in industrial settings. The specific division of the dataset is shown in Table 4.

Table 4. PU dataset partition.

Fault Category	Damage Type	Damage Severity	Bearing Number	Number of Samples			
				Training Set	Validation Set	Test Set	Total
Inner Race	Fatigue Pitting	1	KI14	288	36	36	360
	Fatigue Pitting	3	KI16	288	36	36	360
	Fatigue Pitting	1	KI17	288	36	36	360
	Fatigue Pitting	2	KI18	288	36	36	360
	Fatigue Pitting	1	KI21	288	36	36	360
Outer Race	Fatigue Pitting	1	KA04	288	36	36	360
	Plastic Indentation	1	KA15	288	36	36	360
	Fatigue Pitting	2	KA16	288	36	36	360
	Fatigue Pitting	1	KA22	288	36	36	360
	Plastic Indentation	1	KA30	288	36	36	360
Normal	—	—	K001	1440	180	180	1800

3.2. Analysis of parameter quantity, FLOPs, and inference time for the DBLCNN network model

The experiments were performed on a computational platform equipped with an Intel i5-8265U CPU, 8 GB of RAM, and an NVIDIA MX250 GPU. All models were implemented in Python 3.7 using the TensorFlow 2.2 deep learning framework.

The DBLCNN-based fault diagnosis approach incorporates depthwise separable convolution and global average pooling to significantly decrease parameter complexity. Using the CWRU dataset with Gaussian white noise added at a SNR of -4 dB, the proposed model was compared with two benchmark architectures: WDCNN, which employs a wide first-layer convolutional kernel [25], and the classic LeNet-5 [29]. The assessment focuses on diagnostic accuracy, number of model parameters, FLOPs, and average inference time per sample. To further evaluate the impact of different combinations of global average pooling and depthwise separable convolution on diagnostic performance and model size, multiple comparative experiments are conducted using the CWRU dataset under a 3 HP load condition. The experimental configurations are summarized in Table 5: A denotes the removal of global average pooling from the left 1D branch; B indicates its removal from the right 2D branch; C represents its removal from both branches; and D refers to replacing depthwise separable convolution in the right branch with standard convolution.

As indicated by the comparative experimental results in Table 5, the WDCNN model contains 42,502 parameters, which is significantly fewer than the 61,706 parameters of the classic LeNet-5 architecture. Furthermore, under a -4 dB signal-to-noise ratio, WDCNN attains a fault diagnosis accuracy of 74.50%, surpassing LeNet-5, which achieved only 64.70%. These outcomes confirm that employing a wide convolutional kernel in the initial layer enhances the extraction of discriminative fault-related features in high-noise environments. Building on this insight, the DBLCNN model introduced in this work also integrates a wide kernel design, with corresponding experimental validation discussed in subsequent sections.

In a comparative study examining the impact of different combinations of global average pooling and depthwise separable convolution on model accuracy and parameter count, it was observed that Models A, B, and C, which did not use global average pooling, exhibited significantly lower fault diagnosis accuracy compared to Model D and the DBLCNN model, both of which incorporated this structure. Notably, Model C, which omits global average pooling in both branches, achieved a diagnostic accuracy of 72.61%, higher than Model A at 70.45% and Model B at 70.60%. This phenomenon can be attributed to the fact that when global average pooling is omitted in one branch, that branch's features carry excessive weight in the feature fusion process, thereby diminishing the contribution of the other branch and hindering effective integration of fault features extracted from both one-dimensional and two-dimensional perspectives. In contrast, Model D and the DBLCNN model, which employ global average pooling in both branches, not only achieve significantly higher diagnostic accuracy but also reduce the parameter count and mitigate overfitting.

Compared to Model D, the proposed DBLCNN model reduces the number of parameters by 71.98% and improves fault diagnosis accuracy by 1.31% through the use of depthwise separable convolution. Furthermore, relative to the WDCNN model which has 42,502 parameters and the classic LeNet5 model containing 61,706 parameters, DBLCNN achieves parameter reductions of 32.74% and 53.67% respectively, while increasing accuracy by 10.24% and 20.04%. These results demonstrate that the DBLCNN based bearing fault diagnosis method maintains high fault recognition accuracy with considerably fewer parameters, even under noisy conditions.

Table 5. Comparison of parameters, FLOPs, inference time, and accuracy among different models when SNR is -4dB.

Model	WDCNN	LeNet-5	A	B	C	D	DBLCNN
Params	42502	61706	181547	68907	221867	102138	28587
FLOPs	3.9×10^5	4×10^5	1.2×10^7	1.0×10^7	1.2×10^7	4.1×10^7	9.5×10^6
Inference Time	0.9425ms	0.0121ms	0.0377ms	0.0392ms	0.0377ms	0.0374ms	0.0368ms
Accuracy	74.50%	64.70%	70.45%	70.60%	72.61%	83.43%	84.74%

Based on the comparison of FLOPs among different models in Table 5, the lightweight-designed DBLCNN model significantly outperforms models A, B, C, and D in terms of computational complexity, with its floating-point operations

reduced by an order of magnitude, while also achieving shorter average inference time per sample. In comparison, although the FLOPs of the WDCNN model are lower than those of DBLCNN, its recognition accuracy in a -4dB noisy environment is 10.24 percentage points lower than DBLCNN, and its average inference time per sample is approximately 25 times longer. As for the LeNet-5 model, although both its FLOPs and inference time are lower than those of DBLCNN, its accuracy under -4dB noise conditions is only 64.7%, which is 20.04 percentage points lower than DBLCNN.

3.3. The size of the first convolutional kernel has an impact on noise reduction

To evaluate the impact of the first layer convolution kernel size in both branches of the DBLCNN model on fault bearing recognition accuracy in noisy environments, experiments were conducted under eight different signal-to-noise ratios ranging from -4 dB to 10 dB in 2 dB increments, into the CWRU dataset. The kernel size of the right two-dimensional branch was fixed at 7×7 while testing the left one-dimensional branch, and the kernel size of the left one-dimensional branch was fixed at 64×1 while testing the right two-dimensional branch. Experimental results are presented in Tables 6 and 7.

As shown in Table 6, when the left one-dimensional branch uses a small convolution kernel in the first layer, the recognition accuracy of the DBLCNN model is relatively low. Specifically, at an SNR of -4 dB, the fault diagnosis accuracy is 79.80% with a first-layer kernel size of 16×1, compared to 84.74% with a 64×1 kernel—an improvement of 4.94%. However, when the kernel size is further increased to 128×1, the accuracy drops to 81.64%, which is 3.10% lower than that achieved with the 64×1 kernel. Combined with experimental results at other signal-to-noise ratios presented in Table 6, it can be concluded that both excessively small and overly large kernel sizes in the first layer lead to reduced fault diagnosis accuracy under low SNR conditions. Selecting an appropriately sized convolution kernel helps balance time-domain resolution and feature extraction capability, thereby retaining most fault-related information while minimizing the loss of subtle features in high-noise environments.

As shown in Table 7, the use of a 7×7 convolutional kernel in the first-layer leads to consistently high performance across five comparative trials, with diagnostic accuracy surpassing 99% at signal-to-noise ratios of 4 dB and above. Both Table 6 and Table 7 reveal that variations in the size of the first-layer kernel have minimal impact on the overall accuracy, as the maximum deviation observed is merely 4.94%. This consistency can be explained by the experimental strategy of maintaining a fixed moderate kernel size in one branch during the optimization of the other. The findings affirm that the DBLCNN model successfully enhances bearing fault identification by effectively combining multi-perspective features learned from both one-dimensional and two-dimensional processing paths.

Table 6. Influence of convolution kernel size in the first layer of one-dimensional branch on accuracy in different SNR cases.

Convolution kernel size	**SNR（dB）**							
	-4	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
16	79.80%	87.52%	92.06%	94.62%	98.12%	99.36%	99.36%	99.80%
40	80.22%	88.08%	94.70%	97.62%	99.34%	99.76%	99.90%	**100%**
64	**84.74%**	90.75%	95.55%	**98.55%**	99.50%	99.84%	**100%**	**100%**
96	84.40%	**90.76%**	**95.94%**	98.24%	99.56%	**99.86%**	**100%**	**100%**
128	81.64%	88.74%	95.34%	98.20%	**99.58%**	99.84%	**100%**	**100%**

Table 7. Influence of convolution kernel size in the first layer of 2D branch on accuracy in different SNR cases.

Convolu-tion kernel size	**SNR（dB）**							
	-4	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
3×3	80.64%	87.80%	94.40%	97.08%	98.22%	**99.84%**	99.90%	**100%**
5×5	82.52%	90.10%	94.58%	97.86%	98.88%	99.66%	99.96%	**100%**
7×7	**84.74%**	**90.75%**	95.55%	**98.55%**	**99.50%**	**99.84%**	**100%**	**100%**
9×9	82.94%	90.02%	95.20%	97.70%	99.18%	99.60%	99.92%	**100%**
11×11	82.74%	90.44%	**95.76%**	98.48%	99.18%	99.74%	99.92%	**100%**

3.4. Comparison of noise immunity of different network models under different SNR

To further validate the effectiveness of the proposed model, impulse noise was added to the PU dataset, and the proposed

DBLCNN fault diagnosis model was compared with several established methods across 8 SNR conditions ranging from -4 dB to 10 dB. These methods include the WDCNN model featuring a wide first-layer convolution kernel of size 96×1, the classic LeNet-5 model, and four additional models: 1D-CNN, MLP, KNN, and RF [30]. The results are shown in Table 8, which lists the mean accuracy (M_acc), standard deviation (SD), and 95% confidence interval (CI) for each model over 10 independent runs.

Table 8. Comparison of standard deviation and confidence intervals of accuracy among different models at various signal-to-noise ratios in the PU dataset.

Model		SNR							
		-4	-2	0	2	4	6	8	10
WDCNN	M_acc	98.6%	99.14%	99.57%	99.77%	99.79%	99.78%	99.82%	99.93%
	SD	0.14%	0.07%	0.08%	0.08%	0.06%	0.05%	0.03%	0.03%
	CI	[97.96%-98.16%]	[99.09%-99.20%]	[99.51%-99.62%]	[99.72%-99.82%]	[99.75%-99.84%]	[99.74%-99.82%]	[99.80%-99.85%]	[99.91%-99.95%]
1D-CNN	M_acc	83.88%	85.04%	85.84%	86.18%	86.36%	86.36%	86.42%	86.42%
	SD	0.22%	0.15%	0.14%	0.10%	0.05%	0.07%	0.02%	0.02%
	CI	[83.73%-84.04%]	[84.93%-85.15%]	[85.75%-85.94%]	[86.11%-86.25%]	[86.32%-86.39%]	[86.31%-86.41%]	[86.41%-86.44%]	[86.40%-86.44%]
LeNet-5	M_acc	79.41%	79.73%	80.23%	79.85%	85.96%	86.39%	86.65%	86.65%
	SD	0.06%	0.09%	0.16%	0.04%	0.09%	0.03%	0.01%	0.02%
	CI	[79.37%-79.45%]	[79.67%-79.79%]	[80.11%-80.34%]	[79.82%-79.88%]	[85.90%-86.03%]	[86.37%-86.42%]	[86.64%-86.66%]	[86.64%-86.66%]
MLP	M_acc	74.26%	76.73%	78.95%	89.33%	93.87%	96.54%	97.51%	98.17%
	SD	0.27%	0.26%	0.18%	0.41%	0.22%	0.23%	0.18%	0.10%
	CI	[74.06%-74.45%]	[76.54%-76.92%]	[78.82%-79.08%]	[89.04%-89.61%]	[93.71%-94.03%]	[96.38%-96.71%]	[97.38%-97.64%]	[98.10%-98.25%]
KNN	M_acc	33.59%	33.76%	34.07%	34.17%	34.37%	34.43%	34.57%	34.81%
	SD	1.43%	1.28%	1.38%	1.50%	1.64%	1.73%	0.42%	0.28%
	CI	[32.57%-34.62%]	[32.84%-34.68%]	[33.09%-35.06%]	[33.09%-35.24%]	[33.20%-35.54%]	[33.19%-35.66%]	[34.27%-34.87%]	[34.62%-35.01%]
RandomForest	M_acc	41.57%	46.41%	49.46%	53.57%	57.63%	61.39%	65.84%	67.99%
	SD	1.34%	2.18%	1.86%	2.76%	1.77%	1.18%	2.20%	2.04%
	CI	[40.61%-42.53%]	[44.86%-47.97%]	[48.13%-50.80%]	[51.78%-55.71%]	[56.37%-58.90%]	[60.54%-62.23%]	[64.27%-67.41%]	[66.53%-69.44%]
DBLCNN	M_acc	90.20%	93.56%	97.89%	98.55%	99.64%	99.82%	99.92%	99.99%
	SD	0.36%	0.33%	0.21%	0.17%	0.06%	0.05%	0.02%	0.02%
	CI	[89.95%-90.46%]	[93.32%-93.80%]	[97.74%-98.04%]	[98.42%-98.67%]	[99.59%-99.68%]	[99.79%-99.85%]	[99.92%-99.95%]	[99.97%-100.00%]

Overall, the classification accuracy of all models improved with increasing SNR. Deep learning models, especially WDCNN and the proposed DBLCNN, outperformed traditional machine learning models such as KNN and Random Forest at all SNRs. DBLCCN demonstrated outstanding anti-noise performance and stability: at the low SNR of -4 dB, its accuracy already reached 90.20%, second only to WDCNN. As the SNR increased, the accuracy continued to rise, reaching 99.99% at 10 dB.

At medium-to-high SNRs ($\geqslant 0$ dB), the accuracy of DBLCNN exceeded 97%, and at SNR $\geqslant 6$ dB, it performed better than WDCNN. Although the WDCNN model exhibits high accuracy in certain low signal-to-noise ratio segments of the PU dataset, the DBLCNN model demonstrates higher accuracy than the WDCNN model on the CWRU dataset with Gaussian white noise at a signal-to-noise ratio of -4 dB. This indicates that DBLCNN has better adaptability and generalization ability across different datasets and noise environments.

In summary, DBLCNN performed excellently in various noisy environments, with particularly strong robustness at low SNRs. Moreover, in cross-dataset experiments, it demonstrated stronger generalization performance than WDCNN, making it suitable for classification tasks under complex noise conditions.

3.5. Experimental comparison of left and right branches of DBLCNN model

To validate the effectiveness of the DBLCNN dual-branch structure, this section introduces two typical types of noise interference, Gaussian white noise and impulse noise, into the CWRU dataset and conducts a comparative analysis of the fault diagnosis performance of the one-dimensional branch, the two-dimensional branch, and the complete DBLCNN model. As shown in Table 9, under harsh low SNR conditions ranging from -4 dB to 0 dB, the DBLCNN model demonstrates significantly superior overall robustness compared to any single branch.

Table 9. Compares the recognition accuracy of the left and right branches and the DBCNN model in a noisy environment.

Model	Noise type	SNR（dB）							
		-4	-2	0	2	4	6	8	10
1D Branch	**Gaussian white noise**	77.50%	85.10%	92.00%	95.60%	99.2%	99.40%	99.59%	99.90%
2D Branch		73.70%	82.00%	88.30%	94.20%	97.4%	99.20%	99.70%	99.90%
DBLCNN		84.74%	90.75%	95.55%	98.55%	99.5%	99.84%	100%	100%
1D Branch	**impulse noise**	78.59%	84.56%	91.01%	95.38%	97.86%	97.94%	99.91%	99.44%
2D Branch		64.19%	72.56%	82.47%	87.52%	91.4%	96.72%	98.82%	99.00%
DBLCNN		80.19%	87.69%	91.44%	95.50%	98.53%	99.55%	99.60%	99.92%

In the Gaussian white noise environment, DBLCNN performs particularly well. When the SNR is as low as -4 dB, the diagnostic accuracy of the model reaches 84.74%, which is 7.24 and 11.04 percentage points higher than that of the one-dimensional branch and the two-dimensional branch, respectively. As the SNR increases to 0 dB, DBLCNN maintains a stable lead with an accuracy of 95.55% and approaches 100% accuracy more rapidly, reflecting its excellent noise adaptation capability.

In the impulse noise environment, the complementary effect of the dual branches is more pronounced. The performance of the two-dimensional branch is significantly affected by impulse noise, with an accuracy of only 64.19% at -4 dB, while the one-dimensional branch is relatively robust, achieving 78.59%. However, DBLCNN, by fusing the features of both branches, improves the accuracy to 80.19% under the same conditions, clearly surpassing the diagnostic results of the individual branches.

In summary, DBLCNN achieves the highest diagnostic accuracy and stability under both noise conditions, with its advantages being particularly evident in the low SNR range. This indicates that the dual-branch architecture is not merely a simple structural superposition but achieves information complementarity through effective feature fusion: the one-dimensional branch exhibits good tolerance to noise, while the two-dimensional branch can deeply extract fault features that the one-dimensional branch fails to capture. Consequently, the fused DBLCNN can adaptively enhance robust features, suppress noise interference, and comprehensively improve the model's practicality and reliability in complex noise environments.

3.6. Performance of DBLCNN model under variable operating conditions

In this section, the generalization ability of DBLCNN bearing fault diagnosis model under variable working conditions in a noise-free environment on the CWRU dataset is first experimentally evaluated, and the test results are compared with WDCNN model, LeNet-5 model, 1D-CNN model and MLP model, as shown in Figure 6.

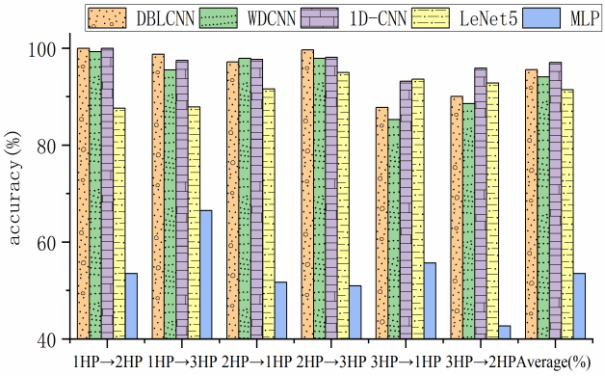


Figure 6. Comparison of the accuracy rates of different models under variable working conditions in a noise-free environment.

As illustrated in Figure 6, the DBLCNN model achieves an average fault diagnosis accuracy of 95.58%. In comparison, the MLP model yields the lowest accuracy across all variable operating condition experiments, with a result of only 42.7% under the 3HP→2HP transfer scenario and an overall average of 53.52%—more than 40 percentage points lower than that of

the DBLCNN model. The WDCNN and LeNet-5 models attain average accuracies of 94.08% and 91.42%, respectively, trailing the proposed model by 1.5% and 4.16%. Although the 1D-CNN model shows a slightly higher average accuracy of 97.07%, exceeding the DBLCNN by 1.49%, the latter demonstrates significantly superior recognition performance under low signal-to-noise ratio conditions, as established in Section 3.4. Thus, the DBLCNN exhibits more balanced and desirable overall characteristics compared to the 1D-CNN.

To further validate the model's cross-condition generalization ability in noisy environments, experiments were conducted using the CWRU dataset with added 0dB Gaussian white noise for comparison. The results are shown in Table 10, which presents the average accuracy (M_acc), standard deviation (SD), and 95% confidence interval (CI) from 10 independent runs of each model. From Table 10, it can be observed that in the two migration tasks of 3HP→1HP and 3HP →2HP, the diagnostic accuracy of the DBLCNN model is 90.37% and 90.44%, respectively, both higher than those of other comparison models. Although in the other four migration tasks, the WDCNN model achieves higher accuracy than DBLCNN, its number of parameters and single-sample inference time are significantly higher than those of DBLCNN. Additionally, under -4dB Gaussian white noise conditions, WDCNN's accuracy is only 74.50%, lower than DBLCNN's 84.7% (see Table 5). Moreover, the average accuracies of the five models—1D-CNN, LeNet-5, MLP, KNN, and Random-Forest—across various migration tasks are all lower than that of the DBLCNN model. Therefore, considering parameters, inference efficiency, and diagnostic accuracy comprehensively, DBLCNN achieves a better balance among parameter efficiency, computational speed, and classification performance, demonstrating overall superior performance compared to WDCNN.

In summary, the DBLCNN model not only exhibits high accuracy and a compact structure under noise-free conditions but also maintains high cross-load recognition accuracy, good stability, and generalization ability in 0dB noisy environments. This fully demonstrates its excellent noise robustness and adaptability to varying working conditions. The model achieves a good balance among parameter efficiency, anti-interference capability, and inference speed, making it suitable for practical deployment and application in complex dynamic industrial scenarios.

Table 10. Based on the CWRU dataset, the average accuracy, standard deviation, and 95% confidence interval results of different models across conditions under 0 dB Gaussian white noise are compared.

Model		SNR					
		1HP→2HP	1HP→3HP	2HP→1HP	2HP→3HP	3HP→1HP	3HP→2HP
WDCNN	M_acc	99.57%	96.51%	96.98%	99.14%	85.20%	89.18%
	SD	0.10%	0.20%	0.17%	0.12%	0.26%	0.22%
	CI	[99.49%-99.64%]	[96.36%-96.65%]	[96.86%-97.10%]	[99.06%-99.23%]	[85.02%-85.39%]	[89.02%-89.33%]
1D-CNN	M_acc	86.99%	82.40%	87.68%	82.71%	81.01%	80.92%
	SD	0.60%	0.43%	0.49%	0.56%	0.17%	0.25%
	CI	[86.56%-87.42%]	[82.09%-82.71%]	[87.33%-88.03%]	[82.31%-83.11%]	[80.89%-81.13%]	[80.74%-81.10%]
LeNet-5	M_acc	56.53%	59.46%	67.70%	71.25%	58.46%	56.13%
	SD	1.56%	0.77%	1.17%	0.33%	0.57%	0.91%
	CI	[55.42%-57.65%]	[58.91%-60.01%]	[66.86%-68.54%]	[71.02%-71.49%]	[58.05%-58.87%]	[55.48%-56.79%]
MLP	M_acc	42.70%	46.48%	27.33%	27.00%	42.88%	32.06%
	SD	1.63%	1.20%	1.68%	1.19%	1.63%	0.93%
	CI	[41.54%-43.86%]	[45.62%-47.34%]	[26.13%-28.53%]	[26.15%-27.85%]	[41.72%-44.04%]	[31.39%-32.73%]
KNN	M_acc	9.62%	10.04%	9.48%	10.04%	9.48%	9.62%
	SD	0.47%	0.56%	0.73%	0.56%	0.73%	0.47%
	CI	[9.28%-9.96%]	[9.64%-10.44%]	[8.95%-10.01%]	[9.64%-10.44%]	[8.95%-10.01%]	[9.28%-9.96%]
RandomForest	M_acc	37.93%	38.13%	36.45%	35.27%	37.32%	36.25%
	SD	1.08%	1.19%	1.05%	1.48%	1.35%	0.93%
	CI	[37.16%-38.71%]	[37.28%-38.98%]	[35.70%-37.20%]	[34.22%-36.33%]	[36.36%-38.29%]	[35.59%-36.92%]
DBLCNN	M_acc	92.23%	87.49%	91.37%	89.17%	90.37%	90.44%
	SD	0.53%	0.85%	0.71%	0.42%	0.58%	0.67%
	CI	[91.94%-92.70%]	[86.88%-88.10%]	[90.87%-91.87%]	[88.87%-89.47%]	[89.97%-90.77%]	[89.96%-90.92%]

3.7. Visual analytics

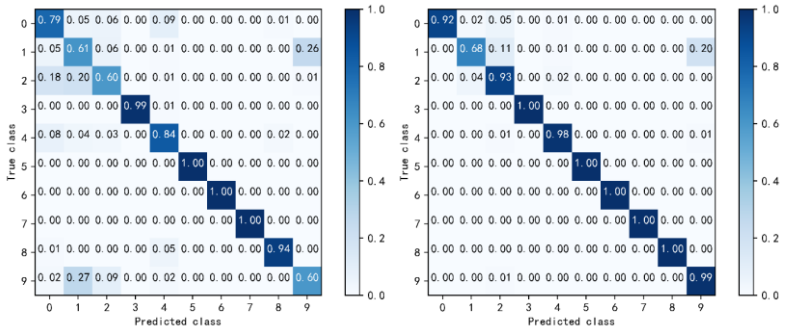
In order to visually evaluate the diagnostic performance of the DBLCNN model for different bearing fault types under different signal-to-noise ratios, Gaussian white noise with SNR of -4 dB, 0 dB, 4 dB, and 10 dB was added to the 3HP load

dataset. The prediction results were then displayed using a confusion matrix. In Figure 7 the abscissa denotes the model's predicted labels, and the ordinate corresponds to the actual labels of the test samples.

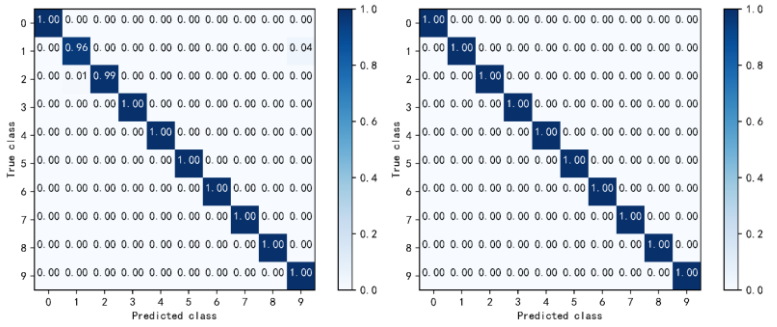
Figure 7 shows the confusion matrix results of the DBLCNN model under Gaussian white noise with different signal-to-noise ratios. It can be observed that the classification performance of the model for bearing conditions varies with the noise level. At a signal-to-noise ratio of -4 dB (Figure 7(a)), the DBLCNN model achieves high recognition accuracy for inner and outer race faults, but noticeable confusion occurs between rolling element faults and the healthy state. Specifically, for rolling element fault samples with a fault diameter of 0.021 inches (Label 2), 18% are misclassified as rolling element faults with a fault diameter of 0.007 inches (Label 0), and 20% are misclassified as rolling element faults with a fault diameter of 0.014 inches (Label 1). Meanwhile, 26% of the rolling element fault samples with a fault diameter of 0.014 inches (Label 1) are incorrectly identified as the healthy state (Label 9). Additionally, in terms of the recognition of the healthy state, the false positive probability of being misclassified as a fault is 0.4, with the probability of being misclassified as a rolling element fault with a fault diameter of 0.014 inches reaching 0.27. On the other hand, the false negative probability of fault samples being misclassified as healthy is 0.26, primarily originating from the misclassification of rolling element faults with a fault diameter of 0.014 inches.

As the signal-to-noise ratio increases, the aforementioned confusion gradually alleviates. At a signal-to-noise ratio of 0 dB (Figure 7(b)), the proportion of 0.014-inch rolling element faults misclassified as healthy decreases to 20%. When the signal-to-noise ratio further increases to 4 dB and 10 dB, the classification performance of the model for various bearing conditions improves significantly, and the overall recognition results tend to be accurate. In summary, the DBLCNN model performs stably under high signal-to-noise ratio conditions. Even under strong noise interference of -4 dB, it maintains good identification capability for inner and outer race faults. This indicates that the model exhibits good effectiveness and robustness in noisy environments.

The confusion matrix shown in Figure 7(b) when the SNR is 0dB is also when the diameter of the ball fault is 0.014 inches, 20% of the samples are identified as normal; When the signal-to-noise ratio is 4dB and 10dB, the DBLCNN model can classify various bearing states more accurately.



(a) -4dB (b) 0dB



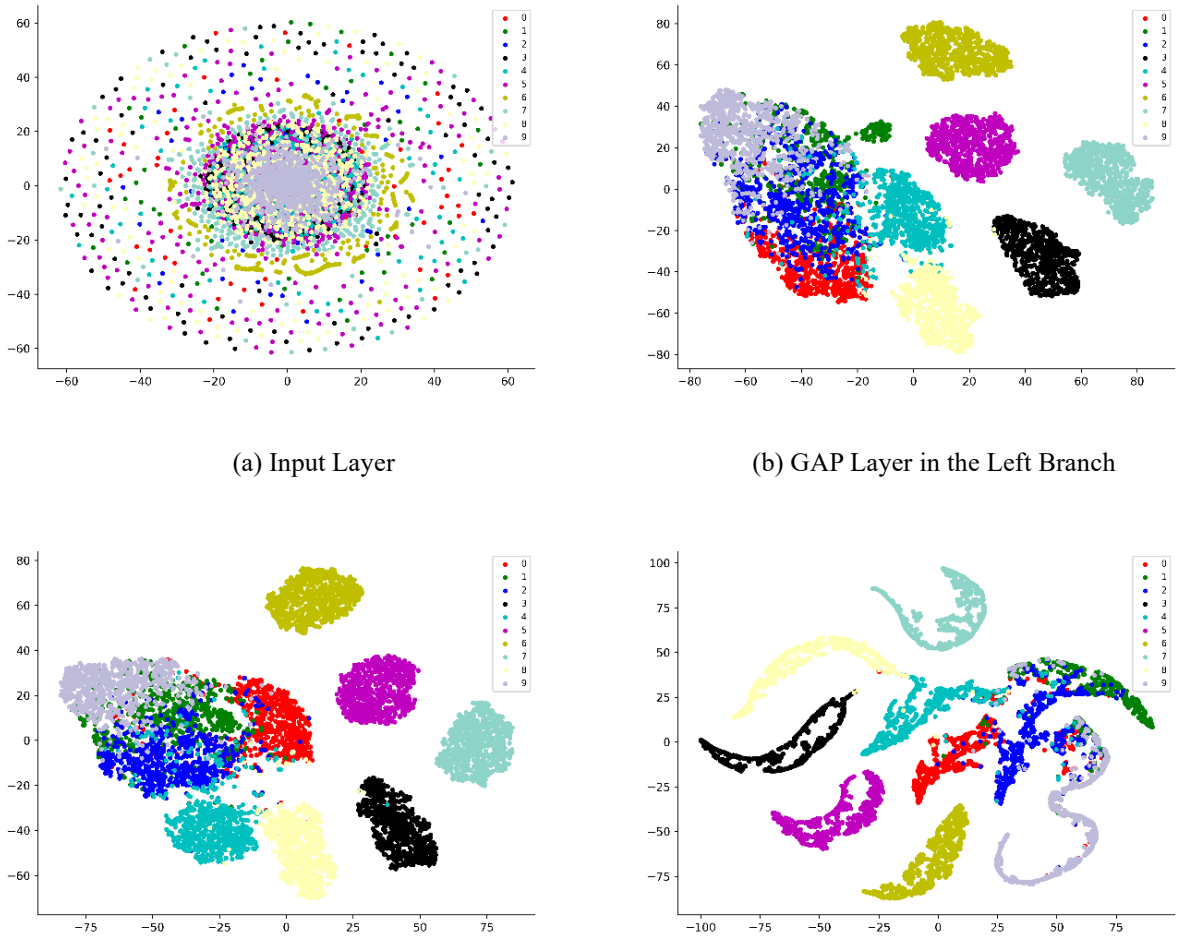
(c) 4dB (d) 10dB

Figure 7. Confusion matrix under different signal-to-noise ratios.

Next, the input layer in the left one-dimensional branch of the DBCNN model, the global average pooling layer, and the feature fusion layer and output layer of the DBCNN model are selected when the SNR is -4dB and 3HP load, and the t-SNE algorithm is used for visualization, and the results are shown in Figure 8.

It can be seen from Figure 8 (a) that the input data are not separable and scattered together. After the feature extraction of the left one-dimensional residual branch, as shown in Figure 8 (b), only the bearing fault types under ball and normal conditions are clustered together, and the DBLCNN model has been able to distinguish the fault states of inner and outer rings.

In Figure 8 (c), after fusing the fault features extracted by the right depth separable convolution, it is found that although the ball fault state and the normal state are still mixed together, the aggregation effect is significantly better than that in Figure 8 (b). This also proves that the DBLCNN fault diagnosis model can fuse the fault features extracted from the two branches to improve the classification accuracy. As can be seen in Figure 8 (d), after the output layer, the bearing state types of ball fault and normal state can be distinguished obviously.



(a) Input Layer

(b) GAP Layer in the Left Branch

(c) Feature Fusion Layer

(d) Output Layer

Figure 8. t-SNE Visualization Results of Different Layers in the DBCNN Model at -4 dB.

4. Discussion

To address the challenge that bearing vibration signals are highly susceptible to noise in real industrial environments, this paper proposes a fault diagnosis model named DBLCNN. The architecture employs a dual-branch design: a one-dimensional residual branch enhances the discriminative power of feature learning under low signal-to-noise ratio conditions, while a two-dimensional branch utilizing depthwise separable convolution improves feature representation capacity. This structure substantially reduces the number of parameters while preserving high diagnostic performance. Validation results indicate that the proposed model achieves a fault recognition

accuracy of 84.74% at –4 dB SNR, outperforming reference models including 1D-CNN, WDCNN, and LeNet-5. In the variable operating conditions test in a noise-free environment, the model achieved an average accuracy of 95.5%. After adding 0 dB Gaussian white noise, its average accuracy under variable operating conditions was 90.17%, which is slightly lower than that of the WDCNN model. However, when considering parameter scale and inference efficiency comprehensively, the DBLCNN model contains only 28,587 parameters, fewer than the 42,502 parameters of the WDCNN model. At the same time, DBLCNN's average inference time per sample is only 0.0368ms, significantly outperforming WDCNN's 0.9425ms. Therefore, the DBLCNN model still demonstrates clear advantages in overall performance. Furthermore, visual analysis of the learned features corroborates the effectiveness of the model in extracting representative fault characteristics. This research offers a high-accuracy and low-complexity solution for bearing fault diagnosis in high-noise industrial applications.

Although the current DBLCNN model has achieved good results, there is still room for optimization of the information interaction and feature fusion mechanism between its dual branches. Future work will focus on exploring the introduction of channel and spatial attention mechanism or Feature Pyramid Network (FPN) to achieve more efficient weighting and fusion of multi-branch features, so as to improve the perception and focus ability of the model on key fault components. Furthermore, research will be conducted on multi-source sensor information fusion technology, followed by its deployment for real-time fault detection on embedded edge devices.

Acknowledgement

This work was funded by the Science and Technology Research Project of Henan Province (Grant Nos. 242102321138、252102210012 and 252102220021), the National Natural Science Foundation of China (Grant No. 62203203), and the Key Research and Development Program of Henan Province (Grant No. 241111223300 and 251111220600), Key Scientific Research Project of Higher Education Institutions in Henan Province (Grant No. 25CY024 and 25A520045).

References

1. Zhang X, Wan S, He Y, Wang X, Dou L. Bearing fault diagnosisbased on iterative 1.5-dimensional spectral kurtosis. IEEE Access 2020; 8: 174233–174243. https://doi.org/10.1109/ACCESS.2020.3024697.
2. Nayana B, Geethanjali P. Analysis of Statistical Time-Domain Features Effectiveness in Identification of Bearing Faults From Vibration Signal. IEEE Sensors 2017; 17(17): 5618–5625. https://doi.org/10.1109/JSEN.2017.2727638.
3. Li D. Research of Fault Diagnosis of Mine Rolling Bearing Based on Time-Frequency Analysis. Journal of Physics: Conference Series, 2022 8th International Forum on Manufacturing Technology and Engineering Materials (IFEMMT 2022) 14/10/2022 - 16/10/2022 Chongqing, China. https://doi.org/10.1088/1742-6596/2459/1/012081.
4. Saravanan N, Kumar Siddabattuni V, Ramachandran K. A comparative study on classification of features by SVM and PSVM extracted using Morlet wavelet for fault diagnosis of spur bevel gear box. Expert Systems with Applications 2008; 35(3): 1351–1366. https://doi.org/10.1016/j.eswa.2007.08.026.
5. Li J, Yao X, Wang X, Yu Q, Zhang Y. Multiscale local features learning based on BP neural network for rolling bearing intelligent fault diagnosis. Measurement 2020; 153: 107419. https://doi.org/10.1016/j.measurement.2019.107419.
6. Sadough Vanini Z, Khorasani K, Meskin N. Fault detection and isolation of a dual spool gas turbine engine using dynamic neuralnetworks and multiple model approach. Information Sciences 2014; 259: 234–251. https://doi.org/10.1016/j.ins.2013.05.032.
7. Wang Y, Tang B, Qin Y, Huang T. Rolling bearing fault detection of civil aircraft engine based on adaptive estimation of instantaneous angular speed. IEEE Transactions on Industrial Informatics 2020; 16(7): 4938–4948. https://doi.org/10.1109/TII.2019.2949000.
8. Fu G, Wei Q, Yang Y. Bearing fault diagnosis with parallel CNN and LSTM. Mathematical Biosciences and Engineering 2024; 21: 2385–2406. https://doi.org/10.3934/mbe.2024105.
9. Ma Y, Jia X, Bai H, Liu G, Wang G, Guo C, Wang S. A new fault diagnosis method based on convolutional neural network and compressive sensing. J Mech Sci Technol 2019; 33: 5177–5188. https://doi.org/10.1007/s12206-019-1007-5.
10. Lu C, Wang Z, Zhou B. Intelligent fault diagnosis of rolling bearing using hierarchical convolutional network based health state classification. Advanced Engineering Informatics 2017; 32: 139–151. https://doi.org/10.1016/j.aei.2017.02.005.

11. Liu H, Yao D, Yang J, Li X. Lightweight convolutional neural network and its application in rolling bearing fault diagnosis under variable working conditions. Sensors 2019; 19: 4827. https://doi.org/10.3390/s19224827.
12. Jin Y, Chen C, Zhao S. Multisource data fusion diagnosis method of rolling bearings based on improved multiscale CNN. Journal of Sensors 2021;2251530: 1–17. https://doi.org/10.1155/2021/2251530.
13. Xing Z, Zhao R, Wu Y, He T. Intelligent fault diagnosis of rolling bearing based on novel CNN model considering data imbalance. Applied Intelligence 2022; 52: 16281–16293. https://doi.org/10.1007/s10489-022-03196-x.
14. Muratbakeev E, Novak, Kozhubaev Y. Investigation of Industrial Bearing Fault Diagnosis Based on 1D-Cnn-Lstm. 2025 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), Sochi, Russian Federation 2025; 1059–1067. https://doi.org/10.1109/ICIEAM65163.2025.11028381.
15. Wang J, Li H, Liu X, Sun B, Liu Y. Bearing fault diagnosis method using CNN with denoising structure under strong noise background. Measurement Science and Technology 2025; 36(3). 10.1088/1361-6501/adaa0a.
16. Ni Y, Li S, Guo P. Discrete wavelet integrated convolutional residual network for bearing fault diagnosis under noise and variable operating conditions. Scientific Reports 2025; 15: 16185. https://doi.org/10.1038/s41598-025-99346-5.
17. Yuan B, Li Y, Chen S. Efficient Gearbox Fault Diagnosis Based on Improved Multi-Scale CNN with Lightweight Convolutional Attention. Sensors 2025; 25: 2636. https://doi.org/10.3390/s25092636.
18. Zhang X, Zhang X, Liang W, He F. Research on rolling bearing fault diagnosis based on parallel depthwise separable ResNet neural network with attention mechanism. Expert Systems With Applications 2025; 286: 128105. https://doi.org/10.1016/j.eswa.2025.128105.
19. Cai S, Wang Y, Cai W, Mo Y, Wei L. One-dimensional time-frequency dual-channel visual transformer for bearing fault diagnosis under strong noise and limited data conditions. Scientific Reports 2025; 15: 26361. https://doi.org/10.1038/s41598-025-12533-2.
20. Sifre L, Mallat S. Rigid-Motion Scattering for Texture Classification. arXiv 2014; 3559: 501-515. https://doi.org/10.48550/arXiv.1403.1687.
21. Howard A, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Hartwig A. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. ArXiv 2017. https://doi.org/10.48550/arXiv.1704.04861.
22. Chollet F. Xception: Deep Learning with Depthwise Separable Convolutions. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017; 1251–1258. https://doi.org/10.48550/arXiv.1610.02357.
23. Guo J, Li Y, Lin W, Chen Y, Li J. Network Decoupling: From Regular to Depthwise Separable Convolutions. 29the British Machine Vision Conference (BMVC 2018), Newcastle, UK, 2018. https://doi.org/10.48550/arXiv.1808.05517.
24. Zhang W, Peng G, Li C, Chen Y, Zhang Z. A New Deep Learning Model for Fault Diagnosis with Good Anti-Noise and Domain Adaptation Ability on Raw Vibration Signals. Sensors 2017; 17(3): 425. https://doi.org/10.3390/s17020425.
25. Wang D, Guo Q, Song Y, Gao S, Li Y. Application of Multiscale Learning Neural Network Based on CNN in Bearing Fault Diagnosis. Journal of Signal Processing Systems 2019; 91(10): 1205–1217. https://doi.org/10.1007/s11265-019-01461-w.
26. Smith W, Randall R. Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study. Mechanical Systems and Signal Processing 2015; 64-65: 100–131. https://doi.org/10.1016/j.ymssp.2015.04.021.
27. Lessmeier C, Kimotho J K, Zimmer D, Sextro W. Condition Monitoring of Bearing Damage in Electromechanical Drive Systems by Using Motor Current Signals of Electric Motors: A Benchmark Data Set for Data-Driven Classification// European Conference of the Prognostics and Health Management Society 2016. https://doi.org/10.36001/phme.2016.v3i1.1577.
28. Chen X, Zhang B, Gao G. Bearing fault diagnosis base on multi-scale CNN and LSTM model. Journal of Intelligent Manufacturing 2021; 32(4): 971–987. https://doi.org/10.1007/s10845-020-01600-2.
29. Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proceedings of the IEEE 1998; 86(11): 2278–2324. https://doi.org/10.1109/5.726791.
30. Chen C-C, Liu Z, Yang G, Wu C-C, Ye Q. An Improved Fault Diagnosis Using 1D-Convolutional Neural Network Model. Electronics 2021; 10: 59. https://doi.org/10.3390/electronics10010059.