# A multi-channel signal fault diagnosis method based on dynamic weighted data fusion and multi-scale feature enhancement

Ke Chen [a,b], Feilong Zhou [a,b], Fangfang Zhang [c], Kunjie Yu [a,b,*], Duo Yang [a]

[a] *School of Electrical and Information Engineering, Zhengzhou University, 450001, Zhengzhou, China*
[b] *Longmen Laboratory, 471000, Luoyang, China*
[c] *School of Engineering and Computer Science, Victoria University of Wellington, 6140, Wellington, New Zealand*

## ARTICLE INFO

## ABSTRACT

The extensive utilisation of rotating machinery has led to escalated demands for effective fault diagnosis method-ologies. Current deep learning methods encounter numerous challenges in practical applications. These include the inability of single-sensor data to comprehensively capture system-level fault information, susceptibility to noise and environmental changes, and the lack of effective mechanisms in multi-sensor fusion methods to address data conflicts and uncertainties. These limitations result in diagnostic accuracy and robustness being compro-mised. To address these issues, this paper proposes a Multi-channel Data and Feature Fusion Diagnostic method (DWDF-MFE) that tackles these challenges through data-level fusion, feature-level fusion, and dynamic feature optimization. Initially, Dempster-Shafer (DS) evidence theory is employed to dynamically compute the weights of multi-sensor signals, resolving signal conflicts and uncertainties while enhancing the reliability of data fusion. Secondly, a Multi-Scale Convolutional Neural Network (MSCNN) is designed to extract multi-scale local and global features through parallel convolutions, enabling precise representation of complex fault patterns. Finally, a channel-based self-attention mechanism (SE module) is incorporated to dynamically adjust feature channel weights, highlighting critical features and significantly improving diagnostic capability in noisy environments. The experimental findings demonstrate that DWDF-MFE attains superior diagnostic accuracy and robustness in comparison to existing methodologies across a range of bearing datasets, particularly in scenarios where signal-to-noise ratios as low as -10 dB prevail. This study provides an innovative and reliable solution for fault diagnosis of rotating machinery under complex working conditions.

## 1. Introduction

Ensuring reliable fault detection in rotating equipment is essential for maintaining both operational efficiency and safety in industrial systems (Wang et al., 2025). Nevertheless, the complexity and diversity of the working and operating conditions of such machinery invariably render fault diagnosis a challenging task (Guo et al., 2024; Lei et al., 2013). In practical industrial scenarios, the operation of rotating machinery is frequently accompanied by significant noise, dynamic load variations, and uncertainty arising from multiple sources of information. The increasing complexity of fault scenarios makes it difficult for any single method to ensure reliable diagnosis. In parallel with industrial advancement, addressing fault detection under complex working conditions has become a central concern in the research on rotating machinery (Zhang et al., 2025a).

Traditional fault diagnosis approaches often rely on signal processing techniques combined with classical machine learning algorithms. Techniques like Fast Fourier Transform (FFT) (Yoo, 2019) and Wavelet Transform (WT) (Zhang et al., 2025b) are commonly applied in signal analysis to extract fault characteristics across time, frequency, and time-frequency domains. Even though these features can provide information about the device's operational status, their extraction requires manual selection and design. Through frequency analysis, researchers can extract fault features such as root-mean-square (RMS) values, crest factors, or spectral components. This process relies heavily on manual experience, which introduces subjectivity. Once the signal features are extracted, traditional methods typically use machine learning models to build fault classifiers, such as Random Forest (Ming et al., 2024) and Support Vector Machine (SVM) (Hou & Huang, 2025). Using mapping relationships between extracted features and fault types, these models

---

classify the operational status of equipment. However, their effectiveness is largely influenced by the quality of the input features. When the features are insufficient to accurately characterize faults, the classification performance can significantly degrade. Furthermore, these models lack adaptability in complex and variable operating conditions and often struggle to process signals in high-dimensional, nonlinear, and noisy environments (Kim et al., 2025).

These limitations are closely related to the intrinsic characteristics of vibration signals in rotating machinery. In rotating machinery, the physical mechanisms of bearing faults give rise to characteristic vibration behaviors. Localized defects in the inner race, outer race, or rolling elements generate periodic impacts that excite harmonics and sidebands in the frequency domain. Meanwhile, vibration responses measured in different directions (e.g., horizontal, vertical, and axial) differ due to structural transmission paths and sensor placement, and these responses are further affected by variations in load, rotational speed, and defect severity. As a result, fault-related spectral components exhibit strong condition dependence, revealing the inherent complexity and multi-scale nature of vibration signals in rotating machinery.

As deep learning technologies continue to evolve, researchers have increasingly explored their application in fault diagnosis. Convolutional neural networks (CNN) (Guo et al., 2025; Yu et al., 2023), deep confidence networks (DBN) (Guo et al., 2021), and long-short memory networks (LSTM) (Chen et al., 2021b) are widely used for automatic feature extraction and fault classification, enabling end-to-end model training and direct extraction of informative diagnostic features from raw sensor signals. However, most existing deep learning methods rely on single-sensor data (Jiang et al., 2019), which limits their ability to perceive system-level information and makes them susceptible to noise and sensor faults.

Multi-sensor data fusion has gained increasing attention in fault diagnosis research, aiming to overcome the inherent constraints of single-sensor approaches (Chen et al., 2021a; Duan et al., 2018; Xu et al., 2023). Fusing data from multiple sensors enables more complete system-level feature representation, thereby enhancing the precision and reliability of fault diagnosis. Multi-sensor fusion methods are commonly classified into data-, feature-, and decision-level approaches based on the stage of integration (Castanedo, 2013; Shao et al., 2021; Wang et al., 2019). Data-level fusion reduces information loss by utilizing multiple sensor signals fully. Wang et al. (2022) proposed a fault diagnosis approach that integrates multi-sensor data weighted fusion (MS-DWF) with depthwise separable convolution (DWSC), facilitating richer feature extraction from raw signals while preserving fine-grained details. Xia et al. (2018) employed a convolutional neural network (CNN) to integrate multiple vibration signals, achieving effective diagnosis of faults in rolling bearings and gearboxes. The utilization of data-layer fusion enhances the fault identification process in rotating equipment by enabling precise feature extraction, reducing information loss, and increasing the dependability of diagnostic outcomes. However, data-level fusion also presents certain challenges, particularly the redundancy inherent in raw sensor data. This redundancy may degrade fusion performance and thereby affect the accuracy of identifying fault conditions (Tong et al., 2022).

From the perspective of fusion strategies, data-level methods can be further grouped into several representative categories. Simple aggregation methods (e.g., signal concatenation, fixed or heuristic weighted averaging) are easy to implement but generally assume consistent sensor reliability and are unable to explicitly handle uncertainty or conflict. Rule-based fusion approaches, such as fuzzy-logic-based schemes, rely on manually designed membership functions and rule bases, which introduce subjectivity and limit adaptability under complex and varying industrial conditions. Probabilistic fusion methods, such as Bayesian inference, provide a rigorous statistical framework, but their performance strongly depends on accurate prior distributions and likelihood models that are often difficult to obtain in practical machinery applications.

In contrast, Dempster-Shafer (DS) evidence theory explicitly models belief, plausibility, and conflict, making it particularly effective for handling uncertainty and contradictory information from sensors without relying on strong prior assumptions. This flexibility makes DS evidence theory an attractive choice for multi-sensor fusion in fault diagnosis, particularly in noisy and uncertain environments typically encountered in industrial settings. The ability to integrate and manage uncertainty across multiple sensors is a crucial advantage, enabling more reliable decision-making compared to other methods that may assume sensor data is either fully reliable or entirely unreliable.

In feature-level fusion, features are first derived from individual sensor outputs and then combined for further analysis. Wan et al. (2023) introduced a multi-sensor information coupling network (MICN) designed for bearing fault diagnosis, which enables progressive and synchronous fusion of sensor features across layers to capture more informative representations. Duan et al. (2023) focuses on multi-source information fusion through an attention mechanism oriented along spatial axes, which facilitates the extraction of fine-grained channel and positional features across different scales. Additionally, the study incorporates a multi-scale dense cyclic (MSDC) structure to improve feature learning, especially when the available training data are limited. Yang et al. (2022) presented a multi-channel graph convolutional network (GCN) framework tailored for rotating machinery fault diagnosis. In this approach, individual sensor signals are converted into graph-structured data, and a parallel graph processing strategy is employed to achieve effective fusion of multi-source features.Feature-level fusion preserves more information than data-level fusion, rendering it especially effective in fault diagnosis, where precise features are essential (Xie et al., 2021). However, it faces challenges such as high-dimensional feature vectors that demand more computational resources. Moreover, since different features contribute unequally to the final decision, it is essential to select informative features and allocate suitable weights to improve fusion performance and reduce information loss.

At the decision level, fusion is achieved by aggregating outputs from multiple sensor-specific classifiers. Typical strategies include voting-based methods and evidence combination techniques based on Dempster-Shafer (DS) theory. Shao et al. (2021) proposed a multi-sensor fault diagnosis method that combines stacked autoencoders with Morlet wavelet functions for feature learning, and employs a flexible weighting scheme for decision-level fusion. Zeng et al. (2019) designed a decision-level fusion framework that leverages integrated weighting approaches and optimized voting schemes, thereby enhancing both the precision and stability of fault diagnosis in rotating machinery. Although decision-level fusion is robust to external interference, its performance is significantly affected by the specific strategy used.

Existing fault diagnosis methods still face challenges, including the limited ability of single-sensor data to capture complex fault patterns, and a lack of effective solutions for managing uncertainty and conflict in multi-sensor fusion. Although multi-sensor fusion has the potential to enhance system-level fault diagnosis, many existing approaches still struggle to fully exploit complementary information from multiple sensors, especially under noisy or varying operating conditions. These limitations highlight the need for a more robust fusion mechanism that can better leverage cross-sensor complementarity and maintain adaptability in complex environments. To address these challenges, this work proposes the DWDF-MFE approach, which integrates dynamic weighted data fusion with multi-scale feature extraction. The proposed method aims to improve diagnostic accuracy and robustness by effectively utilizing complementary information across sensor channels. The primary contributions of this study are as follows:

1. A novel fault diagnosis method named DWDF-MFE is proposed, which employs multi-channel signal fusion by integrating data-level and feature-level information to enhance performance under complex and noisy operating conditions.

2. At the data level, frequency-domain features are extracted from multi-channel signals and fused via a dynamic weighting mechanism based on Dempster-Shafer (DS) evidence theory. The fusion process adaptively assigns weights to each channel by evaluating its relevance, consistency, and conflict degree, thereby improving the robustness and reliability of the aggregated input.

3. For feature-level fusion, a parallel multi-scale convolutional neural network (MSCNN) is constructed to capture features across varying receptive fields. To enhance channel-wise feature representation, a squeeze-and-excitation (SE) attention module is incorporated, enabling adaptive weighting that emphasizes informative signals while suppressing noise.

The organization of the remainder of this paper is as follows. Section 2 introduces the theoretical foundations relevant to the study, including convolutional neural networks, Dempster-Shafer evidence theory, and channel-wise attention. Section 3 presents the DWDF-MFE framework, including data-level fusion, feature-level fusion, and the integration of the attention module. Section 4 outlines the experiment, incorporating data selection, method comparison, and result analysis to assess the model of validity. Section 5 concludes the study and discusses potential directions for future research.

## 2. Related works

Within the domain of rotating machinery fault diagnosis, Convolutional Neural Networks (CNNs), Dempster-Shafer (DS) evidence theory, and the channel attention mechanism have received considerable research interest. This section briefly introduces these techniques and their basic mathematical formulations. Since this paper focuses on extracting deep features from time-series signal vibrations using CNNs, the following section will focus on one-dimensional convolutional neural networks.

### 2.1. Convolutional neural networks (CNNs)

CNNs are deep learning model that perform hierarchical feature extraction through multi-stage convolution and pooling applied to raw input data. The convolution process, which constitutes the fundamental component of CNNs, can be mathematically formulated as follows:

$$Y = X * W + b \tag{1}$$

where $Y$ represents the output feature map, $X$ denotes the input feature map, $W$ is the convolutional kernel, and $b$ is the bias term. The pooling operation reduces the spatial dimension of feature map.

$$Y_{i,j} = \max_{(m,n) \in \mathcal{P}} X_{i+m, j+n} \tag{2}$$

where $X$ is the input feature map, $Y$ is the pooled output, and $\mathcal{P}$ denotes the local pooling window centered at location $(i, j)$. This operation retains the most prominent activation within each region, thereby preserving salient features while reducing computational complexity.

### 2.2. Dempster-Shafer evidence theory

Dempster-Shafer Evidence Theory is widely used for uncertainty modeling and multi-sensor data fusion. It provides a method to combine evidence from multiple sources using belief functions. The Basic Probability Assignment (BPA), quantifying the belief associated with an event $A$, is formulated as follows:

$$m(A) = \frac{\sum_{j=1}^{N} \alpha_j \phi_j(A)}{\sum_{j=1}^{N} \alpha_j} \tag{3}$$

where $\alpha_j$ denotes the confidence weight assigned to the $j$-th source of evidence, and $\phi_j(A)$ represents the degree of support that this source provides for event $A$. This weighted fusion formulation is particularly advantageous in fault diagnosis tasks involving uncertain or conflicting sensor data.

### 2.3. Channel attention mechanism (squeeze-and-excitation)

While the Squeeze-and-Excitation (SE) mechanism is commonly applied in image processing to model inter-channel dependencies in spatial feature maps, this paper extends its use to one-dimensional (1D) vibration signals for fault diagnosis. Unlike conventional applications that operate on multi-dimensional image features, the proposed adaptation applies SE to frequency-domain representations of 1D signals, enabling the model to perform channel-wise reweighting and enhance fault-relevant frequency components.

The squeeze operation summarizes temporal information from the 1D vibration signal by applying global average pooling across the signal length $L$ for each channel. For a given channel $c$, the channel-wise descriptor $z_c$ is computed as:

$$z_c = \frac{1}{L} \sum_{i=1}^{L} x_{c,i} \tag{4}$$

where $x_{c,i}$ denotes the signal value at the $i$-th time step in channel $c$, and $L$ is the total length of the input signal.

The excitation operation produces attention coefficients for each channel based on the compact representations obtained in the squeeze phase. Two fully connected layers are used to calculate these attention weights, and the weights $s$ for each channel are computed as:

$$a = \sigma(\mathbf{V}_2 \cdot \delta(\mathbf{V}_1 \cdot u)) \tag{5}$$

where $\mathbf{V}_1$ and $\mathbf{V}_2$ are trainable weight matrices, $\delta$ denotes the ReLU activation function, and $\sigma$ refers to the Sigmoid function. The resulting vector $a$ contains the attention weights assigned to each channel based on the compressed feature vector $u$.

Finally, the attention weights are applied to the original feature map, recalibrating the features by performing element-wise multiplication:

$$\hat{x}_c = x_c \cdot s_c \tag{6}$$

where $\hat{x}_c$ is the re-weighted feature map for channel $c$, and $s_c$ is the attention weight for that channel. This re-weighting mechanism enhances the sensitivity of the network to fault-related features while diminishing the influence of less informative signals.

## 3. Proposed method

To confront the challenges of fault diagnosis in rotating machinery under complex operational scenarios, this study introduces a multi-channel data fusion approach, termed DWDF-MFE. The method combines data-level fusion using Dempster-Shafer (DS) Evidence Theory and feature-level fusion employing a combination of Multi-Scale Convolutional Neural Networks (MSCNN) and a Squeeze-and-Excitation (SE) module. The following presents the core components and processes of the proposed technique.

### 3.1. Data-level fusion based on DS theory with dynamic weighted adjustment

Signals obtained from multiple sensors typically contain fault-related information from different sources. However, discrepancies in sensor placement, signal interference, redundancy, and noise can cause conflicts between the signals, ultimately reducing diagnostic accuracy. While traditional data fusion methods based on single feature metrics
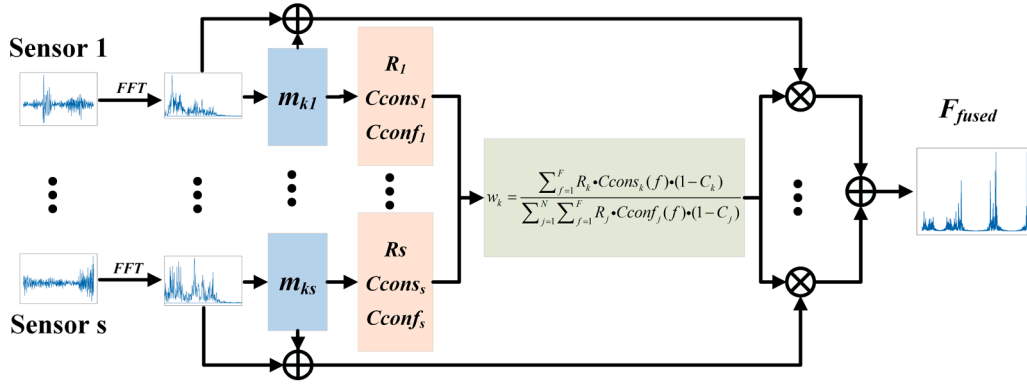
**Fig. 1.** Data fusion with dynamic weighting using Dempster-Shafer evidence theory.

have been successfully applied, they often struggle to capture intricate signal features, especially in noisy and nonlinear environments.To address this challenge, frequency-domain features are chosen in this study. These features are particularly well-suited for capturing fault-related characteristics, such as harmonics and sidebands, which tend to remain stable even under varying operating conditions and noise interference. Compared to time-domain and time-frequency-domain features, frequency-domain representations offer more consistency and reliability, making them ideal for multi-sensor fusion in complex industrial settings. Building on the advantages of frequency-domain features, this study proposes a dynamic weighted data fusion approach utilizing Dempster-Shafer (DS) evidence theory, as illustrated in Fig. 1. This method integrates multi-channel signal characteristics, including energy, consistency, and conflict information, to enhance fault diagnosis performance. The detailed fusion process is outlined as follows:

### 3.1.1. Generating basic probability assignment (BPA)

The BPA is generated based on the frequency energy of each channel. For a given channel $k$, the BPA $m_k(f)$ is defined as:

$$m_k(f) = \frac{E_k(f)}{\sum_{f=1}^{F} E_k(f)} \tag{7}$$

where $E_k(f)$ represents the energy of channel $k$ at frequency $f$, and $F$ is the total number of frequency components.

Higher frequency energy typically indicates more prominent fault characteristics. For instance, in the case of an outer ring fault (OF), characteristic frequencies corresponding to the fault typically show a clear peak in the frequency spectrum. These peaks are more prominent in the frequency domain compared to the surrounding frequencies, providing a clearer indication of fault presence.

The BPA satisfies:

$$\sum_{f=1}^{F} m_k(f) = 1 \tag{8}$$

where $m_k(f)$ represents the assignment weight over $F$ components for channel $k$, ensuring a valid probabilistic distribution.

### 3.1.2. Dynamic weighted fusion

Dynamic weighting adjusts the contribution of each channel using three factors: frequency-domain significance, consistency, and conflict. These factors are calculated as follows:

*Frequency-domain significance.* The frequency-domain significance $R_k$ for channel $k$ is defined as:

$$R_k(f) = \left(m_k(f)\right)^{\alpha} \tag{9}$$

where $m_k(f)$ is the BPA for channel $k$ at frequency $f$, and $\alpha$ is the significance adjustment factor controlling the weight emphasis. Since $m_k(f)$

is derived from the normalized energy distribution in the frequency domain, a larger $m_k(f)$ indicates that channel $k$ exhibits stronger spectral components at frequency $f$, which are typically associated with fault-related features. Physically, $R_k(f)$ therefore reflects how sensitive a given channel is to fault characteristics at that frequency.

*Consistency.* The consistency Consistency$_k$ evaluates the agreement among channels for each frequency:

$$\text{Consistency}_k(f) = \exp\left(-\frac{\text{Var}(m_k(f))}{\mu(m_k(f)) + \epsilon}\right)^{\beta} \tag{10}$$

where $\text{Var}(m_k(f))$ and $\mu(m_k(f))$ denote the variance and mean of $m_k(f)$ across all channels, and $\epsilon$ is a small positive constant introduced to avoid division by zero. The parameter $\beta$ adjusts the influence of consistency. When the BPAs of different channels at the same frequency are similar, the variance becomes small and Consistency$_k(f)$ approaches one, indicating a high level of cross-channel agreement. From a physical perspective, this term measures the stability and reliability of channel $k$ with respect to other sensors at frequency $f$.

*Conflict.* The conflict $C_k$ quantifies the disagreement between channel $k$ and the other channels:

$$C_k = \left(1 - \sum_{B \cap C = \emptyset} m_k(B) \cdot m_j(C)\right)^{\gamma} \tag{11}$$

where $B$ and $C$ are focal elements from channels $k$ and $j$, $\cap$ denotes the intersection of focal sets, and $\emptyset$ represents the empty set. The parameter $\gamma$ adjusts the sensitivity to conflicts. In DS theory, a larger conflict term indicates that the evidence provided by channel $k$ is more contradictory to that of other channels. Physically, this often corresponds to situations where the sensor is affected by severe noise, local disturbance, or abnormal measurements. Thus, $(1 - C_k)$ acts as a penalty factor that suppresses unreliable or conflicting channels.Intuitively, when the overlap between the evidences provided by two channels is small, the frequency components they emphasize differ significantly, leading to a higher conflict degree. This reflects that the corresponding sensor is less consistent with the others and may be affected by noise or abnormal disturbances.

*Dynamic weight calculation.* The dynamic weight $w_k$ for channel $k$ is calculated as:

$$w_k = \frac{\sum_{f=1}^{F} R_k(f) \cdot \text{Consistency}_k(f) \cdot (1 - C_k)}{\sum_{j=1}^{N} \sum_{f=1}^{F} R_j(f) \cdot \text{Consistency}_j(f) \cdot (1 - C_j)} \tag{12}$$

where $N$ is the total number of channels.

The sensor weights are dynamically adjusted based on three key factors: frequency-domain significance, consistency, and conflict. Frequency-domain significance reflects the energy of each channel at fault-related frequencies, consistency measures the agreement between
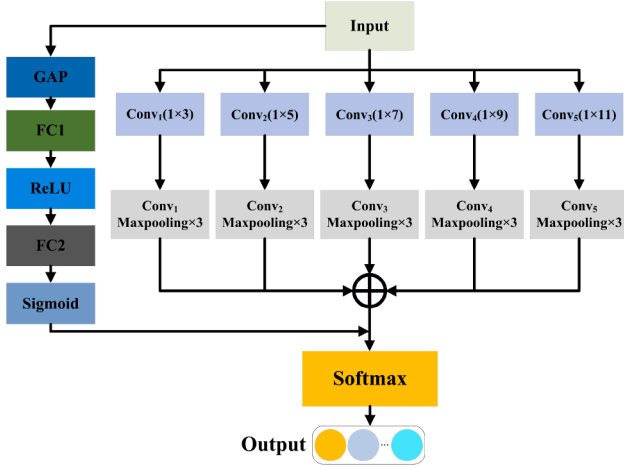
**Fig. 2.** Architecture of a parallel multiscale CNN with SE module.

channels, and conflict quantifies the discrepancies between them. By evaluating these factors, the system assigns higher weights to sensors that provide reliable and consistent data, while reducing the influence of sensors with less reliable or conflicting information. This approach ensures that the fusion process emphasizes reliable diagnostic information while minimizing the impact of uncertain or misleading evidence.

*Fusion of features.* : finally, the BPA and FFT features are fused using the calculated dynamic weights:

$$F_{\text{fused}} = \bigoplus_{k=1}^{N} w_k \cdot \left( F_k + m_k(f) \right) \tag{13}$$

where $\bigoplus$ denotes the concatenation operation, $F_k$ represents the FFT features of channel $k$, and $m_k(f)$ is the BPA for channel $k$.

### 3.1.3. Summary

The proposed method generates BPAs based on frequency domain energy and dynamically adjusts the weights of the multichannel signals by taking into account frequency domain importance, coherence, and conflicts. This fusion mechanism ensures that signals from multiple channels are flexibly combined into a unified signal, which is then used for subsequent training. The parameters $\alpha$, $\beta$, and $\gamma$ are set to 0.8, 1.5, and 1.2, respectively, to balance the frequency significance, consistency, and conflict adjustments.These values were determined through an optimization process involving cross-validation , where $\alpha$ emphasizes the frequency-domain significance, $\beta$ influences the consistency of signals, and $\gamma$ adjusts the sensitivity to conflicts. This combination of parameters was found to provide the best performance in fault detection accuracy and fusion stability.

### 3.2. Feature-level fusion: Multi-scale CNN with SE module

Bearing fault features exhibit a hierarchical structure,where high-level features require extraction using large convolutional kernels, while low-level features are better captured with smaller kernels. In response to this challenge, a feature-level integration approach is constructed by combining MSCNN with the SE module to enhance feature learning. The MSCNN design incorporates multiple convolutional branches with different kernel sizes to capture features at various scales, while the SE module adaptively reweights the channel features through an attention mechanism to enhance fault-relevant information. Fig. 2, shows its structure and the detailed procedure is described as follows:

The MSCNN uses parallel convolution kernels of varying sizes (e.g., 3, 5, 7, 9, 11) to extract features at different scales. In one-dimensional vibration signal analysis, using kernels of different lengths allows the

network to capture information across multiple temporal resolutions. Shorter kernels provide small receptive fields and are therefore effective for extracting localized fine-grained variations in the signal, whereas longer kernels cover a broader temporal neighborhood and help characterize more global waveform structures. By combining these kernels in a parallel multi-scale architecture, the MSCNN obtains a more comprehensive feature representation than would be possible with a single fixed kernel size.

$$f_k(x) = \sigma(w_k * x + b_k) \tag{14}$$

where $f_k(x)$ represents the feature extracted using a convolution kernel of size $k$, $w_k$ and $b_k$ are the kernel weights and bias, and $\sigma$ is the activation function ReLU. This method allows the network to capture fault features at various scales, obtaining a hierarchical understanding of the signal.

To mitigate the risk of overfitting and enhance the generalization capability of the network, two regularization techniques are incorporated into the model architecture. Specifically, a Dropout layer with a rate of 0.5 is introduced in the fully connected stage to prevent co-adaptation of neurons, while Batch Normalization is applied after each convolutional layer to stabilize the distribution of intermediate features and facilitate more robust training.

The feature vectors extracted at multiple scales are concatenated into a unified feature matrix:

$$F_{\text{multi}} = \text{Concat}(f_1(x), f_2(x), \ldots, f_n(x)) \tag{15}$$

where $f_i(x)$ denotes the feature vector obtained from the $i$-th convolutional kernel. This multi-scale concatenation yields a high-dimensional representation that integrates diverse receptive fields, thereby enriching the feature space for subsequent processing.

In the SE module, each channel's weight is dynamically adjusted using the formula:

$$A_c = U_c \cdot \sigma \left( \mathbf{V}_2 \cdot \text{ReLU}\left( \mathbf{V}_1 \cdot \frac{1}{L} \sum_{i=1}^{L} U_{i,c} \right) \right) \tag{16}$$

where $U_c$ denotes the feature vector of the $c$-th channel, $\mathbf{V}_1$ and $\mathbf{V}_2$ are trainable weight matrices, ReLU and $\sigma$ represent the ReLU and Sigmoid activation functions respectively, and $L$ is the feature length. This formulation illustrates how the SE mechanism adaptively scales the features of each channel, emphasizing those relevant to fault patterns while attenuating less informative responses.

In the proposed MSCNN-based feature extraction module, the SE block is applied to the output of each multi-scale convolution branch. After each scale-specific convolutional path generates its 64-channel feature map, the SE module performs channel-wise recalibration by first compressing the feature map through global average pooling to obtain channel descriptors, and then producing attention coefficients via a lightweight excitation network. These coefficients are subsequently used to reweight the feature map through channel-wise multiplication. Through this mechanism, the SE module dynamically modulates the influence of each channel, enabling the network to emphasize features associated with fault patterns and consequently enhance diagnostic performance.

### 3.3. Fault diagnosis procedure of DWDF-MFE

Fig. 3 depicts the proposed fault diagnosis workflow, which is as follows:

**Step 1: Data Acquisition.** Initially, vibration data are acquired from multiple sensors mounted on the rotating machinery. These signals contain fault-related information from different parts of the machinery. However, due to the varying placement of sensors and potential signal
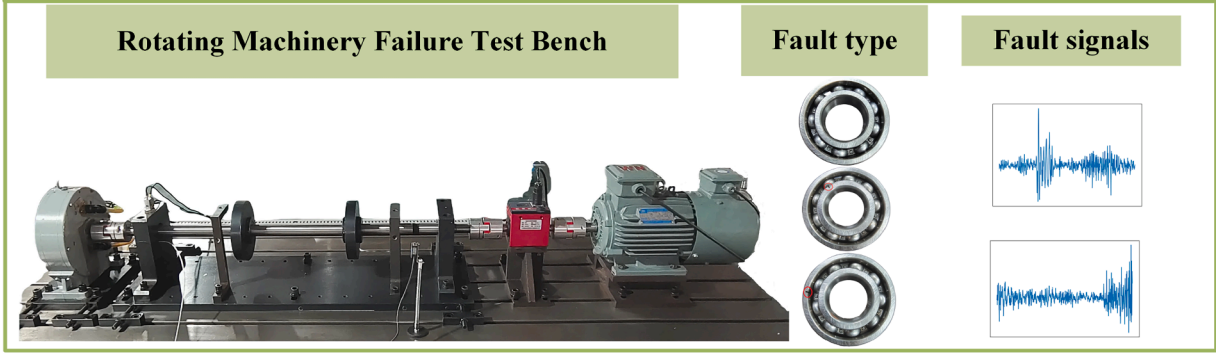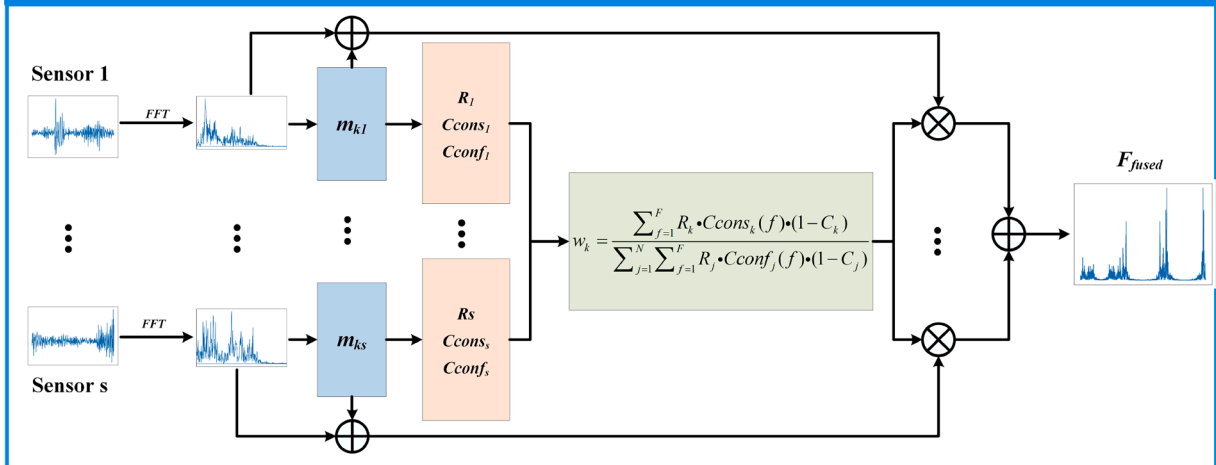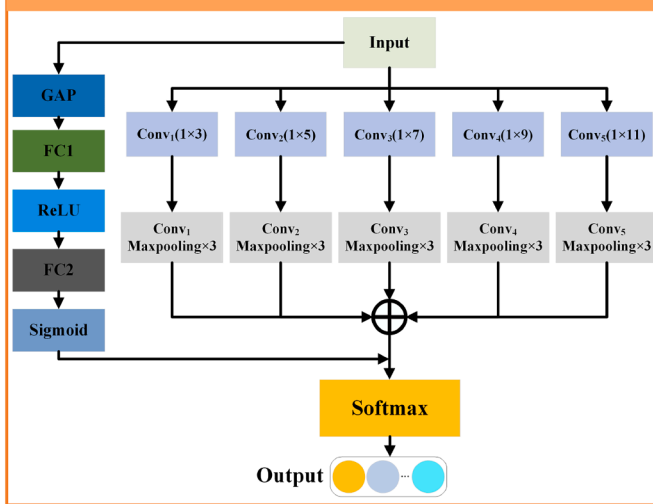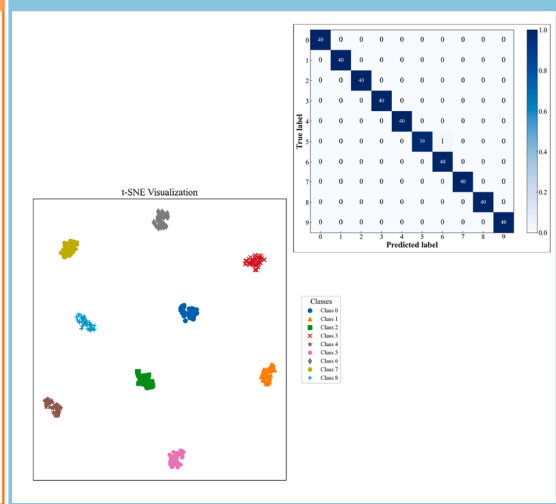
Fig. 3. Schematic of the proposed fault diagnosis workflow.

interference, the raw signals are acquired and processed in the subsequent stages.

**Step 2: Data-Level Fusion.** In the data-level fusion stage, the acquired multi-sensor signals are processed by a dynamic weighting module constructed upon Dempster-Shafer (DS) theory. This module evaluates frequency-domain features and potential conflicts among signals, dynamically sensor-wise weights to produce fused outputs with improved reliability and resistance to uncertainty.

**Step 3: Feature Fusion.** Following the data fusion process, the integrated signals are processed by a MSCNN to extract features. MSCNN employs parallel convolutional kernels of different sizes to capture both local and global features. Subsequently, a SE module performs
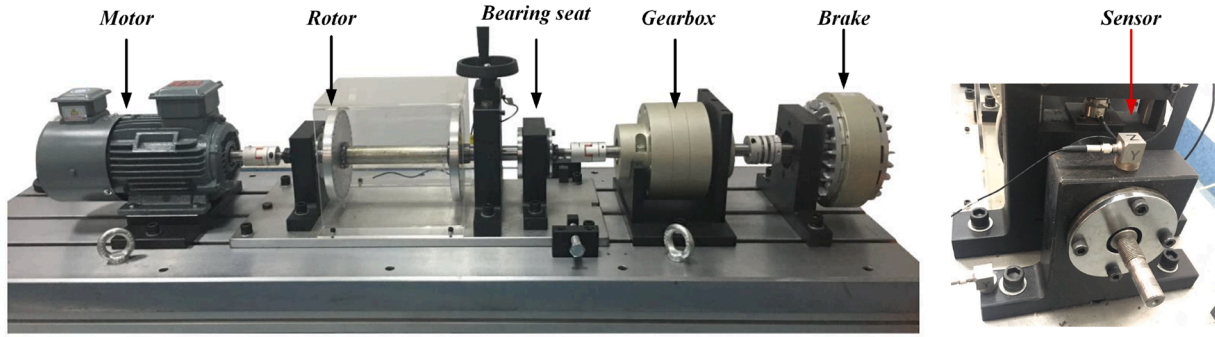
**Fig. 4.** The experimental setup schematic for bearing fault diagnostic in case 1.

channel-wise adaptive weighting through an attention mechanism, enhancing fault-relevant features while suppressing noise.

**Step 4: Diagnostic Results.** The fused feature representations are subsequently passed through fully connected layers and classified via a Softmax function to produce diagnostic outcomes, which provide the foundation for evaluating the performance and reliability of the proposed method.

Although this study focuses on vibration signals collected by triaxial accelerometers, the DWDF-MFE framework is not inherently limited to this type of input. Other sensing modalities used in condition monitoring-such as acoustic emission, airborne sound, or electrical measurements-also contain temporal-spectral patterns associated with mechanical faults and, in principle, could be processed within a similar fusion architecture provided that appropriate feature representations are constructed.

## 4. Experiment

### 4.1. Case 1: SDUST bearing dataset analysis

#### 4.1.1. Experimental dataset selection and preprocessing

The experimental data were obtained from the transmission component fault test platform for rotating machinery at Shandong University of Science and Technology (Han et al., 2023, 2022; Wang et al., 2024; Zhang et al., 2024). The bearing fault simulation test bed is shown in Fig. 4. The experimental platform includes an electric motor, speed controller, test bearing (type 6205 deep groove ball bearing), gearbox, and a magnetic particle brake. Localized faults of different types were introduced on the bearing components using electrical discharge machining (EDM), with three predefined defect severities corresponding to widths of 0.2–0.6 mm. Triaxial accelerometers were mounted on the bearing housing to acquire vibration signals from multiple directions, and sampling was performed at 25.6 kHz over successive 40-second intervals.The dataset covers diverse operating conditions, including six constant speeds (1000-3000 rpm) and four load levels (0, 20, 40, 60 N). Four health states are defined according to the EDM-induced defect location: normal condition (NC), inner ring fault (IF), outer ring fault (OF), and rolling element fault (RF). Additional information regarding the selected samples is presented in Table 1. This dataset is adopted because its triaxial vibration measurements, high sampling rate, and diverse operating and fault conditions provide sufficient multi-channel information for constructing the multi-class diagnostic task in this study.

#### 4.1.2. Ablation experiment

The proposed DWDF-MFE model consists of four core modules: CNN network, data layer fusion, feature layer fusion, and channel attention mechanism. To evaluate the individual impact of each module on the diagnostic capability of the DWDF-MFE framework, ablation studies were performed using the SDUST bearing dataset under a signal-to-noise ratio (SNR) of -10 dB. The Adam optimizer was employed for model training,
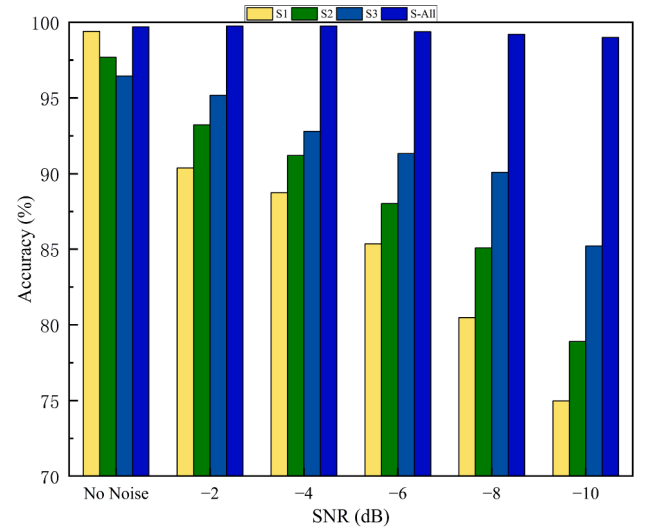


**Fig. 5.** Comparing data source performance under varying noise conditions in Case 1.



**Fig. 6.** Comparison of fault diagnosis results across methods in Case 1.

initialized with a learning rate of 0.001. Cross-entropy served as the loss criterion, and a batch size of 64 was used. The model underwent up to 100 training epochs for parameter updates. The model parameters are detailed in Table 2. To enhance result reliability, each experimental configuration was independently repeated ten times, and the aggregated
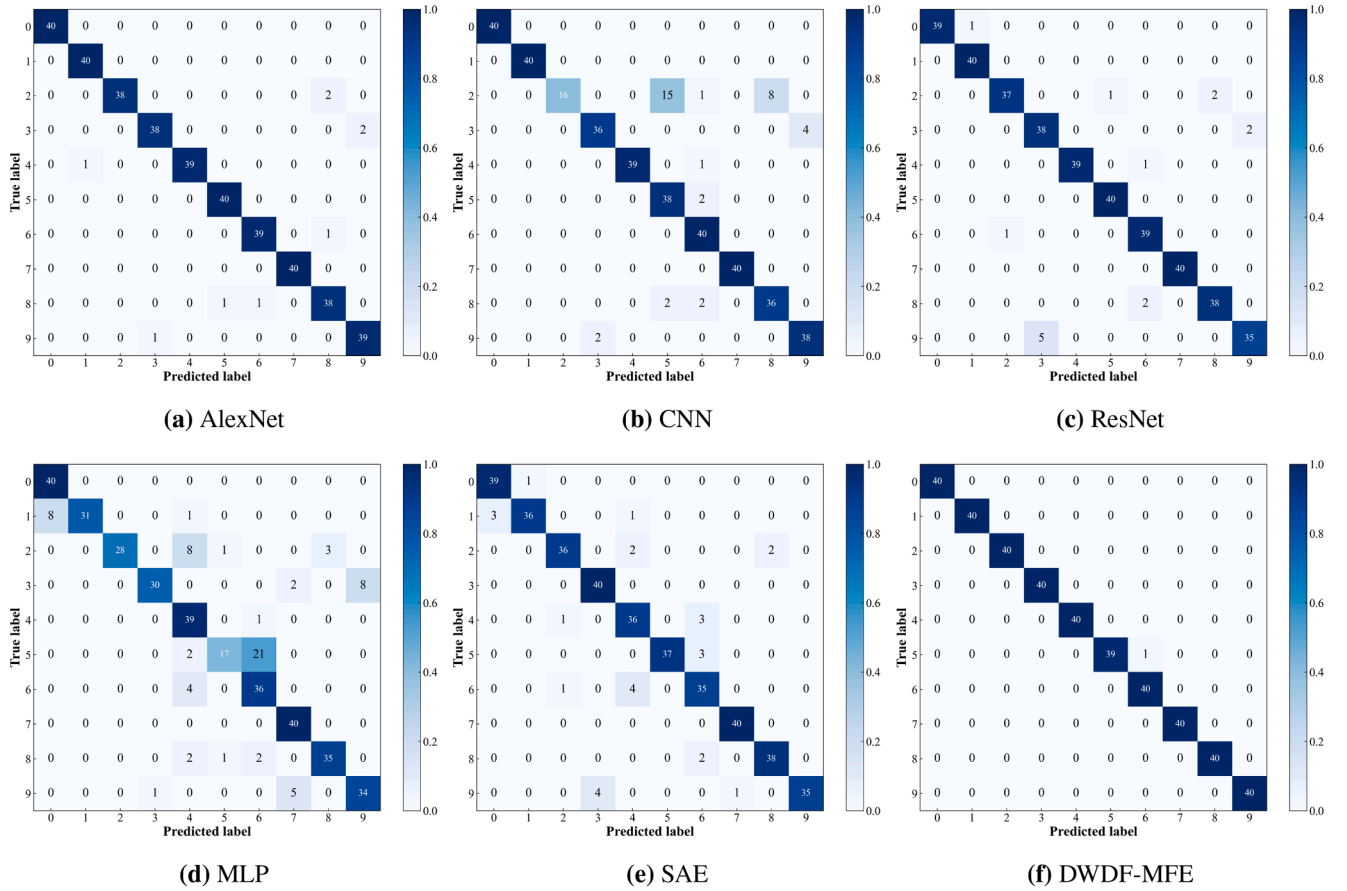
**Fig. 7.** Confusion matrices generated by six methods in Case 1.

**Table 1**
Selected sample information for Case 1.

| Fault type | Fault size | Sample | Label |
|---|---|---|---|
| NC | \ | 200 | 0 |
| IF | 0.2 | 200 | 1 |
| OF | 0.2 | 200 | 2 |
| RF | 0.2 | 200 | 3 |
| IF | 0.4 | 200 | 4 |
| OF | 0.4 | 200 | 5 |
| RF | 0.4 | 200 | 6 |
| IF | 0.6 | 200 | 7 |
| OF | 0.6 | 200 | 8 |
| RF | 0.6 | 200 | 9 |

**Table 2**
The parameter settings of DWDF-MFE.

| Type | Kernel Size | Channel | Activation Function | Has BN | Output Size |
|---|---|---|---|---|---|
| Input | – | – | – | No | 1, 3096 |
| Con_1 | 3*1 5*1 7*1 9*1 11*1 | 8 | ReLU | Yes | 8, 3096 |
| GMP_1 | 2*1 | – | – | No | 8, 1548 |
| Con_2 | 3*1 5*1 7*1 9*1 11*1 | 16 | ReLU | Yes | 16, 1548 |
| GMP_2 | 2*1 | – | – | No | 16, 774 |
| Con_3 | 3*1 5*1 7*1 9*1 11*1 | 32 | ReLU | Yes | 32, 774 |
| GMP_3 | 2*1 | – | – | No | 32, 387 |
| Con_4 | 3*1 5*1 7*1 9*1 11*1 | 64 | ReLU | Yes | 64, 387 |
| GMP_4 | 2*1 | – | – | No | 64, 193 |
| Concat | – | – | – | No | 120, 193 |
| FC | – | – | ReLU | – | 256 |
| Softmax | – | – | – | – | 10 |

outcomes are reported in Table 3. The SNR metric quantifies the relative intensity of the signal compared to background noise in the measured data. The formula for SNR is given by:

$$SNR = 10 \log_{10}\left(\frac{P_{\text{signal}}}{P_{\text{noise}}}\right) \tag{17}$$

where $P_{\text{signal}}$ and $P_{\text{noise}}$ denote the respective energy magnitudes of the signal and noise components.

In industrial environments, vibration measurements are often contaminated by broadband sensor interference originating from acquisition electronics, ambient disturbances, and structural vibrations, which are widely reported to exhibit approximately Gaussian and wideband characteristics. Therefore, Gaussian white noise of varying intensities was introduced into the sensor vibration signals to mimic realistic industrial conditions. The corresponding experimental configurations are detailed as follows:

1. **DWDF-MFE**: The comprehensive diagnostic model that integrates all essential modules, including the CNN network, data-level fusion, feature-level fusion, and channel attention mechanism.
2. **DWDF-MFE-1**: DWDF-MFE variant excluding the channel attention mechanism.
3. **DWDF-MFE-2**: DWDF-MFE variant omitting both channel attention and feature-level fusion components.
4. **DWDF-MFE-3**: DWDF-MFE version with all fusion modules (channel, data, and feature) removed, directly utilizing the concatenated signals from the three sensor channels as input.

Table 3 shows that the DWDF-MFE-1 configuration achieves an average diagnostic accuracy of 98.57%, which is 0.43% lower than the complete model, underscoring the significance of incorporating the channel attention mechanism within the DWDF-MFE framework. When the

**Fig. 8.** Comparative visualization of five diagnostic approaches in case 1.



**Fig. 9.** The experimental setup schematic for bearing fault diagnostic in case 2.
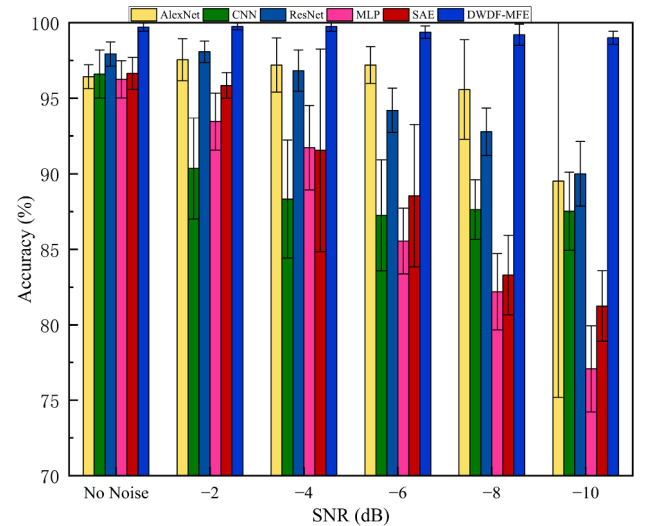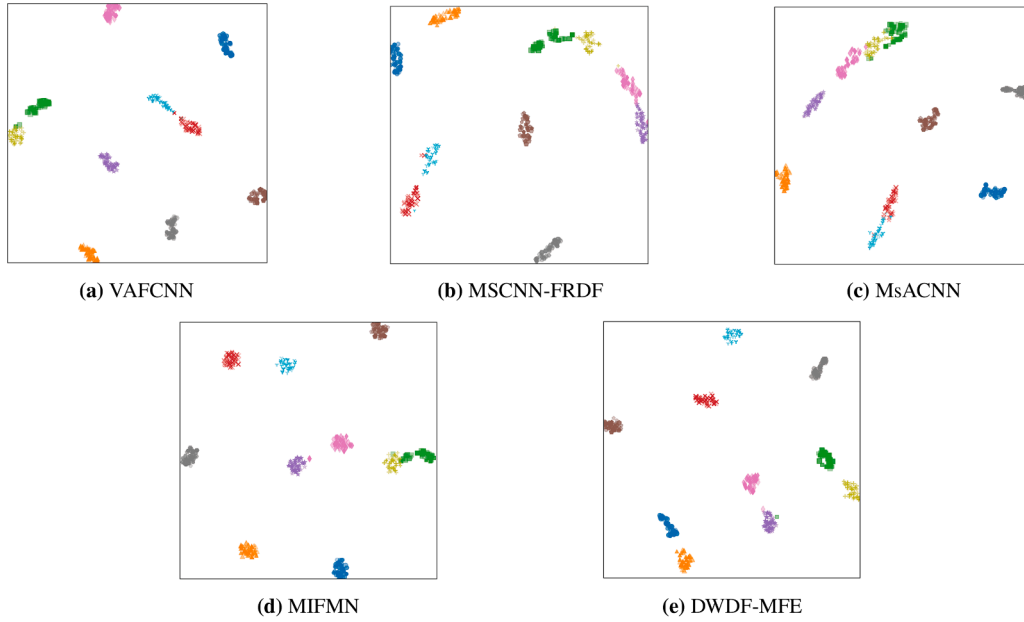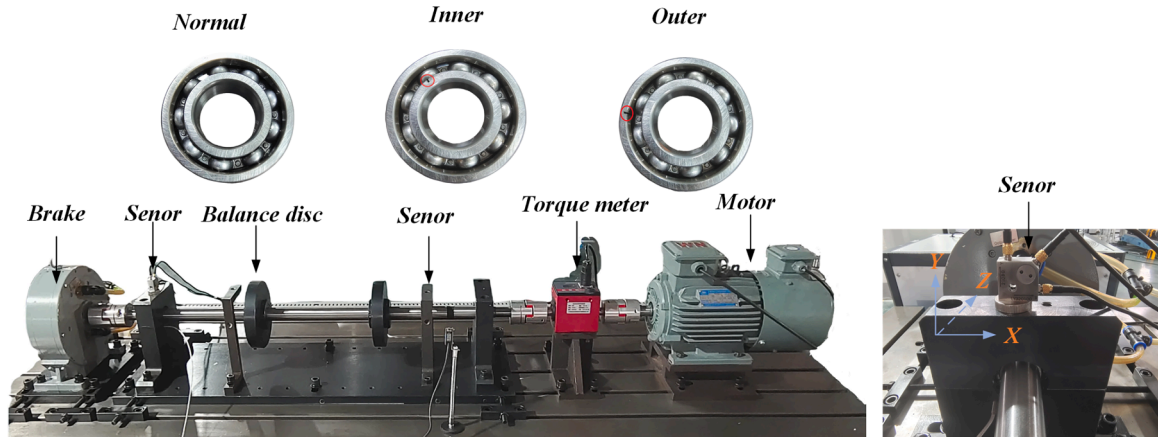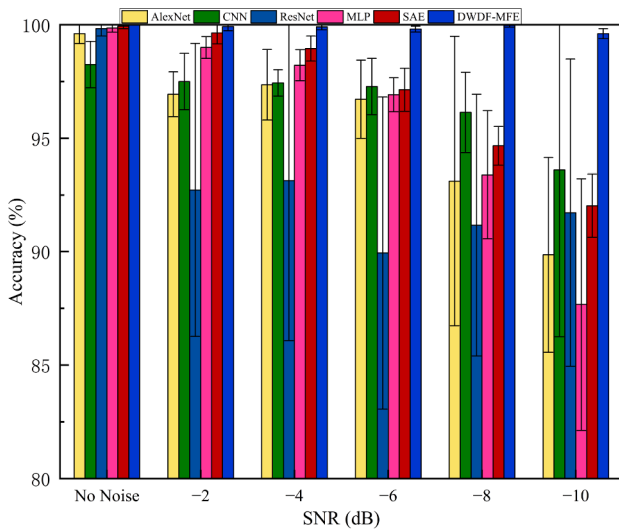


**Fig. 10.** Comparison of fault diagnosis results across methods in Case 2.

**Table 3**
An ablation study of the DWDF-MFE(%).

| Methods | CNN | Data Fusion | Channel Attention | Feature Fusion | Results |
|---|---|---|---|---|---|
| DWDF-MFE | ✓ | ✓ | ✓ | ✓ | **99.00 ± 0.43** |
| DWDF-MFE-1 | ✓ | ✓ | ✓ | | 98.57 ± 0.59 |
| DWDF-MFE-2 | ✓ | ✓ | | | 96.75 ± 1.01 |
| DWDF-MFE-3 | ✓ | | | | 95.78 ± 2.59 |

feature fusion module is removed (DWDF-MFE-2), the accuracy drops to 96.75%, which is 1.82% decrease from DWDF-MFE-1, emphasizing the crucial contribution of the feature fusion component to enhancing diagnostic effectiveness. Further, the accuracy of DWDF-MFE-3 is 95.78%, which is 0.97% decrease from DWDF-MFE-2, indicating that the data fusion method makes a non-negligible contribution to the model performance. A progressive decline in diagnostic performance is observed as essential components are incrementally excluded from the DWDF-MFE architecture.

### 4.1.3. Performance comparison across data sources

To assess the diagnostic performance and advantages of DWDF-MFE, comparative analysis was conducted against single-sensor signal-based

**(a)** AlexNet      **(b)** CNN      **(c)** ResNet

**(d)** MLP      **(e)** SAE      **(f)** DWDF-MFE
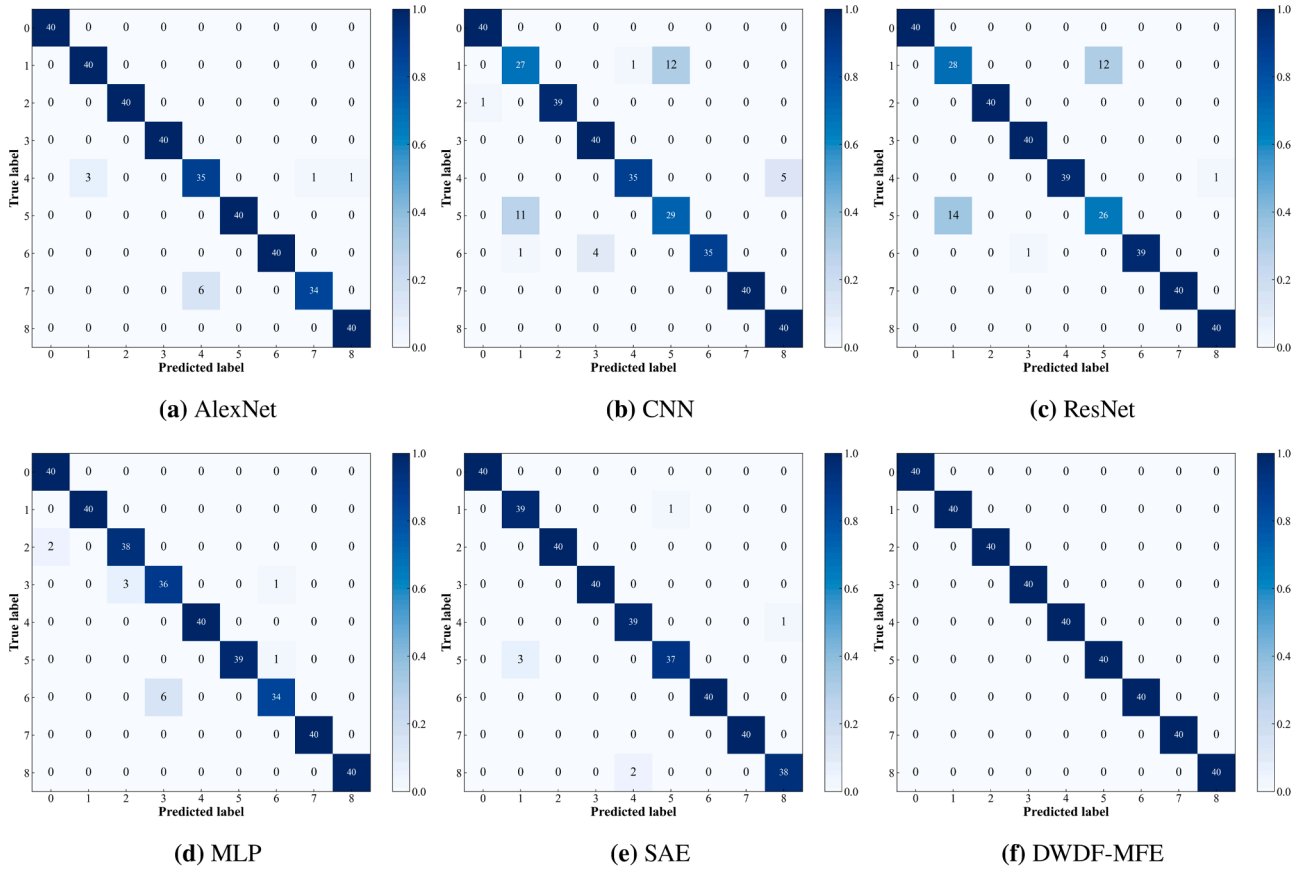
**Fig. 11.** Confusion matrices generated by six methods in Case 2.



**(a)** VAFCNN      **(b)** MSCNN-FRDF      **(c)** MsACNN

**(d)** MIFMN      **(e)** DWDF-MFE
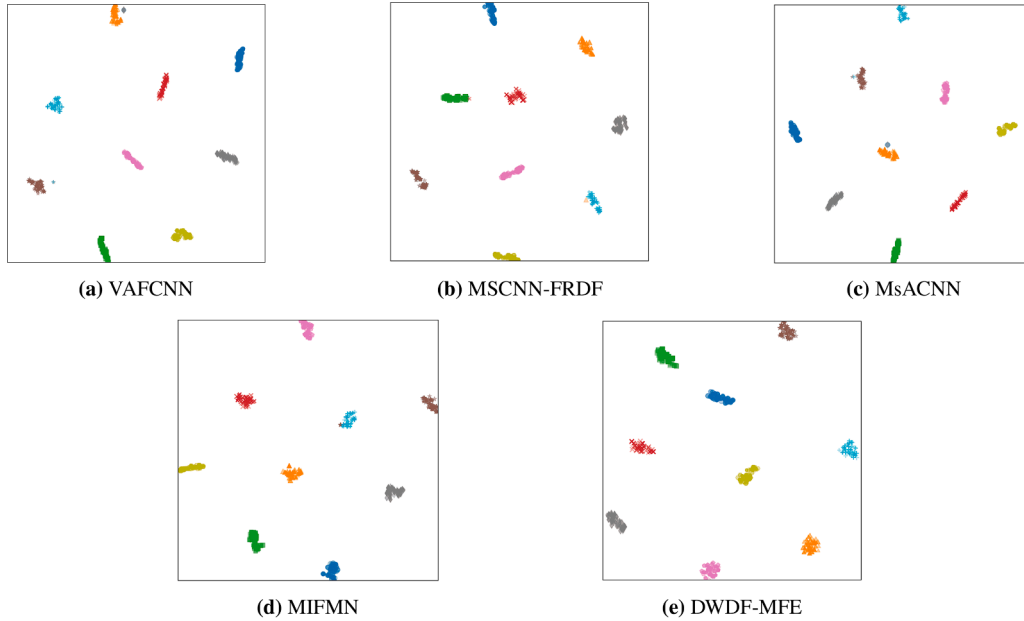
**Fig. 12.** Comparative visualization of five diagnostic approaches in case 2.

diagnosis. In detail, Sensor 1 (S1), Sensor 2 (S2), and Sensor 3 (S3) individually acquire single-channel signals, whereas "S-All" denotes the integrated data produced by fusing outputs from all three sensors using the DWDF-MFE method. In this comparison, the proposed diagnostic model remains consistent, the only difference being the input signal.

The diagnostic outcomes corresponding to various data sources are detailed in Table 4.

The findings indicate that the S-All configuration achieves superior average accuracy with lower standard deviation (std) compared to single-sensor inputs. This performance is attributed to its ability to

**Table 4**
Diagnostic accuracy of ten trials using multiple data sources in Case 1(%).

| Text Number | S1 | S2 | S3 | S-All |
|---|---|---|---|---|
| 1 | 99.50 | 98.25 | 94.00 | 99.50 |
| 2 | 99.25 | 97.50 | 97.50 | 99.75 |
| 3 | 99.75 | 98.00 | 97.00 | 100.00 |
| 4 | 99.75 | 98.50 | 97.00 | 99.75 |
| 5 | 99.50 | 97.50 | 95.75 | 100.00 |
| 6 | 100.00 | 95.25 | 97.00 | 99.25 |
| 7 | 98.25 | 97.00 | 96.75 | 100.00 |
| 8 | 99.25 | 98.00 | 95.75 | 99.75 |
| 9 | 99.25 | 98.50 | 96.00 | 99.75 |
| 10 | 99.50 | 98.25 | 97.50 | 99.25 |
| Mean | 99.40 | 97.675 | 96.45 | **99.70** |
| Std | 0.4500 | 0.9290 | 1.0112 | **0.2693** |

**Table 5**
Comparison of three evaluation indicators across six methods in Case 1(%).

| Method | Accuracy | Precision | F1-score |
|---|---|---|---|
| MLP | 85.55 ± 2.18 | 86.85 ± 1.70 | 85.42 ± 2.27 |
| SAE | 88.55 ± 4.70 | 89.55 ± 3.51 | 88.23 ± 5.28 |
| ResNet | 94.20 ± 1.47 | 94.59 ± 1.26 | 94.20 ± 1.47 |
| AlexNet | 97.20 ± 1.21 | 97.35 ± 1.02 | 97.19 ± 1.22 |
| CNN | 87.25 ± 3.67 | 89.86 ± 3.98 | 86.08 ± 4.49 |
| DWDF-MFE | **99.38 ± 0.41** | **99.40 ± 0.38** | **99.37 ± 0.41** |

integrate complementary information from multiple sensors, enhancing both precision and consistency in fault identification. The three sensors, which record vibration responses at different positions and directions, each provide partial and noise-affected information. By fusing these data, the S-All configuration retains useful fault-related responses while reducing random noise and unstable components from individual channels. As a result, the fused S-All signal contains more complete fault-related information and exhibits stronger stability, leading to higher accuracy and smaller fluctuations in repeated experiments. This multi-sensor data fusion approach mitigates the limitations of single-sensor inputs, which are often susceptible to noise and lack comprehensive diagnostic details, ultimately improving diagnostic stability and accuracy.

To assess the robustness of DWDF-MFE under noisy conditions, varying intensities of noise were introduced into the signals collected from all three sensors. The diagnostic results under these interference levels were then compared. Fig. 5 visualizes the comparative performance of single-sensor methods and multi-sensor fusion across multiple SNR settings. The results demonstrate that the multisensor fusion method S-All outperforms the singl -sensor data (S1, S2, and S3) in all noise conditions, thus demonstrating the significant advantage of integrating multi-sensor information. Notably, when the SNR declines to -10dB, the S-All method exhibits a high degree of diagnostic accuracy, underscoring its notable resilience to noise. This enhanced performance can be attributed to the data complementary characteristics inherent in multi-sensor fusion, which engender more stable and reliable diagnostic results. Compared to the fusion approach, the single-sensor method shows noticeable limitations, particularly in its susceptibility to noise, which significantly reduces its diagnostic precision.

*4.1.4. Comparison with conventional fault diagnosis methods*

Most existing diagnostic approaches make use of vibration signals acquired from a single sensor to identify fault conditions in machinery. To highlight the strengths of the DWDF-MFE method, it is benchmarked against several widely adopted fault diagnosis techniques, including multilayer perceptron (MLP) (Rana et al., 2018), stacked autoencoder (SAE) (Qi et al., 2017), residual network (ResNet) (He et al., 2016), deep convolutional neural network (AlexNet) (Krizhevsky et al., 2012) and basic convolutional neural network (CNN) (Liu et al., 2019).

In the specific experimental setup, the MLP architecture is structured with three hidden layers containing 512, 256, and 128 neurons, respectively., and the input batch size is set to 64, with ReLU activation function for all layers. The SAE model is designed as an input layer containing 1024 neurons, followed by three hidden layers with 768, 512, and 128 neurons, respectively, and the base CNN The underlying CNN architecture consists of three convolutional layers (with kernel sizes of 15, 10, and 5, respectively), two maximal pooling layers, and a globally averaged pooling layer. The ResNet model employs a modified deep architecture that utilizes hopping connections to maintain the gradient flow during training. The model configuration is based on the standard

ResNet guidance scheme, adapted to the troubleshooting dataset. The AlexNet model consists of five convolutional layers, with varying kernel sizes, and a series of maximally-pooled layers. All models are optimized using the Adam algorithm (Jais et al., 2019) with a learning rate of 0.001. The number of training epochs was limited to 100, and the loss function adopted was cross-entropy. The signal sources of all models are the original vibration data collected from the X, Y, and Z directions. The comparison method utilises multi-channel spliced signals as model inputs following a unified preprocessing procedure. Noise corresponding to different SNRs was injected into the data to emulate the interference commonly encountered in practical industrial settings. Fig. 6 displays the experimental results.

Under noise-free conditions, most models achieve similar diagnostic accuracy, with ResNet, AlexNet, CNN, SAE, and MLP maintaining performance around 96%. As the noise level increases-particularly at SNR = -8 dB and -10 dB-the accuracy of most models declines notably.

For instance, MLP drops from 96.25% to 77.08% at -10 dB, and SAE decreases from 96.65% to 81.25%. CNN shows slightly better performance, decreasing from 96.60% to 87.53%. Although AlexNet and ResNet maintain relatively stable performance under low noise levels, their accuracy still declines to 89.53% and 90.00% at -10 dB, respectively.

In comparison, the proposed DWDF-MFE method maintains higher accuracy across all noise levels. At SNR = -10 dB, it achieves 99.00% accuracy. The findings demonstrate that the proposed approach maintains superior robustness and stability when subjected to noise interference.

To further evaluate the DWDF-MFE method in diagnosing various fault types under noisy conditions, experiments were performed at an SNR of -6 dB, and the average diagnostic performance of each model was recorded. Fig. 7 presents the confusion matrix, enabling a visual comparison of how each model performs across the ten fault categories. In the experiments, the DWDF-MFE method shows excellent fault diagnosis ability, with minimal misclassification observed in certain fault categories. Compared to DWDF-MFE, alternative methods yielded a higher number of misclassifications, especially MLP and SAE, which perform poorly in certain categories. In addition, ResNet and AlexNet also exhibit high confusion rates in certain types of fault classification, which indicates their lack of ability to distinguish features.

For a more thorough assessment of model performance, three indicators-accuracy, precision, and F1 score-were employed. Eqs. (18)–(20) represent the accuracy, precision and F1 score, respectively. Table 5 lists the specific evaluation values of all models under the condition of SNR of -6dB. Relative to the baseline methods, the DWDF-MFE model demonstrates consistently superior performance across all evaluation metrics. This advantage arises from its capacity to integrate complementary information from multiple sensors and extract rich frequency-domain features, thereby enhancing both the robustness and precision of fault detection.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{18}$$

$$Macro\text{-}Precision = \frac{1}{C} \sum_{i=1}^{C} \frac{TP_i}{TP_i + FP_i} \tag{19}$$

**Table 6**
Assessing diagnostic performance of multi-sensor fusion methods in Case 1(%).

| Model Type | No Noise | −2dB | −4dB | −6dB | −8dB | −10dB |
|---|---|---|---|---|---|---|
| VAFCNN | 96.43 ± 1.08 | 97.70 ± 1.61 | 97.60 ± 1.31 | 94.83 ± 3.06 | 90.73 ± 2.71 | 89.43 ± 1.19 |
| MSCNN-FRDF | 96.10 ± 1.22 | 92.53 ± 7.17 | 92.68 ± 5.52 | 88.80 ± 8.89 | 84.60 ± 4.74 | 83.15 ± 9.38 |
| MsACNN | 96.02 ± 1.04 | 92.25 ± 8.74 | 93.48 ± 4.24 | 95.65 ± 2.05 | 93.53 ± 3.15 | 93.68 ± 2.06 |
| MIFMN | 97.53 ± 0.87 | 98.68 ± 1.26 | 98.85 ± 1.04 | 97.80 ± 0.86 | 96.93 ± 0.86 | 96.73 ± 1.02 |
| DWDF-MFE | **99.70 ± 0.27** | **99.75 ± 0.22** | **99.75 ± 0.32** | **99.38 ± 0.41** | **99.20 ± 0.70** | **99.00 ± 0.43** |

**Table 7**
Selected sample information for Case 2.

| Fault type | Speed(rpm) | Sample | Label |
|---|---|---|---|
| NC | 600 | 200 | 0 |
| IF | 600 | 200 | 1 |
| OF | 600 | 200 | 2 |
| NC | 900 | 200 | 3 |
| IF | 900 | 200 | 4 |
| OF | 900 | 200 | 5 |
| NC | 1200 | 200 | 6 |
| IF | 1200 | 200 | 7 |
| OF | 1200 | 200 | 8 |

**Table 8**
Comparing data source performance under varying noise conditions in Case 2(%).

| Sensor | No Noise | −2dB | −4dB | −6dB | −8dB | −10dB |
|---|---|---|---|---|---|---|
| S1 | 98.47 | 97.25 | 97.00 | 94.42 | 93.14 | 90.17 |
| S2 | 99.33 | 98.81 | 97.47 | 96.42 | 92.22 | 88.83 |
| S-All | **100.00** | **99.92** | **99.91** | **99.80** | **99.72** | **99.61** |

$$\text{Macro-F1} = \frac{1}{C} \sum_{i=1}^{C} \frac{2 \cdot \text{Precision}_i \cdot \text{Recall}_i}{\text{Precision}_i + \text{Recall}_i} \tag{20}$$

### 4.1.5. Comparison with multi-source fusion methods

To evaluate the effectiveness of the proposed DWDF-MFE approach in fault diagnosis, it is benchmarked against four multi-sensor fusion methods (including VAFCNN Wang et al., 2021, MSCNN-FRDF Tong et al., 2023a, MsACNN Tong et al., 2023b and MIFMN Song et al., 2024) in detail. During the comparison, all models share the same input format, where raw vibration data collected along the X, Y, and Z axes is preprocessed and merged into multi-channel input form.For fairness, all comparative models were trained under identical hyperparameter settings, including learning rate, batch size, number of epochs, and optimization strategy. To simulate real industrial conditions, the experiments are conducted with varying levels of SNR interference. Table 6 presents the diagnostic performance and corresponding standard deviations of each method across various SNR levels.

Under noise-free conditions, all evaluated methods achieve diagnostic accuracies exceeding 95%, as reflected in the experimental results, which exhibits a strong fault recognition capability. With increasing noise levels, noticeable differences emerge in the performance of the methods. the accuracy of MSCNN-FRDF decreases most significantly to about 84.6% at SNR = -8 dB. MsACNN and VAFCNN are slightly better at resisting noise, with accuracies of 93.52% and 90.73% at SNR = −8 dB, respectively. Although these methods show some immunity to interference, they still cannot ensure stable diagnosis under high noise conditions.

In contrast, MIFMN and DWDF-MFE show significant robustness to noise. At SNR = -8 dB, MIFMN achieves an accuracy of 96.93%, while DWDF-MFE achieves an accuracy of 99.20%. It is noteworthy that under more severe noise conditions (SNR = -10 dB), DWDF-MFE still achieves more than 99% accuracy, clearly outperforming the other compared methods. These results indicate that the multi-sensor fusion mechanism employed by DWDF-MFE effectively integrates fault-relevant information and suppresses noise interference, thereby achieving more stable and accurate diagnostic performance. In conclusion, compared with other multisensor fusion methods, DWDF-MFE exhibits excellent robustness and diagnostic performance under complex noise conditions. Especially in high noise environments,DWDF-MFE shows significant advantages.

To evaluate the ability of each model to distinguish between different fault categories, t-SNE is employed in this study to visualize the extracted feature representations. The t-SNE visualization under an SNR of -6 dB is illustrated in Fig. 8. As can be seen from the figure, DWDF-MFE demonstrates excellent feature clustering performance. Samples of the same category are tightly grouped with clear boundaries between categories, showing good separability. In contrast, the clustering effect of VAFCNN and MSCNN-FRDF is weak, with overlapping samples in some categories. In particular, in the MsACNN and MIFMN methods, a significant mixing of samples from different categories is observed, resulting in unclear category boundaries and a reduced ability to distinguish fault types.

These results indicate that DWDF-MFE efficiently leverages information from multiple sensing sources to learn more distinguishable features and demonstrates stronger abilities in both feature representation and classification. Compared with other methods, DWDF-MFE shows significant advantages in complex noise environments, and is able to recognize and distinguish various fault types more clearly, thus providing strong support for high-precision fault diagnosis.

### 4.2. Case 2: ZZU bearing dataset analysis

#### 4.2.1. Experimental dataset selection and preprocessing

The experimental dataset was obtained from the mechanical failure simulator testbed of Zhengzhou University. The setup consists of a motor, speed regulator, signal acquisition system, and accelerometer. 6205-type deep groove ball bearings were used as test subjects. Failures of varying severity were artificially introduced via electrical discharge machining (EDM). Triaxial accelerometers were mounted on the test bearing housing to capture vibration data in the horizontal (X), vertical (Y), and axial (Z) directions. Signals were sampled at 12 kHz. Test conditions included three rotational speeds (600, 900, and 1200 rpm) and torque levels (0, 3.65, and 7.3 Nm). The bearing states-normal, inner race fault, and outer race fault-were labeled as NC, IF, and OF, respectively.In addition, the dataset provides triaxial vibration measurements acquired under multiple speed-torque combinations, with stable sampling conditions and clearly distinguishable fault responses, which enable the construction of a consistent multi-class diagnostic task within this study. The experimental setup schematic for bearing fault diagnostic in Fig. 9. In this study, only the NC, IF, and OF classes of the ZZU dataset are used, as these representative fault modes provide complete measurements under all operating speeds considered in this work and form a consistent diagnostic setting for the analysis. Additional information on the selected samples is summarized in Table 7.

#### 4.2.2. Performance comparison across data sources

To evaluate the robustness of DWDF-MFE under noisy conditions, signals from different sensors were subjected to varying levels of noise interference. The single-channel data were obtained from vibration

**Table 9**

Assessing diagnostic performance of multi-sensor fusion methods in Case 2(%).

| Model Type | No Noise | −2dB | −4dB | −6dB | −8dB | −10dB |
|---|---|---|---|---|---|---|
| VAFCNN | 98.69 ± 0.54 | 96.86 ± 2.25 | 96.69 ± 3.17 | 97.28 ± 3.05 | 98.47 ± 1.99 | 98.75 ± 1.94 |
| MSCNN-FRDF | 98.14 ± 0.82 | 97.81 ± 1.39 | 97.53 ± 1.48 | 97.36 ± 1.71 | 98.06 ± 1.47 | 97.58 ± 1.11 |
| MsACNN | 99.17 ± 0.53 | 95.78 ± 2.42 | 96.86 ± 1.06 | 96.61 ± 1.39 | 96.39 ± 2.2 | 96.44 ± 1.3 |
| MIFMN | 99.72 ± 0.08 | 97.78 ± 0.78 | 95.11 ± 3.46 | 96.83 ± 1.93 | 95.47 ± 1.34 | 95.69 ± 1.76 |
| DWDF-MFE | **100.00** | **99.92 ± 0.04** | **99.91 ± 0.06** | **99.80 ± 0.10** | **99.72 ± 0.08** | **99.61 ± 0.08** |

signals along the X and Y axes, while "S-All" denotes the fused data from both directions processed using DWDF-MFE. It is noteworthy that the diagnostic model remains constant in this comparative analysis; the sole variation pertains to the input signal. All other experimental conditions remained consistent. Each experiment was conducted ten times to enhance result stability, with the accuracy from every run individually recorded. To finalize the evaluation, statistical metrics such as the average and standard deviation were calculated across all experimental runs. Table 8 reports the accuracy outcomes obtained from both standalone and fused sensor inputs under a range of noise levels (SNRs). The experimental findings demonstrate that multi-channel fusion (S-All) significantly surpasses single-channel signal diagnosis, exhibiting higher average accuracy and reduced standard deviation. This result demonstrates that the fusion strategy enables the extraction of reliable diagnostic features from multiple sensors, which in turn improves robustness by minimizing the impact of noise.

*4.2.3. Comparison with conventional fault diagnosis methods*

To further validate the superiority of DWDF-MFE, we compare it with five traditional conventional fault diagnosis methods, including ResNet, AlexNet, CNN, SAE, and MLP. In the comparison, all the models use the same hyper-parameter settings and input signals while simulating the effect of noise in real industrial scenarios, and the experiments are introduced with varying degrees of SNR noise interference.

Fig. 10 presents the diagnostic accuracy for individual models under various noise conditions, with the corresponding results analyzed as follows: The MLP and SAE perform well under noisy conditions with diagnostic accuracies of 99.86% and 99.94%, respectively. However, their performance deteriorates sharply under high noise conditions. For example, at SNR = −8 dB, the accuracy of MLP drops to 93.39%, while the accuracy of SAE slightly improves to 94.67%, which indicates their limitations in dealing with high-noise environments. AlexNet shows relatively better noise immunity by virtue of its feature extraction capability. However, they still struggle in high noise conditions. For example, AlexNet achieves an accuracy of 96.94% at SNR = 0 dB, but its accuracy drops significantly to 89.86% at SNR = −10 dB. The experiment revealed that ResNet did not perform as expected. Its diagnostic accuracy under noisy conditions is relatively low, especially at SNR = −8 dB and −10 dB, which are only 91.16% and 91.72%, respectively. Furthermore, its performance is even inferior to that of MLP and SAE. Compared with other models, CNN exhibits inferior performance and fails to maintain robustness under varying noise levels.

In contrast, the proposed DWDF-MFE method demonstrates significantly superior performance across all noise levels. Even under strong noise conditions (SNR = -10dB), DWDF-MFE still achieves an accuracy of 99.61%, a substantial improvement over alternative methods. This outcome demonstrates the DWDF-MFE method's remarkable robustness and resilience to noise in complex environments.

To better assess the performance of the DWDF-MFE model in noisy environments, experiments were carried out at an SNR of -6 dB, and the optimal results from each method were recorded. The confusion matrices obtained from the experiments are shown in Fig. 11, and these results enable a clear assessment of how accurately each model identifies the fault types.

In the experimental phase, the DWDF-MFE method showed excellent fault diagnosis performance, with only a few misclassifications in certain

fault types. These misclassifications can be disregarded. In contrast, alternative methods exhibited a greater number of misclassified samples, particularly AlexNet and CNN, whose performance in certain categories was suboptimal. Furthermore, ResNet and SAE exhibited higher confusion rates in classifying certain fault types, indicating a lack of discriminative power in their features.

*4.2.4. Comparison with multi-source fusion methods*

To assess the performance of the DWDF-MFE approach in multi-sensor fault diagnosis tasks, a detailed comparison with four other multisensor fusion diagnostic methods, including VAFCNN, MSCNN-FRDF, MsACNN, and MIFMN, is conducted. In the experiments, all models take the multi-sensor fusion data (S-All) produced by the DWDF-MFE method as input and are tested under various SNR levels to assess how each method performs in noisy and complex environments. To ensure a fair comparison, all models were trained using consistent hyperparameters, including the same optimizer , learning rate , batch size , number of epochs , and loss function . The diagnostic performance of different methods across varying SNR conditions is presented in Table 9.

Under noise-free conditions, all methods achieved diagnostic accuracies exceeding 98%, highlighting their effectiveness in accurately identifying fault types. Under elevated noise conditions, significant differences in diagnostic performance are observed across the methods. Specifically, DWDF-MFE achieves 99.61% accuracy under high noise conditions, while the other methods have relatively low accuracy. For example, MsACNN achieves 96.44% accuracy under these conditions, MSCNN-FRDF achieves 97.58% accuracy, MIFMN achieves 95.69% accuracy, and VAFCNN achieves 98.75% accuracy. These results show that DWDF-MFE exhibits excellent robustness in dealing with different levels of noise interference and can provide more accurate fault diagnosis in complex environments. In conclusion, DWDF-MFE proves effective in fault diagnosis scenarios involving multi-sensor data fusion, achieving high accuracy and robustness under both noiseless and noisy conditions. This validates the advantages of DWDF-MFE in complex and noisy environments, highlighting its applicability in real-world industrial fault diagnosis.

To evaluate how effectively each model separates different fault categories, t-SNE is utilized in this study for visualizing the learned feature distributions. The t-SNE visualization results under a noise level of −6dB are presented in Fig. 12. The visualization reveals that DWDF-MFE achieves well-defined feature clusters, reflecting strong discrimination between different fault categories. Same-class samples exhibit high intra-class compactness, and inter-class separability is clearly maintained. In contrast, the clustering effect of VAFCNN and MSCNN-FRDF is weak, and the samples of some categories overlap. In particular, MsACNN and MIFMN methods show significant mixing of samples from different categories with fuzzy category boundaries, resulting in the inability to effectively differentiate different fault categories.

**5. Conclusion**

This work proposed DWDF-MFE, a multi-channel fusion approach built on deep learning, to address challenges in diagnosing faults in rotating machinery under variable working environments. The network integrates data-level fusion using Dempster-Shafer (DS) Evidence Theory and feature-level fusion through a combination of MSCNN and SE

module. The DWDF-MFE network is designed to effectively handle multi-source signals, providing accurate fault diagnosis even under high noise conditions.

Experimental results show that DWDF-MFE maintains over 99 % classification accuracy on ZZU and SDUST bearing datasets, even under strong noise conditions, significantly outperforming traditional methods and other deep learning models. The t-SNE algorithm is employed to visualize the high-dimensional features extracted by the proposed DWDF-MFE method. The resulting distributions exhibit clear inter-class separability and intra-class compactness, providing intuitive evidence of the discriminative power of the model. These results validate the robustness and effectiveness of the proposed DWDF-MFE method in diagnosing faults under complex and noisy industrial conditions, demonstrating superior reliability compared to conventional approaches.

Future work may investigate the potential applicability of the DWDF-MFE framework to other types of rotating machinery, such as pumps and fans, under similar monitoring conditions, explore multi-level fusion techniques spanning the data, feature, and decision stages, incorporate heterogeneous sensor sources to enhance multi-modal diagnostic performance, and leverage transfer learning to improve adaptability across different working conditions.

## CRediT authorship contribution statement

**Ke Chen:** Conceptualization, Investigation, Resources, Supervision, Project administration, Validation, Funding acquisition, Writing – review & editing; **Feilong Zhou:** Methodology, Investigation, Software, Validation, Formal analysis, Data curation, Writing – original draft; **Fangfang Zhang:** Supervision, Writing – review & editing; **Kunjie Yu:** Funding acquisition, Validation, Supervision, Writing – review & editing; **Duo Yang:** Supervision, Visualization.

## Data availability

The data that has been used is confidential.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

## References

Castanedo, F. (2013). A review of data fusion techniques. *The Scientific World Journal*, *2013*, 704504. https://doi.org/10.1155/2013/704504.

Chen, G., Liu, Z., Yu, G., & Liang, J. (2021a). A new view of multisensor data fusion: Research on generalized fusion. *Mathematical Problems in Engineering*, *2021*, 5471242. https://doi.org/10.1155/2021/5471242.

Chen, X., Zhang, B., & Gao, D. (2021b). Bearing fault diagnosis base on multi-scale CNN and LSTM model. *Journal of Intelligent Manufacturing*, *32*(4), 971–987. https://doi.org/10.1007/s10845-020-01600-2.

Duan, X., Xue, L., Lei, C., & Li, J. (2023). Rolling bearing fault diagnosis method based on multi-information fusion characteristics under complex working conditions. *Applied Acoustics*, *214*, 109685. https://doi.org/10.1016/j.apacoust.2023.109685.

Duan, Z., Wu, T., Guo, S., Shao, T., Malekian, R., & Li, Z. (2018). Development and trend of condition monitoring and fault diagnosis of multi-sensors information fusion for rolling bearings: A review. *International Journal of Advanced Manufacturing Technology*, *96*(1–4), 803–819. https://doi.org/10.1007/s00170-017-1474-8.

Guo, J., Liu, C., Cao, J., & Jiang, D. (2021). Damage identification of wind turbine blades with deep convolutional neural networks. *Renewable Energy*, *174*, 122–133. https://doi.org/10.1016/j.renene.2021.04.040.

Guo, W., Li, X., & Shen, Z. (2024). A lightweight residual network based on improved knowledge transfer and quantized distillation for cross-domain fault diagnosis of rolling bearings. *Expert Systems with Applications*, *245*, 123083. https://doi.org/10.1016/j.eswa.2023.123083.

Guo, Z., Du, W., Liu, Z., Hu, T., Yu, Y., & Li, C. (2025). Few-shot sample multi-class incremental fault diagnosis for gearbox based on convolutional-attention fusion network. *Expert Systems with Applications*, *264*, 125918. https://doi.org/10.1016/j.eswa.2024.125918.

Han, B., Jiang, X., Wang, J. et al. (2023). A novel domain adaptive fault diagnosis method for bearings based on unbalance data generation. *IEEE Transactions on Instrumentation and Measurement*, *72*, 3519911. https://doi.org/10.1109/tim.2023.3284131.

Han, B., Yang, Z., Zhang, Z., Bao, H., Wang, J., Liu, Z., & Li, S. (2022). A novel rolling bearing fault diagnosis method based on generalized nonlinear spectral sparsity. *Measurement*, *198*, 111131. https://doi.org/10.1016/j.measurement.2022.111131.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770–778). https://doi.org/10.1109/cvpr.2016.90.

Hou, L., & Huang, Q. (2025). A smart WSNs node with sensor computing and unsupervised one-class SVM classifier for machine fault detection. *Measurement*, *242*, 115843. https://doi.org/10.1016/j.measurement.2024.115843.

Jais, I. K. M., Ismail, A. R., & Nisa, S. Q. (2019). ADAM optimization algorithm for wide and deep neural network. *Knowledge Engineering and Data Science*, *2*(1), 41–46. https://doi.org/10.17977/um018v2i12019p41-46.

Jiang, G., He, H., Yan, J., & Xie, P. (2019). Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox. *IEEE Transactions on Industrial Electronics*, *66*(4), 3196–3207. https://doi.org/10.1109/TIE.2018.2844805.

Kim, S.-G., Park, T., Kang, R., & Kwak, Y. (2025). Domain-collaborative multimodal transformer for fault diagnosis of rotating machines under noisy environments. *Advanced Engineering Informatics*, *68*, 103763. https://doi.org/10.1016/j.aei.2025.103763.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, *25*. https://doi.org/10.1145/3065386.

Lei, Y., Lin, J., He, Z., & Zuo, M. J. (2013). A review on empirical mode decomposition in fault diagnosis of rotating machinery. *Mechanical Systems and Signal Processing*, *35*(1), 108–126. https://doi.org/10.1016/j.ymssp.2012.09.015.

Liu, X., Zhou, Q., Zhao, J., Shen, H., & Xiong, X. (2019). Fault diagnosis of rotating machinery under noisy environment conditions based on a 1-d convolutional autoencoder and 1-d convolutional neural network. *Sensors*, *19*(4), 972. https://doi.org/10.3390/s19040972.

Ming, Y., Shao, H., Cai, B., & Liu, B. (2024). Rgfc-forest: An enhanced deep forest method towards small-sample fault diagnosis of electromechanical system. *Expert Systems with Applications*, *238*, 122178. https://doi.org/10.1016/j.eswa.2023.122178.

Qi, Y., Shen, C., Wang, D., Shi, J., Jiang, X., & Zhu, Z. (2017). Stacked sparse autoencoder-based deep network for fault diagnosis of rotating machinery. *IEEE Access*, *5*, 15066–15079. https://doi.org/10.1109/access.2017.2728010.

Rana, A., Rawat, A. S., Bijalwan, A., & Bahuguna, H. (2018). Application of multi layer (perceptron) artificial neural network in the diagnosis system: A systematic review. In *2018 International conference on research in intelligent and computing in engineering (RICE)* (pp. 1–6). http://dx.doi.org/10.1109/RICE.2018.8509069.

Shao, H., Lin, J., Zhang, L., Galar, D., & Kumar, U. (2021). A novel approach of multi-sensory fusion to collaborative fault diagnosis in maintenance. *Information Fusion*, *74*, 65–76. http://dx.doi.org/10.1016/j.inffus.2021.03.008.

Song, S., Zhang, S., Dong, W., Li, G., & Pan, C. (2024). Multi-source information fusion meta-learning network with convolutional block attention module for bearing fault diagnosis under limited dataset. *Structural Health Monitoring*, *23*(2), 818–835. https://doi.org/10.1177/14759217231176045.

Tong, J., Liu, C., Bao, J., Pan, H., & Zheng, J. (2023a). A novel ensemble learning-based multisensor information fusion method for rolling bearing fault diagnosis. *IEEE Transactions on Instrumentation and Measurement*, *72*, 1–12. https://doi.org/10.1109/tim.2022.3225910.

Tong, J., Liu, C., Pan, H., & Zheng, J. (2022). Multisensor feature fusion based rolling bearing fault diagnosis method. *Coatings*, *12*(6), 866. https://doi.org/10.3390/coatings12060866.

Tong, J., Liu, C., Zheng, J., & Pan, H. (2023b). Multi-sensor information fusion and coordinate attention-based fault diagnosis method and its interpretability research. *Engineering Applications of Artificial Intelligence*, *124*, 106614. https://doi.org/10.1016/j.engappai.2023.106614.

Wan, S., Li, T., Fang, B., Yan, K., Hong, J., & Li, X. (2023). Bearing fault diagnosis based on multisensor information coupling and attentional feature fusion. *IEEE Transactions on Instrumentation and Measurement*, *72*, 1–12. https://doi.org/10.1109/tim.2023.3269115.

Wang, J., Fu, P., Zhang, L., Gao, R. X., & Zhao, R. (2019). Multilevel information fusion for induction motor fault diagnosis. *IEEE/ASME Transactions on Mechatronics*, *24*(5), 2139–2150. http://doi.org/10.1109/TMECH.2019.2928967.

Wang, J., Zhang, X., Zhang, Z., Han, B., Jiang, X., Bao, H., & Jiang, X. (2024). Attention guided multi-wavelet adversarial network for cross domain fault diagnosis. *Knowledge-Based Systems*, *284*, 111285. https://doi.org/10.1016/j.knosys.2023.111285.

Wang, T., Xu, X., Pan, H., Chang, X., Yuan, T., Zhang, X., & Xu, H. (2022). Rolling bearing fault diagnosis based on depth-wise separable convolutions with multi-sensor data weighted fusion. *Applied Sciences*, *12*(15), 7640. https://doi.org/10.3390/app12157640.

Wang, X., Mao, D., & Li, X. (2021). Bearing fault diagnosis based on vibro-acoustic data fusion and 1d-CNN network. *Measurement, 173*, 108518. https://doi.org/10.1016/j.measurement.2020.108518.

Wang, Z., Liu, Y., Bai, R., Chen, H., Li, J., Chen, X., Yao, L., Zhao, J., & Chu, F. (2025). Multi-modal multi-scale multi-level fusion quadrant entropy for mechanical fault diagnosis. *Expert Systems with Applications, 281*, 127715. https://doi.org/10.1016/j.eswa.2025.127715.

Xia, M., Li, T., Xu, L., Liu, L., & De Silva, C. W. (2018). Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks. *IEEE/ASME Transactions on Mechatronics, 23*(1), 101–110. https://doi.org/10.1109/TMECH.2017.2728371.

Xie, T., Huang, X., & Choi, S. K. (2021). Multi-sensor data fusion for rotating machinery fault diagnosis using residual convolutional neural network. In *Proceedings of the ASME design engineering technical conference*. https://doi.org/10.1115/DETC2021-67406.

Xu, Z., Tang, X., & Wang, Z. (2023). A multi-information fusion vit model and its application to the fault diagnosis of bearing with small data samples. *Machines, 11*(2), 277. https://doi.org/10.3390/machines11020277.

Yang, C., Liu, J., Zhou, K., Jiang, X., & Zeng, X. (2022). An improved multi-channel graph convolutional network and its applications for rotating machinery diagnosis. *Measurement, 190*, 110720. https://doi.org/10.1016/j.measurement.2022.110720.

Yoo, Y. J. (2019). Fault detection of induction motor using fast fourier transform with feature selection via principal component analysis. *International Journal of Precision Engineering and Manufacturing, 20*(10), 1543–1552. https://doi.org/10.1007/s12541-019-00176-z.

Yu, G., Wu, P., Lv, Z., Hou, J., Ma, B., & Han, Y. (2023). Few-shot fault diagnosis method of rotating machinery using novel MCGM based CNN. *IEEE Transactions on Industrial Informatics, 19*(11), 10944–10955. https://doi.org/10.1109/TII.2023.3242813.

Zeng, F., Li, Z., Zhou, Z., & Du, S. (2019). Fault classification decision fusion system based on combination weights and an improved voting method. *Processes, 7*(11), 783. https://doi.org/10.3390/pr7110783.

Zhang, X., Wang, J., Zhang, Z. et al. (2024). Integrated decision-making with adaptive feature weighting adversarial network for multi-target domain compound fault diagnosis of machinery. *Advanced Engineering Informatics, 62*, 102730. https://doi.org/10.1016/j.aei.2024.102730.

Zhang, Y., Li, C., Su, L., & Zhang, C. (2025a). Degradation trend prediction for centrifugal blowers based on multi-sensor information fusion and attention mechanism. *Expert Systems with Applications, 276*, 127195. https://doi.org/10.1016/j.eswa.2025.127195.

Zhang, Y., Zhao, X., Peng, Z., Xu, R., & Chen, P. (2025b). Wd-kantf: An interpretable intelligent fault diagnosis framework for rotating machinery under noise environments and small sample conditions. *Advanced Engineering Informatics, 66*, 103452. https://doi.org/10.1016/j.aei.2025.103452.