Eng. Sci., 2025, **38**, 1882

Engineered Science
DOI: https://dx.doi.org/10.30919/es1882

A Multi-Sensor Information Fusion Fault Diagnosis Method for Hydropower Turbine Bearings Based on Multi-Scale Spatiotemporal Graph Neural Networks

Yongzhong Zeng,[1,2] Liuyan Tang,[1,2] Shuquan Xiao,[3] Zhilin Dong,[4,*] Xueyi Li,[3,5,*] Tianyang Wang,[5] Xiaobing Liu,[2] Yuanjiang Ma[6] and Fulei Chu[5]

Abstract

Multi-sensor information fusion plays a critical role in hydroturbine bearing fault diagnosis. Traditional approaches often struggle to capture complex spatiotemporal dependencies, limiting their ability to fully exploit the latent features in multi-sensor data. To address this challenge, this paper proposes a novel Multi-Scale Spatiotemporal Graph Neural Network (MSTGNN) for hydroturbine bearing fault diagnosis, aiming to comprehensively model the spatiotemporal correlations within sensor data. First, an adaptive receptive field convolutional layer is designed to extract key information from each sensor. Then, a dynamic graph structure is constructed to capture spatial dependencies among sensors. A multi-scale spatiotemporal convolution module is subsequently introduced to extract temporal features, enhancing the model's capability to identify fault patterns across different time scales. Extensive experiments on two bearing case studies demonstrate that MSTGNN achieves outstanding diagnostic performance, with accuracies of 99.96% on the Harbin Institute of Technology (HIT) bearing dataset and 99.27% on a thrust bearing dataset from a hydropower station in Sichuan Province. These results exceed the best-performing baseline method by 2.29% and 1.13%, respectively. Finally, ablation studies confirm the effectiveness of each proposed component, further validating the potential of MSTGNN in hydroturbine bearing fault diagnosis.

Keywords: Multi-sensor information fusion; Graph neural network; Spatiotemporal correlation; Dynamic graph structure.
Received: 14 September 2025; Revised: 30 October 2025; Accepted: 10 November 2025
Article type: Research article.

1. Introduction

With the ongoing global transition toward clean energy, hydropower—recognized as a renewable, mature, and highly efficient energy source—continues to hold a central position in the global energy mix. According to a report released by the International Renewable Energy Agency (IRENA), by the end of 2024, the total global installed capacity of renewable energy had reached 4,448 GW, as illustrated in Fig. 1. Among this, hydropower accounted for 1,283 GW, representing 29% of the total capacity, with the majority of growth driven by China.[1]

As the core mechanical component of a hydropower station, the hydroturbine plays a crucial role in determining the overall energy efficiency and operational safety of the generating unit.[2] Among its components, the hydroturbine bearing serves as a critical support element in the high-speed rotating system and operates under harsh conditions such as high load, high humidity, and intense vibration. These factors make it highly susceptible to failures including wear, excessive vibration, and lubrication degradation.[3,4] Hydropower, as a backbone energy source among renewables, is essential for enhancing grid regulation capabilities and increasing the share of clean energy. However, due to the complex and dynamic hydraulic environments in which hydropower units operate over extended periods, they are prone to faults such as rotor imbalance, shaft misalignment, and frictional contact. These issues may reduce operational efficiency or, in severe cases, lead to safety hazards.[5,6] Therefore, developing a fault diagnosis model with high accuracy and strong robustness is key to enabling intelligent maintenance and predictive condition monitoring of hydropower units.

Currently, although end-to-end intelligent diagnosis methods based on deep learning have achieved remarkable progress, most studies still rely on single-sensor data, which results in limited information dimensions, insufficient feature representation, and inadequate adaptability to complex

[1]*Key Laboratory of Fluid and Power Machinery, Ministry of Education, Xihua University, Chengdu, 610039, China*
[2]*School of Energy and Power Engineering, Xihua university, Chengdu, 610039, China*
[3]*College of Mechanical and Electrical Engineering, Northeast Forestry University, Harbin, 150040, China*

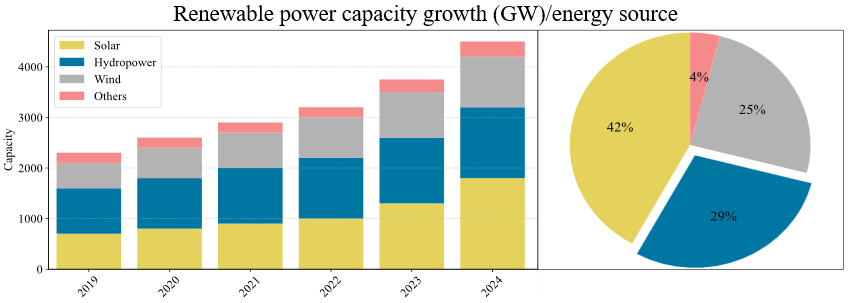

Fig. 1: Growth in renewable energy capacity.

operating conditions. A single signal channel often fails to comprehensively characterize the operating state of equipment, while traditional approaches face limitations in capturing cross-sensor correlations and multi-scale spatiotemporal dependencies, thereby constraining diagnostic performance. In contrast, multi-sensor fusion diagnosis methods enable complementary utilization of spatiotemporal information, providing an effective approach to improving fault identification capability and enhancing adaptability under complex conditions.[7,8] By integrating models such as CNNs, GRUs, and Transformers, multi-sensor frameworks exhibit enhanced capabilities in mining non-stationary signals and weak fault features. This research direction not only contributes theoretical innovation but also provides practical engineering value in ensuring the safe operation of hydropower equipment and promoting the development of intelligent power stations.

With the increasing scale and parameterization of hydropower units, their operating conditions have become significantly more complex. Traditional signal processing methods based on single-sensor data-such as Fourier transform, wavelet packet decomposition, and empirical mode decomposition-often struggle to effectively extract accurate fault features from non-stationary vibration signals, thereby leading to less reliable diagnostic results.[9,10] As a result, the diagnostic accuracy deteriorates, especially in scenarios involving coexisting multiple faults or the early stages of weak faults, where single-source information is prone to ambiguity and misjudgment. In recent years, the development of deep learning technology has significantly advanced methods for mechanical fault diagnosis.[11,12] Intelligent algorithms based on models such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and long short-term memory networks (LSTMs) can effectively extract features and model complex nonlinear relationships, and have been widely applied to condition monitoring and fault diagnosis of rotating machinery.[13] However, the performance of deep learning models still heavily depends on the quality and completeness of input data features. Relying solely on single-sensor signals often fails to comprehensively capture the operating state of complex equipment, thereby limiting the model's generalization ability and robustness.

To enhance diagnostic accuracy and stability, multi-sensor fusion technologies have emerged as a prominent research focus. Guo *et al.*[14] proposed a fusion method based on the improved cyclic spectral covariance matrix (ICSCM), which effectively captures the interactions among multi-source signals by combining current signals with vibration data. However, this method still has certain limitations when dealing with non-stationary signals or high-dimensional channel information. Wang *et al.*[15] Wang *et al.* extracted richer local feature information by altering the input channel dimensions and developed a collaborative self-attention module to enhance the model's focus on important channels. However, systematic research on multi-sensor feature fusion and temporal dependency modeling is still lacking. Yan *et al.*[16] designed a Cross-Domain Time-Frequency Attention Fusion Network (CDTFAFN) that leverages the complementary nature of acoustic and vibration signals to maintain high recognition rates even under severe noise conditions. Li *et al.*[17] proposed a multi-level fusion strategy that combines graph convolution with an information entropy-based decision mechanism, significantly improving the robustness and anti-interference capability of the fusion diagnosis system. Zhou *et al.*[18] extracted time-domain and frequency-domain features in parallel and introduced an attention mechanism to emphasize key feature dimensions, ultimately building a fusion model suitable for multi-fault pattern recognition.

In addition, Sun *et al.*[19] enhanced the coupling learning between different sensors by adding a 2D convolutional layer, thereby improving the model's robustness and adaptability. Shao *et al.*[20] developed a stacked multi-granularity scanning module, which enhances the diversity of feature extraction by integrating the advantages of multiple base learners and effectively filters out irrelevant information. Yan *et al.*[21] employed a spatial–temporal fragment construction mechanism to reconstruct multi-channel vibration signals into feature sequences with strong spatial–temporal correlations, effectively alleviating the problem of coupled information loss

[4]*College of Engineering, Zhejiang Normal University, Jinhua, 321004, China*
[5]*Department of Mechanical Engineering, Tsinghua University, Beijing, 100084, China*
[6]*Yingxiuwan Hydro Power Plant, State Grid Sichuan Electric Power Company, Chengdu, 611830, China*
[*]*Email:* d18133679022@zjnu.edu.cn (Zhilin Dong), lixueyiphm@163.com (Xueyi Li)

caused by traditional independent processing methods. Wang *et al.*[22] introduced a CNN-based method that converts multi-source vibration signals into images and integrates them for recognition, significantly improving diagnostic accuracy. Zhang *et al.*[23] conducted a joint analysis of infrared images and vibration signals to achieve the identification and classification of coupled faults. Zhao *et al.*[24] employed Dempster-Shafer (D-S) evidence theory to perform confidence-weighted fusion of multi-source signals, greatly reducing the misjudgment rate. Guo *et al.*[25] combined wavelet transform with an enhanced wavelet reconstruction mechanism, effectively improving the separability and robustness of multi-channel signals. He *et al.*[26] proposed a dual-correlation adaptive gated graph convolutional network, which first employs a predefined dual-correlation graph to transform feature channel data into source and target matrices for creating an adaptive graph, thereby constructing an adaptive graph structure to enhance feature modeling capability. In summary, research on hydropower and rotating machinery fault diagnosis based on multi-sensor fusion has established a comprehensive technical framework encompassing data acquisition, feature extraction, fusion strategies, and deep modeling. This research direction provides essential support for building intelligent operation and maintenance systems with strong anti-interference capability, high diagnostic accuracy, and real-time processing capacity.

Graph Neural Networks (GNNs)[27] have provided new avenues for multi-sensor information fusion due to their ability to explicitly model the spatiotemporal dependencies between sensor nodes. For example, Li *et al.*[28] proposed a Dynamic Graph Attention Network (DyGAT), which adaptively adjusts sensor node weights to achieve efficient fusion of cross-sensor features, thereby improving the accuracy of gearbox fault diagnosis in wind turbines. Liu *et al.*[29] designed a Spatiotemporal Graph Convolutional Network (ST-GCN) that combines dynamic temporal information with spatial topological structure, demonstrating exceptional feature extraction capabilities in rolling bearing fault diagnosis. However, existing GNN-based methods still have certain limitations:

(1) Inadequate adaptability of the fixed convolution receptive field. Traditional convolution operations use static-sized kernels to extract features from local regions, which struggle to adapt to the non-stationary characteristics such as rapid frequency changes and time-varying impacts in hydroturbine bearing signals. This leads to the weakening or omission of key diagnostic information during the convolution process, negatively affecting the model's responsiveness to early-stage faults.

(2) Limited ability to model multi-source dynamic relationships. Current multi-sensor data processing frameworks still rely on a static input assumption, ignoring the inherent coupling patterns of sensor-acquired signals over time. This results in the failure to effectively integrate spatial layout and temporal evolution characteristics, thus restricting the expression and utilization of cross-sensor joint features.

(3) Insufficient feature redundancy control and node attribute representation. Existing multi-sensor graph modeling strategies primarily focus on graph structure construction, lacking deep exploration and optimization mechanisms for node-level attribute information.

To address the challenges in multi-sensor signal diagnosis of hydroturbine systems, such as the difficulty in extracting multi-scale impact features, insufficient modeling of dynamic spatiotemporal dependencies, and interference from redundant information, this paper proposes a Spatiotemporal Graph Neural Network Framework (MSTGNN) for multi-sensor applications. By introducing a dynamic receptive mechanism, adaptive graph convolution, and a sparsity-preserving strategy, the proposed method achieves a synergistic enhancement of feature representation and diagnostic performance, ultimately improving the accuracy of hydroturbine bearing fault diagnosis. The main contributions of this work are summarized as follows:

(1) Proposed a dynamic receptive field convolution mechanism to enhance the model's adaptive perception of non-stationary signal features.

(2) Developed a multi-scale graph convolution structure integrated with contrastive learning, improving the accuracy of modeling multi-source spatiotemporal dependencies.

(3) Introduced a sparsity-guided deep non-negative matrix factorization method, achieving a balance between information compression and the preservation of key features.

(4) Constructed a hydroturbine fault diagnosis framework, validating its excellent performance on a real hydroturbine bearing dataset.

The remainder of this paper is organized as follows: In Section 2, the relevant theoretical background and related work are introduced. In Section 3, the proposed method framework is described in detail. Section 4 presents experimental validation using the Harbin Institute of Technology (HIT) bearing dataset and the thrust bearing dataset from a hydropower station in Sichuan Province. Ablation studies are conducted at the end of each subsection to validate the effectiveness of each module. The conclusions are provided in Section 5.

2. Basic concepts of graph neural networks

2.1 GNN framework

With the development of graph neural networks (GNNs), researchers have proposed various general frameworks, such as message passing neural networks (MPNNs),[30] hypergraph neural networks (HGNNs),[31] and graph convolutional networks (GCNs).[32] Among them, GCN as a fundamental model, performs convolution operations on graphs based on spectral graph theory, effectively capturing the relationships between nodes and their neighbors. It strikes a balance between computational complexity and interpretability, and thus serves as the foundation for many variants (*e.g.*, GAT, GraphSAGE).

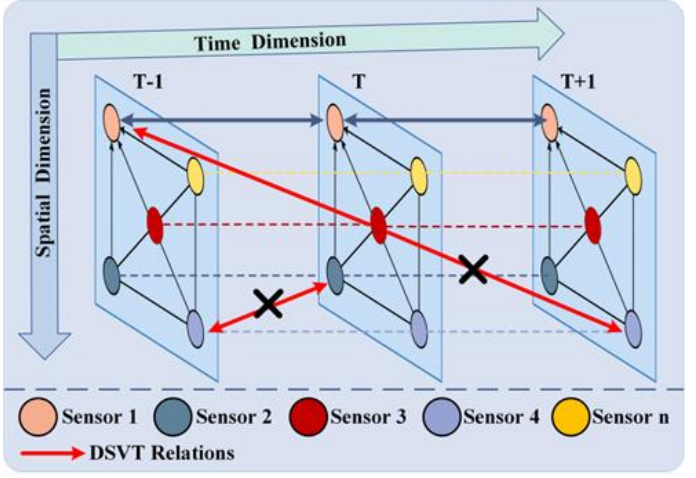

Fig. 2: Multi-sensor construction of spatiotemporal graph.

In GNN modeling, a graph can be represented as $G = (V, E, X)$, where V denotes the set of nodes, E the set of edges, and X the node feature matrix. Through the message passing mechanism, node features are iteratively updated among adjacent nodes, achieving the integration of local information and global structure. This mechanism enables flexible modeling of the correlation structures among multiple sensors in complex systems and enhances the extraction capability of spatiotemporal features.

2.2 Fusion of multi-sensor data

In hydroturbine bearing fault diagnosis, the vibration signals of bearings are often influenced by environmental factors and the operating conditions of the equipment. A single sensor is insufficient to comprehensively capture the fault features. Therefore, Multi-Sensor Data Fusion (MSDF) has become a key technology for enhancing equipment health monitoring and fault prediction capabilities.[33,34] By deploying multiple sensors at different locations on the turbine and collecting data from various physical channels, the bearing condition can be analyzed from multiple perspectives, providing richer and more accurate fault information.

Current spatiotemporal modeling methods for multi-sensor data can be broadly categorized into two types. The first focuses on temporal dependencies using models such as RNNs and LSTMs, but mainly relies on single-sensor sequences and overlooks inter-sensor interactions, limiting joint feature extraction under complex turbine conditions.[35] The second emphasizes spatial features through CNNs or SOM, sometimes combined with LSTM for sequential modeling. While effective in capturing spatial-temporal features, these approaches often rely on fixed convolution kernels, which restrict generalization across different fault modes.

Existing methods often assume fixed spatial relationships among sensors and focus only on features within the current time window, neglecting dynamic changes across time steps, which leads to incomplete spatiotemporal modeling. GNNs can efficiently capture sensor relationships through flexible graph structures. For example, Zhao *et al.*[36] proposed a multi-graph convolutional network combining spatiotemporal attention and dynamic graphs for multi-sensor data fusion. However, many GNN-based approaches still rely on predefined adjacency matrices, making it difficult to adapt to dynamic correlation changes under different fault modes, thereby limiting their ability to fully model spatiotemporal dependencies. As illustrated in Fig. 2, sensor signals may remain relatively independent at the early stage of a fault but become more correlated as the fault progresses.

In hydroturbine bearing fault diagnosis, the correlation between sensor signals varies across fault stages. This study proposes an adaptive receptive field convolution layer that dynamically adjusts the convolution kernel size to more effectively capture multi-scale features. Combined with a GNN, the model captures complex spatiotemporal dependencies and, by integrating global and local information, strengthens relationships between distant sensors, thereby improving fault diagnosis accuracy.

3. Method

3.1 Adaptive receptive field convolutional layer (ARFConv)

Traditional convolutional layers have a fixed receptive field, typically expanding the receptive field by adjusting hyperparameters such as the kernel size and stride. However, for bearing fault diagnosis tasks, the time-frequency features of the signals may exhibit highly variable local structures and multiscale characteristics, which makes it difficult for fixed receptive field convolutions to fully capture key features at different scales. Therefore, in this paper, we design a new convolutional layer-Adaptive Receptive Field Convolutional Layer, as shown in Fig. 3, to efficiently extract features from multi-sensor bearing vibration signals.

First, the width of the convolution kernel is learned, and its mathematical formulation is given in Eq. (1):

$$y_i = f_\theta(X) \tag{1}$$

where, $X \in \mathbb{R}^{L \times C}$, C represents the number of input data channels, where C=3, and L denotes the length of the signal in each channel. f_θ is the function for predicting the convolution

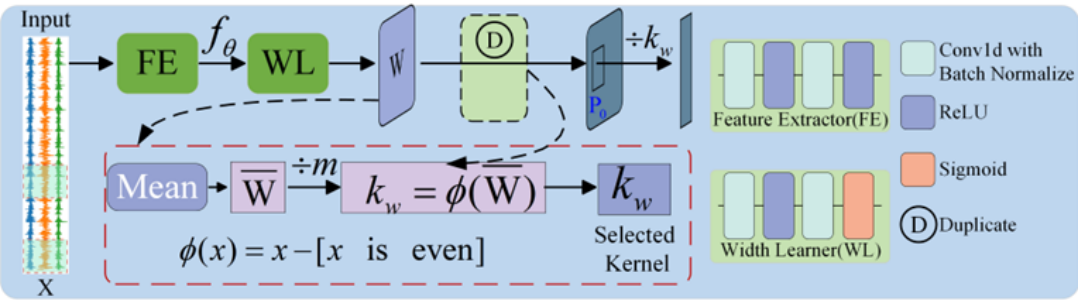

Fig. 3: Structure of ARFConv.

kernel width, and the output feature map is represented as $y_i \in \mathbb{R}^{L\times C}$.

Secondly, the number of sampling points for the convolution kernel is selected as k_w, and the width feature is replicated k_w times. The average value of y is computed to obtain the learned average width of the convolution kernel. The mathematical expression for the number of sampling points in the convolution kernel The mathematical expression for the number of sampling points in the convolution kernel is given in Eq. (2):

$$k_w = \phi(\frac{\overline{y}}{m}) \tag{2}$$

where, m represents the fine-tuning coefficient that maps the convolution kernel width to the number of sampling points. The mathematical expression for function $\phi()$ is given as follows:

$$\phi(x) = x - [x \text{ is even}] \tag{3}$$

when the width of the convolution kernel is given, a larger m value results in fewer sampling points, leading to a sparser signal distribution. From Eq. (3), it can be observed that we select a convolution kernel with an odd number of sampling points. If $\overline{y}/m$ is an even number, we choose the nearest smaller odd number. Finally, the sampling points are determined by N=k_w.

In standard convolution, this process involves sampling from the input feature map X using a regular grid G, followed by a weighted summation of these sampled values with weights w. The mathematical expression is given in Eq. (4):

$$G = \{(-1,-1),(-1,0),\cdots,(1,0),(1,1)\} \tag{4}$$

The standard convolution operation at position p_0 can be expressed as Eq. (5):

$$y(p_0) = \sum_{g_n \in G} w(g_n) \cdot x(p_0 + g_n) \tag{5}$$

where, y represents output feature map, w represents convolution kernel parameters, g_n indicates the offset relative to position p_0.

For the adaptive receptive field convolutional layer proposed in this paper, $G \in \mathbb{R}^{k_w}$ represents the offset matrix. The convolution operation is expressed in Eq. (6):

$$y(p_0) = \sum_{r_n \in R} w(r_n) \cdot t(p_0 + r_n) \tag{6}$$

where, the elements in G are represented by g_i, defined as $g_i = (2i - k_w - 1)/2$, and the offset matrix at positions p_0 is denoted by R, defined as $R = Z_0 \odot G$. Z_0 represents the scale matrixp_0, and $\odot$ denotes element-wise multiplication.

Hydroturbine bearing fault signals typically contain periodic impacts and non-stationary features. Fixed-size convolution kernels struggle to effectively capture features at different scales. Therefore, we apply the adaptive receptive field convolution layer defined above to feature extraction, allowing for more flexible capture of fault characteristics across varying scales.

3.2 Multi-Scale chebyshev convolution and contrastive learning

This paper uses a Graph Generation Layer (GGL) to learn the dynamic matrix A from multi-sensor input data and construct the instance graph.[37] The process is shown in Fig. 4.

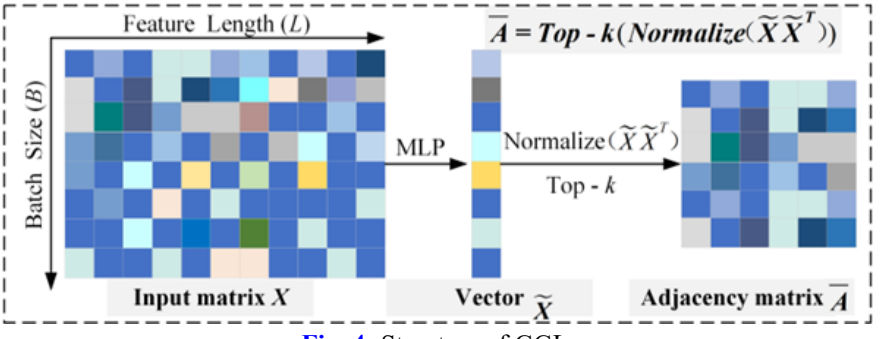

Fig. 4: Structure of GGL.

First, the extracted feature matrix is input into a Multi-Layer Perceptron (MLP) for nonlinear mapping. Then, by multiplying the MLP output features with their transpose matrix, the adjacency matrix is generated, thereby constructing the graph structure. The specific expression is given in Eq. (7):

$$A = \mathrm{MLP}(X_{\mathrm{input}}) \cdot \mathrm{MLP}(X_{\mathrm{input}})^T \tag{7}$$

To improve the model's capacity to capture fault propagation patterns, this paper introduces a cross-time-step sensor connection mechanism, allowing information from the same sensor at different time steps to interact, thereby better modeling temporal features.[38]

Specifically, in the process of constructing the adjacency matrix mentioned above, time information is introduced, and the structure along the time dimension is defined in Eq. (8):

$$A_{\mathrm{time}}^{(t)} = \alpha A^{(t)} + (1-\alpha)A^{(t-1)} \tag{8}$$

where, $A^{(t)}$ and $A^{(t-1)}$ represent the adjacency matrices at the time steps t and $t-1$, respectively. α is a balancing coefficient ($0 \le \alpha \le 1$), used to adjust the contribution ratio between the present and past time steps.

Then, the cross-time-step information is added to the adjacency matrix $A^{(t)}$, as specifically expressed in Eq. (9):

$$A_{\mathrm{final}}^{(t)} = A^{(t)} + A^{(t,t-1)} \tag{9}$$

Through the above process, each sensor at a given time step not only establishes connections with other sensors at the current time step but also interacts with relevant sensors from the previous time step, thereby more effectively modeling the fault propagation patterns across time.

In this model, Chebyshev convolutions of different orders (K=1,2,3) are employed to capture multi-scale features [39]. The specific structure is shown in Fig. 5. For the multi-sensor input feature X, the output of the multi-scale Chebyshev convolution given in Eq. (10):

$$H_{\mathrm{multi}} = \underbrace{(\sum_{k=0}^{k_1} T_k(\tilde{L})XW_k^{(1)}}_{\text{Local (K=1)}} \| \underbrace{\sum_{k=0}^{k_2} T_k(\tilde{L})XW_k^{(2)}}_{\text{Medium (K=2)}} \| \underbrace{\sum_{k=0}^{k_3} T_k(\tilde{L})XW_k^{(3)})}_{\text{Global (K=3)}} \tag{10}$$

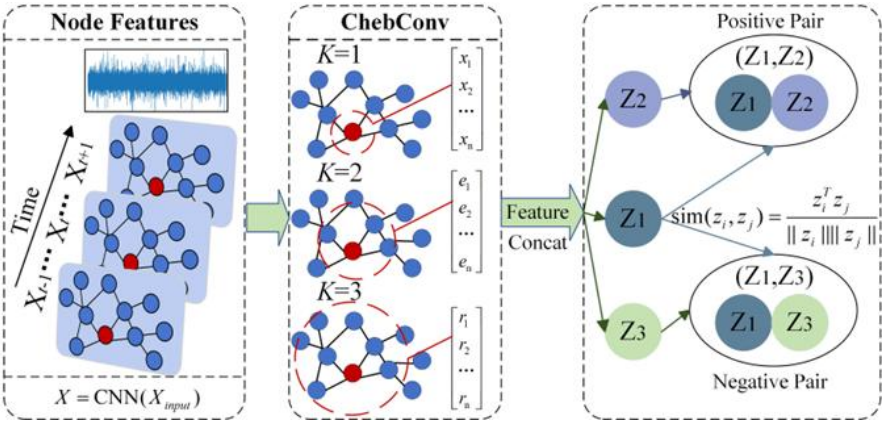

Fig. 5: Multi-scale Chebyshev convolution.

where, $\|$ denotes feature concatenation, $W_k^{(m)} \in \mathbb{R}^{C \times d_m}$ represents the parameter matrix for each scale.

During model training, a contrastive learning strategy is introduced to enhance the model's discriminative capability by constructing positive and negative sample pairs. The contrastive loss function is based on the Normalized Temperature-scaled Cross Entropy loss (NT-Xent),[40] and is defined in Eq. (11):

$$L_{\mathrm{contrastive}} = -\sum_{(i,j)\in P} log \frac{exp(\mathrm{sim}(z_i, z_j)/T)}{\sum_{k \ne i} exp(\mathrm{sim}(z_i, z_k)/T)} \tag{11}$$

where, z_i represents the node embedding after Multi-Scale Chebyshev processing, P denotes the index set of positive sample pairs, T is the temperature coefficient. $\mathrm{sim}(z_i, z_j)$ represents the similarity measure.

In addition, a negative contrastive loss term is introduced to prevent the model from overfitting, as specifically expressed in Eq. (12):

$$L_{\mathrm{negative}} = -\sum_{(i,j)\in N} log(1 - sin(z_i, z_j)) \tag{12}$$

The overall loss function is defined in Eq. (13):

$$L = \lambda_{\mathrm{rec}} L_{\mathrm{rec}} + \lambda_{\mathrm{conc}} L_{\mathrm{contrastive}} + \lambda_{\mathrm{negc}} L_{\mathrm{negative}} + \lambda_{\mathrm{reg}} L_{\mathrm{reg}} \tag{13}$$

where, $\lambda_{\mathrm{rec}}, \lambda_{\mathrm{conc}}, \lambda_{\mathrm{negc}}, \lambda_{\mathrm{reg}}$ is a hyperparameter used to adjust the weight of each loss term.

3.3 Deep non-negative matrix factorization

To capture deeper information from the original network structure, this paper reconstructs the adjacency matrix A obtained earlier into a deeper structure to characterize the information representation at different levels. Specifically, the

Deep Non-negative Matrix Factorization (DNMF) method is used, as shown in Fig 6, to extract community structures layer by layer and optimize node representations. The specific expression is given in Eq. (14):

$$\min_{U_i,V_p} L_D = ||A - U_1U_2 \dots U_pV_p||_F^2 \quad s.t. V_p \geq 0, U_i \geq 0, \forall i = 1,2,\dots,p \tag{14}$$

where, $V_p \in \mathbb{R}_+^{r\times n}$, $U_i \in \mathbb{R}_+^{r_{i-1}\times r_i}(i = 1,2,\dots,p)$. $||||_F$ represents the Frobenius norm of the matrix. Each matrix $U_i(i = 1,2,\dots,p)$ can be viewed as the ith feature matrix containing information at various levels. Matrix V_p represents features after deep transformation, with each column reflecting the trend of a node across different communities.

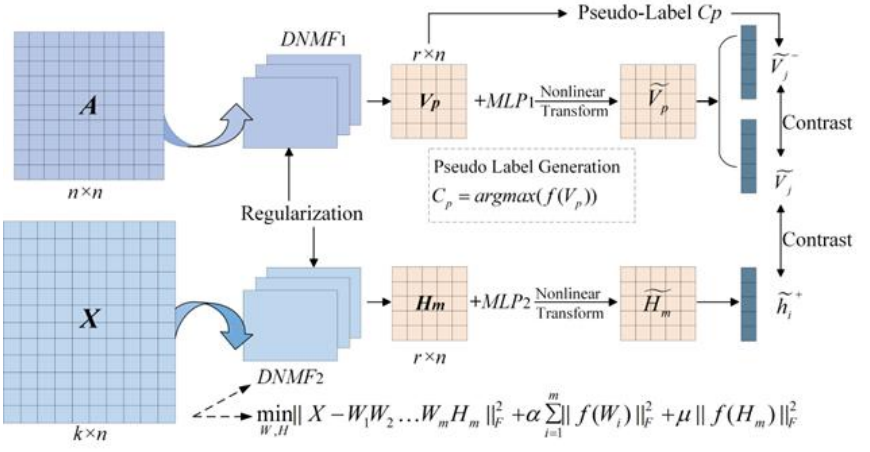

Fig. 6: Deep Non-negative Matrix Factorization (DNMF).

To optimize the above decomposition problem and further enhance the model's generalization ability, this paper introduces non-negative constraints and regularization terms to convert the constrained optimization problem into an unconstrained one.[41] First, we define a transformation function f for any given matrix $B \in \mathbb{R}^{a\times b}$ that satisfies the following conditions in Eq. (15):

$$f(B) = \begin{cases} B_{ij}, B_{ij} < 0 \\ 0, B_{ij} \geq 0. \end{cases} \tag{15}$$

In short, the function f is used to convert the positive elements of the input matrix to 0, while leaving the negative elements unchanged. The penalty term for matrix $U_i(i = 1,2,\dots,p)$ is defined in Eq. (16):

$$\begin{cases} \min_{U_i} f(U_i)_F^2 \\ \min_{V_p} f(V_p)_F^2 \end{cases} \tag{16}$$

By combining Eq. (16), we transform the optimization problem in Eq. (14) into the following unconstrained objective function in Eq. (17):

$$\min_{U_i,V_p} L_A = ||A - U_1U_2 \dots U_pV_p||_F^2 + \alpha\left(\sum_{i=1}^{p} ||f(U_i)||_F^2 + ||f(V_p)||_F^2\right) \tag{17}$$

where, $\alpha > 0$ represents the coefficient of the non-negative penalty. For the network's node feature matrix X, the following objective function is obtained in Eq. (18):

$$\min_{W_j,H_m} L_X = ||X - W_1W_2 \dots W_mH_m||_F^2 + \alpha\left(\sum_{j=1}^{m} ||f(W_j)||_F^2 + ||f(H_m)||_F^2\right) \tag{18}$$

The objective function for the DNMF term is obtained in Eq. (19):

$$\min_{U_i,V_p,W_j,H_m} L_{DNMF} = L_A + L_X \tag{19}$$

Furthermore, to maintain the intrinsic topological relationships between nodes in the deep-layer representation, this paper introduces a graph regularization strategy.[42] This strategy ensures that the embedding representations genuinely reflect the network geometry by minimizing the trace norm of the graph Laplacian matrix, as shown in Eq. (20):

$$\frac{1}{2}\sum_i\sum_j A(i,j)||V_p(:,i) - V_p(:,j)||_2^2 = \mathrm{tr}(V_pLV_p^T) \tag{20}$$

where, $L = D - A$ represents the Laplacian matrix of the graph, D is the diagonal matrix and tr()represents the matrix trace.

Finally, by minimizing the graph regularization across the two views, the objective function is obtained in Eq. (21):

$$\min_{V_p,H_m} L_{reg} = L_{reg_A} + L_{reg_X} = \mathrm{tr}(V_pLV_p^T) + \mathrm{tr}(H_mLH_m^T) \tag{21}$$

3.4 Proposed method

This study proposes a multi-sensor information fusion fault diagnosis method for hydroturbine bearings based on a multi-scale spatiotemporal graph neural network (MSTGNN). The overall framework is illustrated in Fig. 7, and the algorithmic workflow is presented in Table 1. The detailed steps are as follows:

Step 1: Multi-sensor bearing vibration signal acquisition: The collected multi-sensor vibration signals are segmented using an overlapping rectangular sliding window. The resulting dataset is divided into training and testing sets to ensure the generalization capability of the model.

Step 2: Multi-scale spatiotemporal feature interaction: First, adaptive receptive fields are employed to extract fault-related features from multi-sensor data. The resulting feature matrix is passed into a graph construction layer, where a spatiotemporal graph is built based on a cross-time-step connection mechanism. Subsequently, multi-scale Chebyshev graph convolutions (with orders K=1,2,3) are applied to the dynamic graph to capture spatiotemporal dependencies at varying receptive field sizes, thereby enhancing the model's feature representation capability.

Step 3: Graph structure reconstruction: The dynamic adjacency matrix is decomposed in a hierarchical manner. A non-negativity constraint is imposed to preserve physically interpretable fault propagation paths, thus improving the generalization and robustness of the model.

Step 4: Bearing fault diagnosis: The bearing datasets corresponding to different fault types are split into training and testing sets with a ratio of 3:1. During the training phase, model parameters are optimized, and performance is evaluated on the testing set.

This study proposes a multi-sensor information fusion fault diagnosis method for hydroturbine bearings based on a multi-scale spatiotemporal graph neural network (MSTGNN). The overall framework is illustrated in Fig. 7, and the algorithmic workflow is presented in Table 1. The detailed steps are as follows:

Step 1: Multi-sensor bearing vibration signal acquisition: The collected multi-sensor vibration signals are segmented using an overlapping rectangular sliding window. The resulting dataset is divided into training and testing sets to ensure the generalization capability of the model.

Step 2: Multi-scale spatiotemporal feature interaction: First, adaptive receptive fields are employed to extract fault-related features from multi-sensor data. The resulting feature matrix is passed into a graph construction layer, where a spatiotemporal graph is built based on a cross-time-step connection mechanism. Subsequently, multi-scale Chebyshev graph convolutions (with orders K=1,2,3) are applied to the dynamic graph to capture spatiotemporal dependencies at varying receptive field sizes, thereby enhancing the model's feature representation capability.

Step 3: Graph structure reconstruction: The dynamic adjacency matrix is decomposed in a hierarchical manner. A non-negativity constraint is imposed to preserve physically interpretable fault propagation paths, thus improving the generalization and robustness of the model.

Step 4: Bearing fault diagnosis: The bearing datasets corresponding to different fault types are split into training and testing sets with a ratio of 3:1. During the training phase, model parameters are optimized, and performance is evaluated on the testing set.

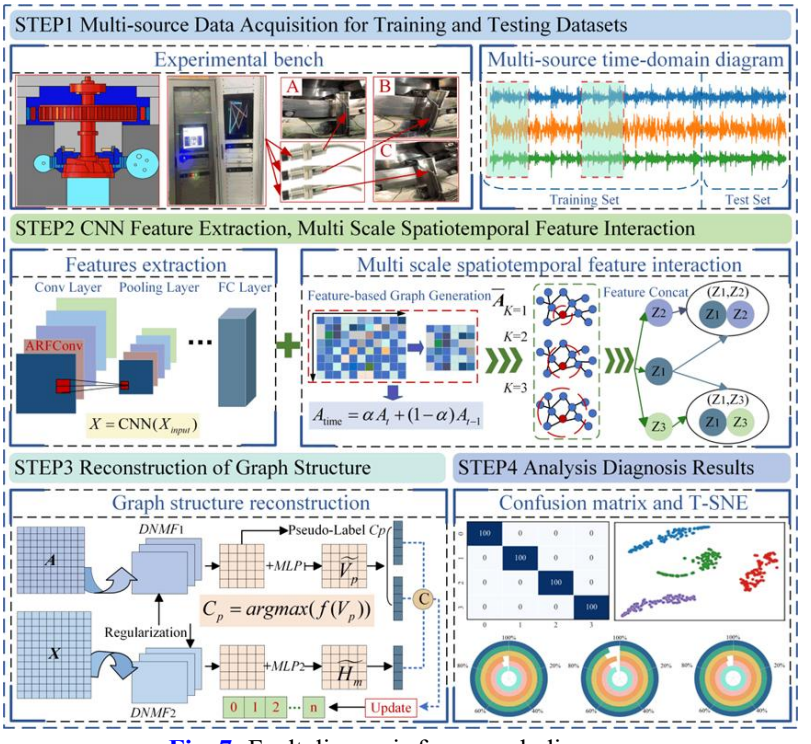

Fig. 7: Fault diagnosis framework diagram.

Table 1: Training process.

Algorithm1 MSTGNN Training Process
Input: Multi sensor signal sequence $S \in \mathbb{R}^{N\times T\times C}$
Hyper-Parameters: Graph sparsity k, Chebyshev order K, DNMF decomposition layers p.
Output: Fault diagnosis probability distribution$y$
1. Data Preprocessing: Sliding Window Segmentation, Z-score Normalization.
2. Adaptive Receptive Field Feature Extraction:
$\mathrm{ARFConv}\ (S) \rightarrow X$
$\mathrm{GenerateGraph}\ (X) \rightarrow G$
3. Multiscale Spatiotemporal Modeling:
$\mathrm{ChebConv(G,X,K)}\ \{k = 1 \ldots K\} \bigoplus \rightarrow H$
4. Interpretability reconstruction:
$\mathrm{DNMF}(A,\ p) \rightarrow A^{'}$
$L = \lambda_{\mathrm{rec}} L_{\mathrm{rec}} + \lambda_{\mathrm{conc}} L_{\mathrm{contrastive}} + \lambda_{\mathrm{negc}} L_{\mathrm{negative}} + \lambda_{\mathrm{reg}} L_{\mathrm{reg}}.$
5.End
$\mathrm{Softmax}(H) \rightarrow y$
Return y

4. Validation of experiments

This section aims to validate and analyze the application of the proposed MSTGNN method in bearing fault diagnosis, experiments will be conducted and results will be analyzed from different perspectives. All experiments were conducted on a computer with the following configuration: Pytorch 1.12.0+cu113, CPU: i5-12400F, GPU: NVIDIA GeForce RTX 2060 SUPER, 16 GB RAM. The batch size was set to 64, and the Adam optimizer was used to dynamically adjust the learning rate and update model weights, with an initial learning rate of 0.0001 and 200 training epochs. To ensure fairness in the experiments, all hyperparameters for the comparison methods mentioned in this paper were set the same as those of the proposed method.

4.1 Comparison methods

A comprehensive comparison was conducted between the proposed method and several state-of-the-art fault diagnosis approaches. The specific experimental conditions are as follows: all compared methods were trained on datasets of the same size, and all experiments were conducted under identical conditions. For data preprocessing, multi-sensor vibration signals were segmented using a sliding window and standardized via Z-score normalization. During model training, the batch size was set to 64, the Adam optimizer was used to dynamically adjust the learning rate and update model weights, the initial learning rate was set to 0.0001, and the number of training epochs was 200.

(1) Convolutional Neural Network model for capturing spatial and temporal features in signals:[43] This model uses fixed convolutional kernels to extract useful features from raw vibration signals, making fault diagnosis more efficient.

(2) FGDAE model utilizing fully connected graphs and graph-adaptive encoders:[44] This model aggregates multi-view feature information between channels by adaptively adjusting to varying operating conditions and flexibly adjusts the reconstruction loss weight.

(3) Periodic MFD model for multi-source cross-speed:[45] By constructing periodic samples, this model captures periodic features to resolve the inconsistency in the feature dimensions of multi-source samples.

(4) HAGCN model utilizing hierarchical attention graph convolutional networks:[46] This model designs regularized self-attention graph pooling to achieve effective information fusion from sensors.

(5) DAGCN model utilizing graph convolutional networks and maximum mean discrepancy:[47] This model mines the structural feature relationships of samples, constructs instance graphs, and performs modeling through graph convolutional networks.

4.2 HIT bearing dataset

4.2.1 Data introduction

The Harbin Institute of Technology (HIT) bearing dataset[48] used in this study is derived from real-world data collected from an actual aero-engine. The fault diagnosis testbed consists of a modified aero-engine—retaining the core dual-rotor structure while removing rotor blades, the combustion chamber, and certain accessory housings-along with a motor drive system, a lubrication system, two eddy current sensors, and four accelerometers (the detailed structure is illustrated in Fig. 8). The sensor specifications are K900XL displacement sensors (positions 1 and 2) for measuring horizontal and normal displacements, and KISTLER 8776A50M1 accelerometers (positions 3 and 4). Each group of sensor signals is sampled at a frequency of 25,000 Hz, with a continuous sampling duration of 15 seconds.

The bearing dataset comprises four health conditions, as illustrated in Fig. 9 and detailed in Table 2. The fault types were fabricated using wire electrical discharge machining (EDM), including normal bearings (N), outer race faults (OF), and inner race faults (IF), with fault lengths of 0.5 mm and 1.0 mm, respectively. The vibration signals under different health states from various sensors are visualized in Fig. 10.

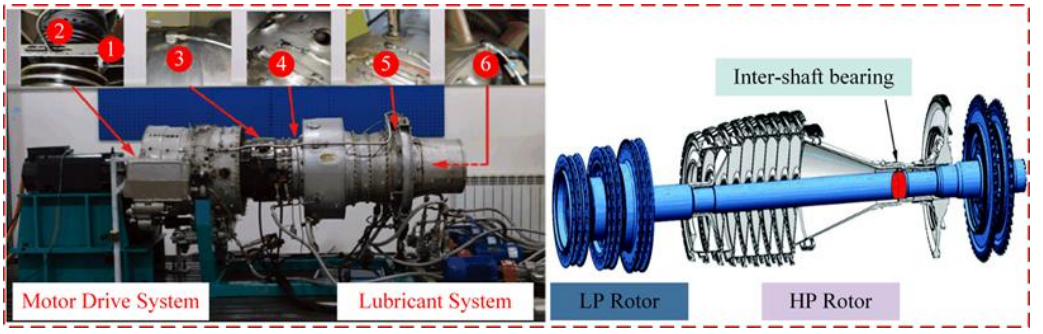

Fig. 8: Specific Structure of the HIT Fault Diagnosis Test Bench. Reproduced from.[48]

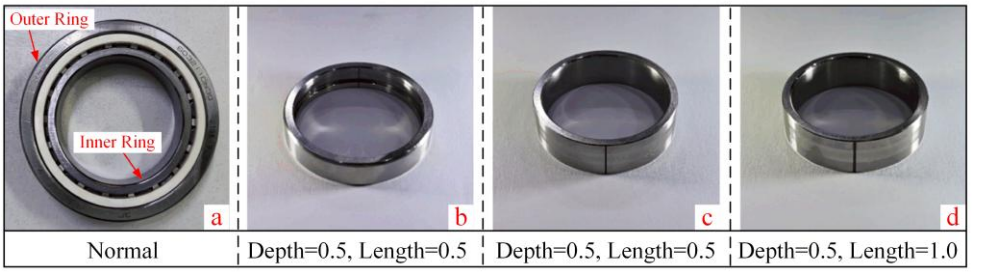

Fig. 9: HIT Bearing Health States. Reproduced from.[48]

Table 2: Introduction to HIT bearing dataset.

Health Condition	Label	Depth of fault/mm	Depth of fault/mm
Normal Bearing（a）	0	—	—
Outer Race Fault（b）	1	0.5	0.5
Inner Race Fault（c）	2	0.5	0.5
Inner Race Fault（d）	3	0.5	1.0

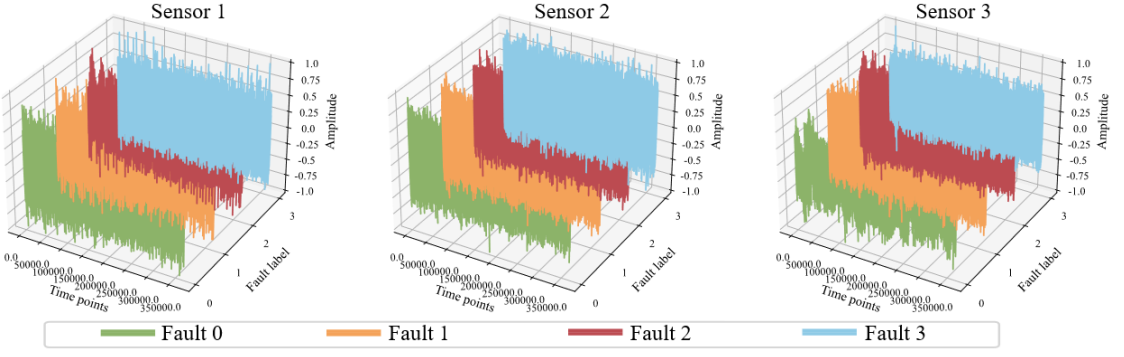

Fig. 10: Visualization of vibration signals from each sensor.

In engineering practice, fault diagnosis of mechanical bearing components aims to detect potential failures in a timely manner and prevent accidents. Therefore, accurate diagnosis of bearing faults is of critical importance. A sliding window approach is used to segment the fault data of bearing components under different health conditions, with each sample containing 1,024 data points. An overlapping sampling strategy is employed to enhance data diversity, and 360 samples are collected for each category. Among them, 70% are used for training and 30% for testing.

4.2.2 Results and analysis of the experiments

This paper employs the t-SNE algorithm to visualize the feature distributions learned by different models, as shown in Fig. 11(1)-11(6), illustrating the effectiveness of fault feature extraction for each model. As observed in Fig. 11(6), the proposed method demonstrates significantly better classification performance compared to other approaches. The same fault types exhibit more compact distributions across different data domains, while the separability between different fault categories is more pronounced. This indicates that the proposed model can efficiently extract critical fault features from multi-sensor vibration signals.

To further analyze the diagnostic performance of the multi-sensor MSTGNN method under different health conditions, this paper presents the confusion matrix after the model has

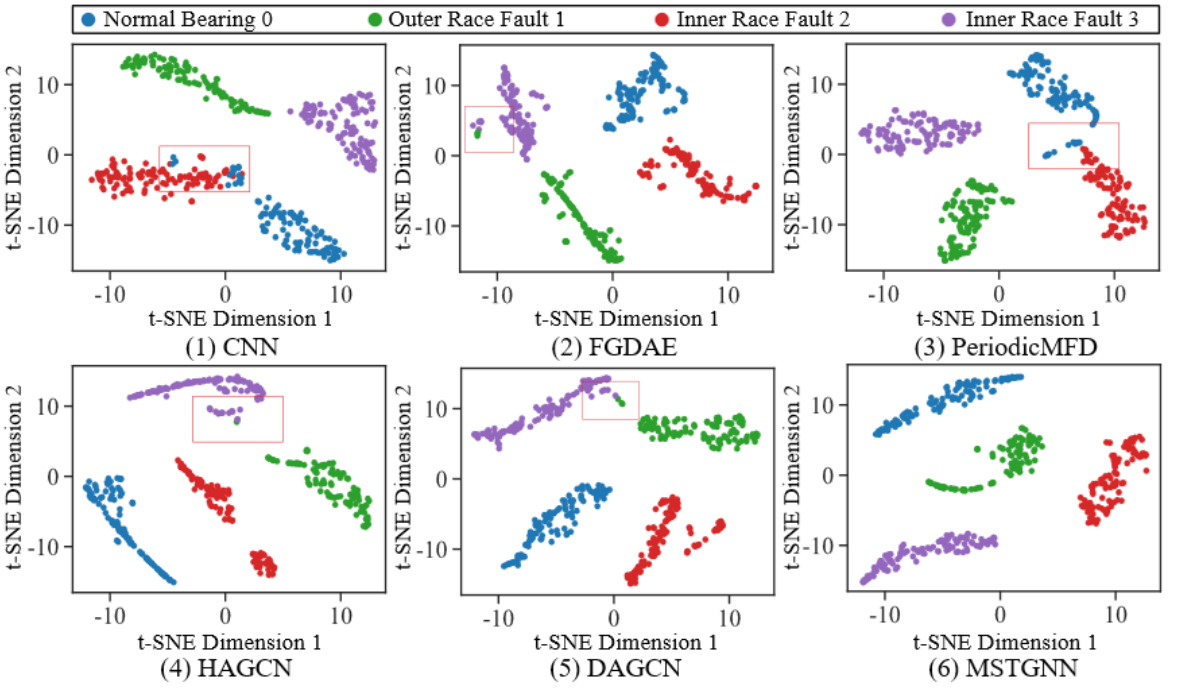

Fig. 11: Feature visualization of different methods on the HIT dataset.

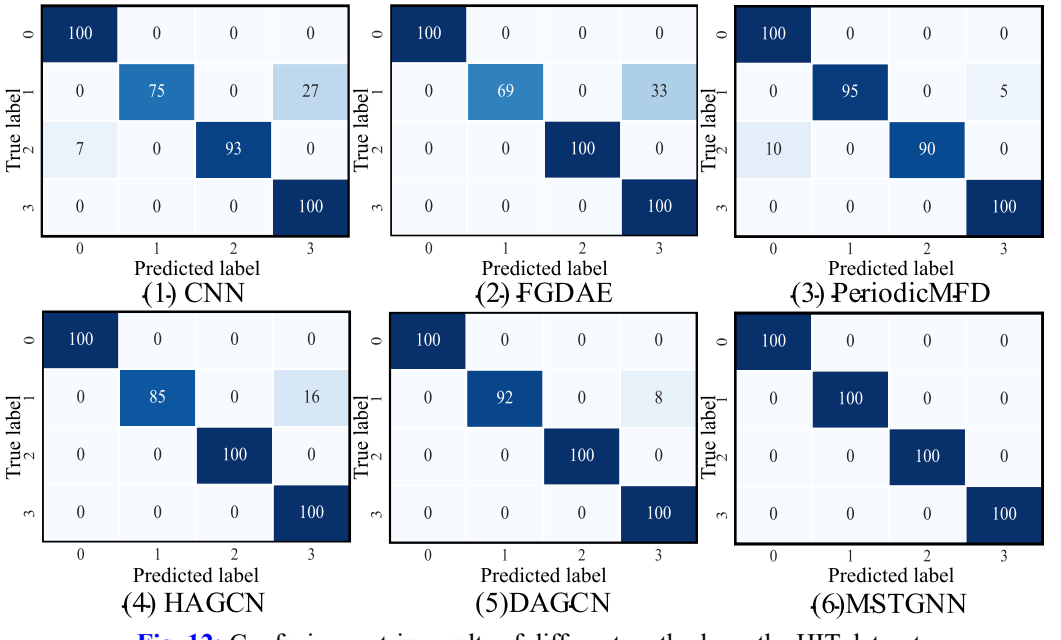

Fig. 12: Confusion matrix results of different methods on the HIT dataset.

stabilized, as shown in Fig. 12.

The values on the main diagonal of the confusion matrix represent the diagnostic accuracy for each health condition, while the values in the off-diagonal positions reflect the misclassification between different categories. It can be observed that there is a noticeable feature entanglement between label 1 and label 2, resulting in a higher misclassification rate. The proposed method effectively alleviates the misdiagnosis caused by feature entanglement, significantly improving the overall diagnostic performance.

To evaluate the stability and performance fluctuations of the proposed method during the training process, 10 repeated experiments were conducted. The performance data of each model at the final 50 training epochs were averaged, and the evaluation results are shown in Table 3, where the optimal value in each column is highlighted in bold.

To analyze the classification performance of each model on different fault categories more intuitively, this paper conducts 10 repeated experiments and averages the F1-Score metric from the last 50 epochs. The resulting donut chart is shown in Fig. 13. The gap in the chart represents the value of 100 - F1-Score, with a larger gap indicating worse classification performance. This visualization method more clearly explains the classification performance of the proposed method on each

Table 3: Evaluation metrics of different methods on the HIT dataset.

Method	F1	Pr(%)	Re(%)	Accuracy(%)
FGDAE	94.32	95.25	94.21	94.48±0.018
CNN	96.78	97.16	96.76	96.64±0.017
PeriodicMFD	96.22	96.41	96.22	96.17±0.008
DAGCN	96.57	96.97	96.94	96.69±0.022
HAGCN	97.31	97.62	97.31	97.67±0.016
MSTGNN	99.98	99.97	99.98	99.96±0.003

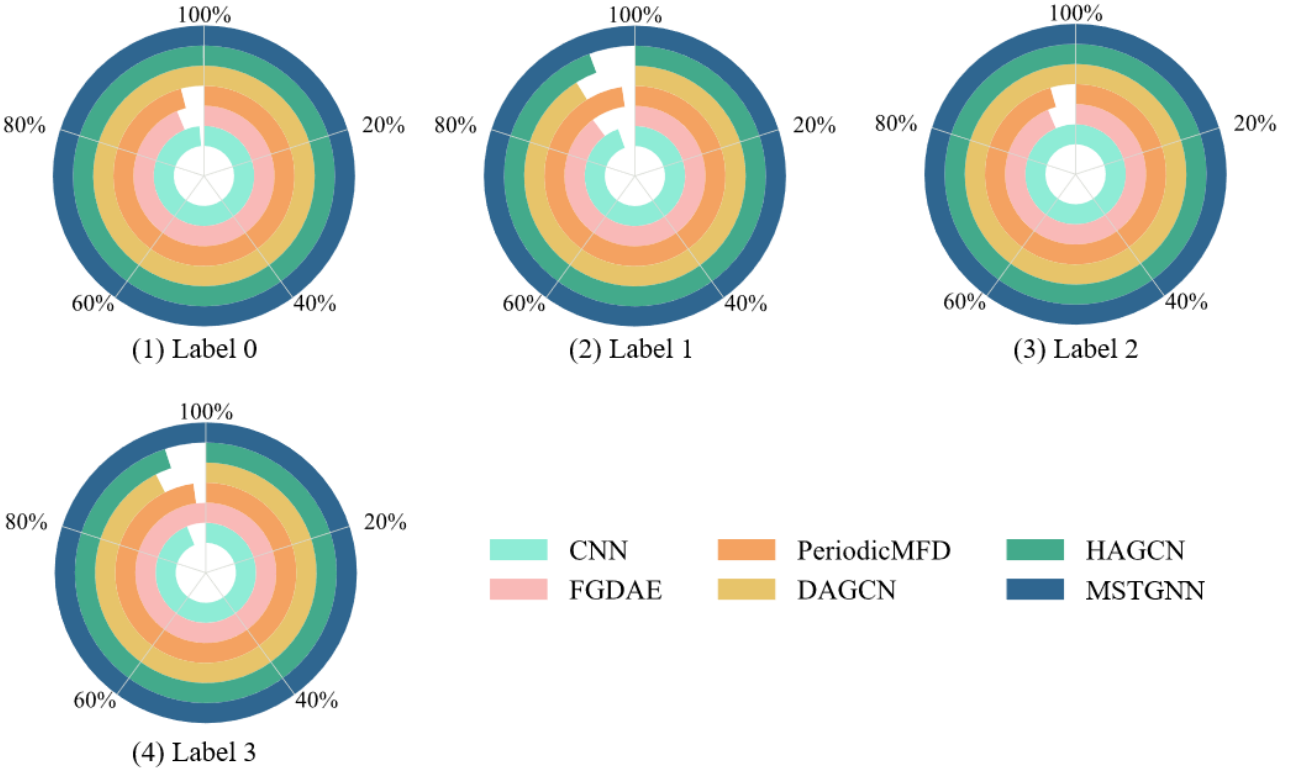

Fig. 13: F1-Score performance of different models on different labels.

fault label.

4.2.3 Ablation study

To better validate the diagnostic effectiveness of the proposed MSTGNN method, this section conducts an ablation study. In this study, the adaptive receptive field convolution layer module is named Module A, the multi-scale Chebyshev convolution module is named Module B, and the deep non-negative matrix factorization module is named Module C. When all three modules are combined, they form the MSTGNN method proposed in this paper. The experiments run for 200 epochs, repeated ten times, and the average data from the last 50 epochs are taken for validation. The diagnostic results for all modules are shown in Table 4.

The results in Table 4 indicate that the dynamic adaptive receptive field convolutional layer enhances local feature extraction, while the multi-scale Chebyshev convolution improves the model's sensitivity to fault features at different scales, collectively boosting generalization ability. The ablation study shows significant synergy among the modules: A+B complements local features with multi-scale structural information, A+C leverages latent feature decomposition to assist local feature learning, and B+C further strengthens the representation of complex fault patterns. When all modules are integrated to form MSTGNN, the model achieves its highest performance (Accuracy = 99.28%), demonstrating that the module interactions are complementary rather than redundant, effectively enhancing the model's robustness and diagnostic capability across multiple fault types.

4.3 Sichuan province hydropower station thrust bearing dataset

4.3.1 Data introduction

The experiment uses the real measured thrust bearing data from a hydropower station in Sichuan Province for validation. The thrust bearing data from the Sichuan hydropower station were collected using three eddy current sensors, measuring the axial displacements A, B, and C of the unit. The sensors are evenly arranged at 120° intervals in a clockwise direction, using the +X direction as the reference, and are mounted on the upper surface of the unit's mirror. The data were sampled at 10 kHz for 20 seconds, covering both normal operation and

Table 4: Accuracy and recall of each module.

Method	F1	Pr(%)	Re(%)	Accuracy(%)
A	96.82	96.51	96.11	96.65±0.018
A+B	97.14	97.37	97.86	97.33±0.008
A+C	97.94	97.51	97.39	97.28±0.012
B+C	98.78	98.73	98.98	98.48±0.007
MSTGNN	99.70	99.61	99.55	99.28±0.003

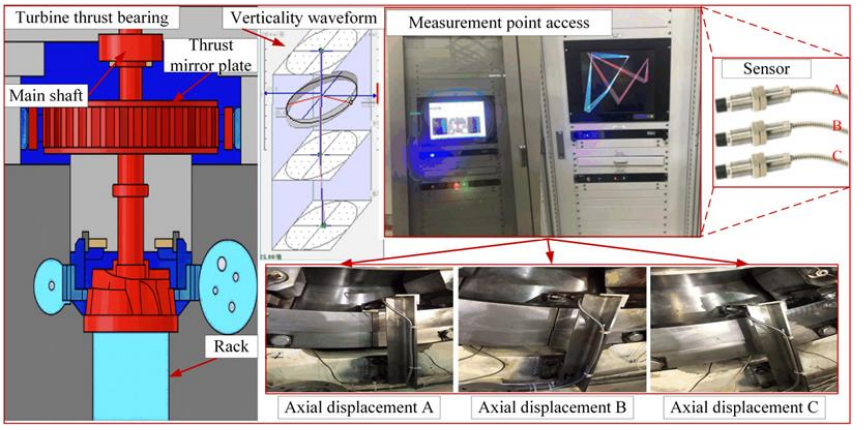

Fig. 14: Data collection flowchart of the thrust bearing in the hydropower station.

various load conditions of the unit, as shown in Fig. 14.

The collected thrust bearing data includes nine different health conditions. The health conditions and their corresponding labels are detailed in Table 5. The vibration signals under each sensor for different health conditions are visualized in Fig. 15.

To evaluate the fault diagnosis performance of the proposed method, 1024 data points from each fault type are selected as a sample, using an overlapping sampling method. A total of 300 samples are obtained from each health condition of the bearing, with 70% used for the training set and 30% for the test set.

4.3.2 Results and analysis of the experiments

Using the t-SNE method, the feature differences between different data formats are displayed to extract the high-dimensional features of each label, as shown in Fig. 16.

By comparing the high-dimensional features of different models, it is found that there exists some mutual coupling between features. The proposed method demonstrates a better clustering effect among similar features, indicating that it can accurately distinguish the data features of the same category. This explains the superior performance of the method in fault diagnosis and visually verifies the credibility of the diagnostic results.

Table 5: Introduction to the thrust bearing data of the hydropower station.

Health Condition	Label	Health Condition	Label
Normal Bearing	0	Corrosion Fault	5
Wear Fault	1	Cavitation Fault	6
Spalling Fault	2	Clearance Abnormality Fault	7
Crack Fault	3	Shaft Misalignment Fault	8
Scuffing Fault	4		

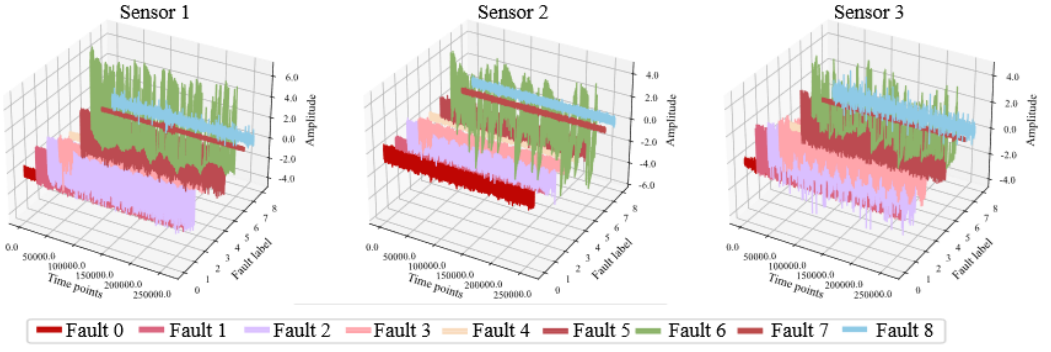

Fig. 15: Visualization of vibration signals from each sensor.

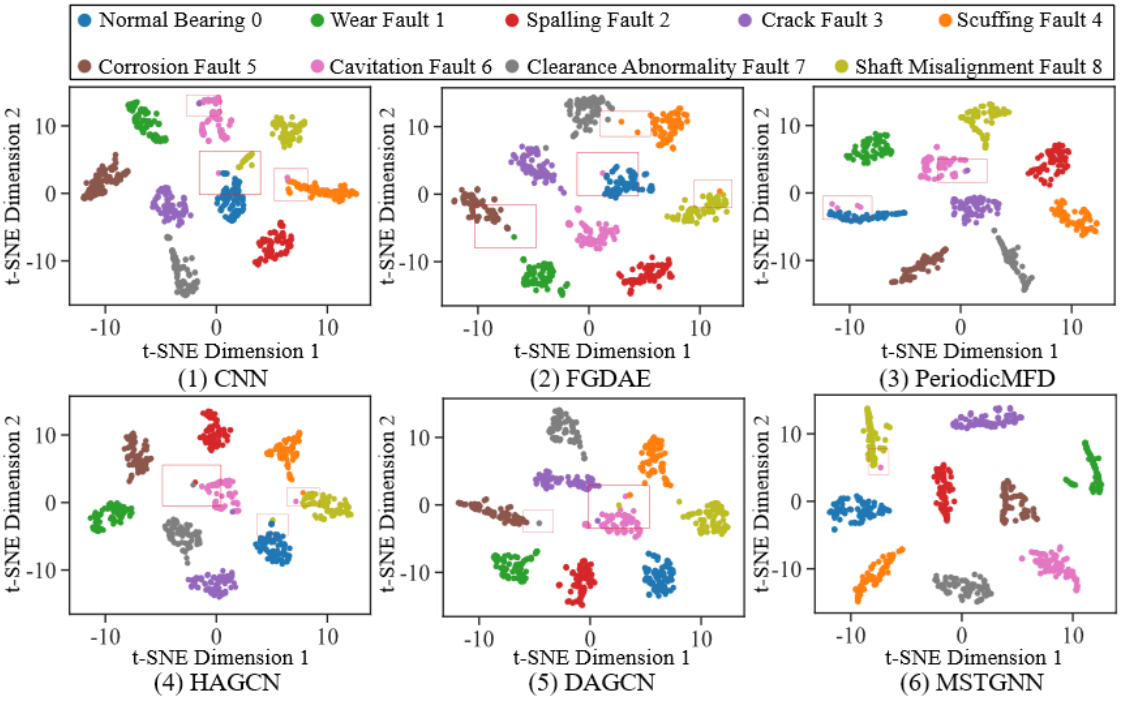

Fig. 16: Feature visualization of different methods on the thrust bearing dataset.

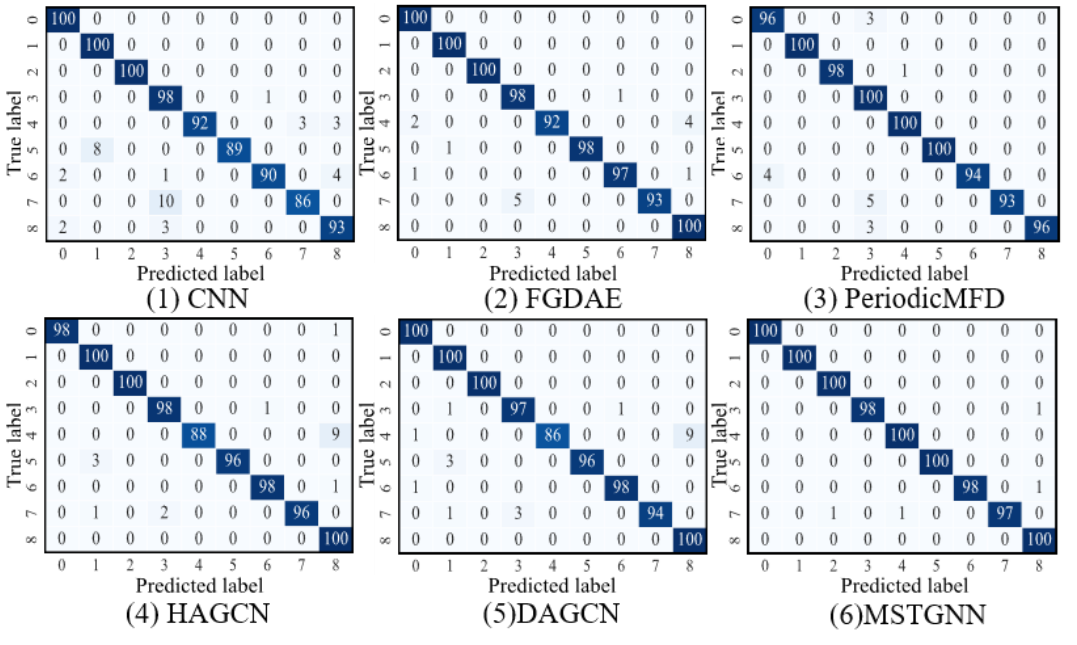

Fig. 17: Confusion matrix results of different methods on the thrust bearing dataset.

The confusion matrix results of each comparison model on the hydropower station thrust bearing dataset are shown in Fig. 17. When handling labels 5, 6, and 7, some models exhibit feature entanglement, leading to inaccurate fault type diagnosis. In contrast, the proposed method demonstrates excellent performance in fault recognition across all label types, further validating the model's ability to effectively distinguish fault features under varying dataset conditions.

To evaluate the stability and performance fluctuations of the proposed method during the training process, ten repeated experiments were conducted, and the performance data of each model at the final 50 training epochs were averaged. The evaluation results are shown in Table 6, where the optimal value in each column is highlighted in bold.

As shown in the pie chart in Fig. 18, the proposed model exhibits good F1 scores across different categories. However, in the corrosion fault condition of the hydropower station thrust bearing, the classification performance of all models decreases. The main reason for this phenomenon is that the feature frequencies of faults such as wear and spalling are more distinct in the thrust bearing dataset, making them easier for the model to recognize. In contrast, the features of corrosion faults are relatively hidden, making them more difficult to detect accurately.

Table 6: Evaluation metric values of different methods.

Method	F1	Pr(%)	Re(%)	Accuracy(%)
CNN	95.41	95.89	95.88	95.82±0.019
HAGCN	97.27	97.38	97.47	97.18±0.006
DAGCN	97.09	97.28	97.09	97.54±0.009
FGDAE	97.85	97.76	97.20	97.69±0.007
PeriodicMFD	98.27	98.38	98.26	98.14±0.006
MSTGNN	99.25	99.27	99.24	99.27±0.004

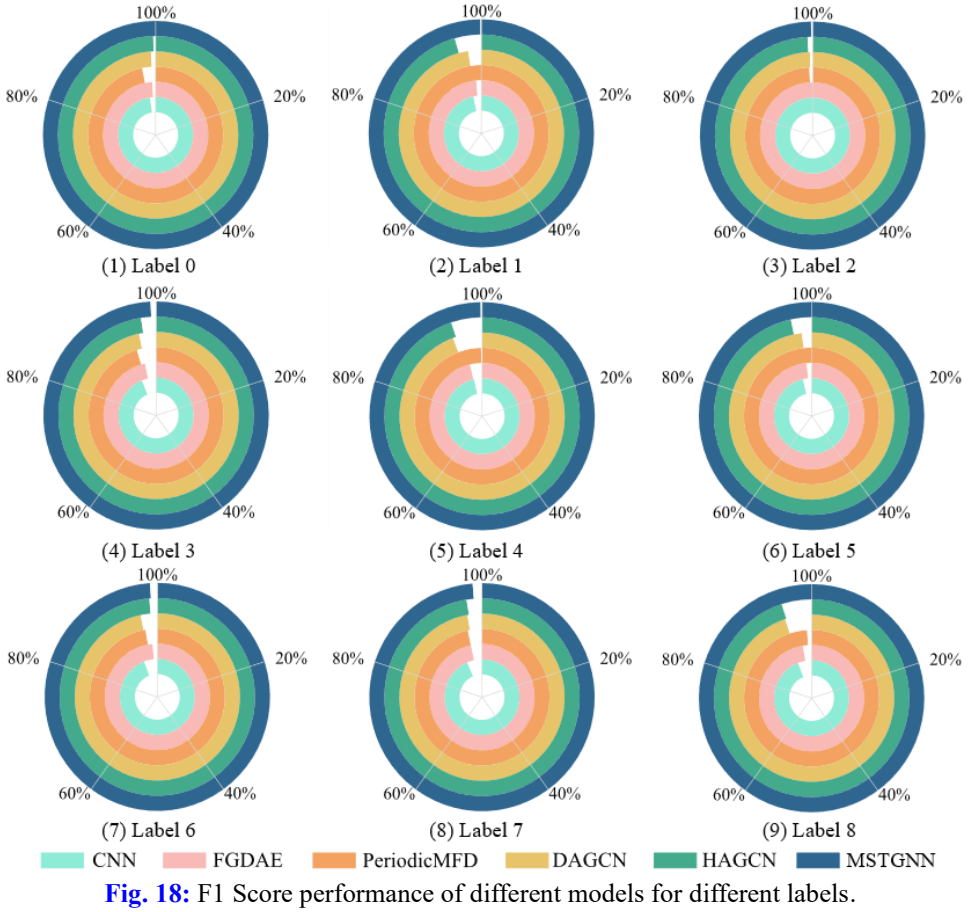

Fig. 18: F1 Score performance of different models for different labels.

Table 7: Accuracy and recall of each module.

Method	F1	Pr(%)	Re(%)	Accuracy(%)
A	96.25	96.58	96.83	96.88±0.020
A+B	97.72	97.72	97.17	97.56±0.006
A+C	97.84	97.84	97.27	97.12±0.013
B+C	98.14	98.97	98.19	98.24±0.006
MSTGNN	99.17	99.11	99.43	99.12±0.003

4.3.3 Ablation study

In this section, the impact of different component modules on the MSTGNN performance is evaluated through an ablation study, where each module is integrated for testing. The experiment uses the same training parameters, runs for 200 epochs, and repeats the test ten times, taking the average of the last 50 epochs for stable data. The results are shown in Table 7, with the conclusions as follows.

The results in Table 7 indicate that each module contributes to the performance of MSTGNN. Module A enhances local feature extraction, achieving an F1-score of 96.25% and an accuracy of 96.88% when used alone. With the addition of Module B (A+B), the F1-score and accuracy increase to 97.72% and 97.56%, respectively, demonstrating that multi-scale graph convolution effectively captures spatiotemporal dependencies and improves model robustness. Module C (A+C) further optimizes feature extraction by uncovering deeper structural information within the network. The combination of B+C further strengthens the representation of complex fault patterns. When all modules are integrated to

form MSTGNN, the F1-score reaches 99.17% and the accuracy 99.12%, indicating that the modules work synergistically and complementarily, significantly enhancing the model's robustness and diagnostic accuracy across multiple fault types.

5. Conclusion

To address the limitations of current multi-sensor fault diagnosis methods with low accuracy, this paper proposes a multi-sensor spatiotemporal correlation-based Graph Neural Network (GNN) method for hydropower turbine bearing fault diagnosis. This method combines adaptive receptive field convolution and multi-scale Chebyshev convolution to effectively extract both local and global spatiotemporal features, while integrating contrastive learning and deep non-negative matrix factorization to further enhance the discriminative power of fault features. Experimental results show that, on the HIT bearing dataset and the thrust bearing dataset from a hydropower station in Sichuan Province, this method achieves a diagnostic accuracy of 99%, with a recall rate improvement of over 3.36% compared to baseline methods. These results fully validate the excellent performance and good adaptability of MSTGNN across different bearing datasets.

Building upon the achievement of high-accuracy fault diagnosis, future research will focus on the following aspects:

Exploring methods for cross-domain and multi-source data fusion to enhance the model's robustness and adaptability across different sensor types and multimodal data.

Future collaboration with hydropower enterprises will involve collecting thrust bearing data from various types of sensors to further validate the proposed method and improve the model's interpretability and reliability.

Acknowledgments

This work is supported in part by the Key Program of National Natural Science Foundation of China (U23A20669), in part by the Supported by the Open Research Subject of Key Laboratory of Fluid and Power Machinery (Xihua University), Ministry of Education (grant number szjj2014-037), in part by the National Natural Science Foundation of China (52505091).

Conflict of Interest

There is no conflict of interest.

Supporting Information

Not applicable.

CRediT Statement

Yongzhong Zeng: Led the research work, proposed the research concept and technical route, designed the experimental scheme, carried out experimental validation and model optimization, drafted the initial manuscript, coordinated research resources, and was responsible for project funding application and management. **Liuyan Tang:** Responsible for the collection and preprocessing of vibration signal data, and assisted in the design and implementation of the experimental setup. **Shuquan Xiao:** Participated in the acquisition and processing of vibration signal data, and was responsible for part of the visualization and chart preparation of the results. **Xueyi Li**, **Tianyang Wang** and **Fulei Chu:** Provided overall supervision of the research, coordinated project progress and task allocation, and participated in the review and scientific validation of the manuscript. **Xiaobing Liu:** Participated in project management and technical support, coordinated the integration of various modules in the research process, and was responsible for manuscript language editing and formatting. **Zhilin Dong** and **Yuanjiang Ma:** Responsible for the experimental design, on-site data collection, and preprocessing of thrust bearing data from real hydropower stations. All authors have reviewed and approved the final version of the manuscript, ensuring the accuracy and completeness of the research.

References

[1] H. Li, T. Wang, F. Zhang, F. Chu, Enhanced adaptive high-frequency resonance technology for bearing fault detection, *Structural Health Monitoring*, 2024, **23**, 3430-3445, doi: 10.1177/14759217231224034.

[2] F. Chen, Z. Zhao, X. Hu, D. Liu, Z. Kang, Z. Ma, P. Xiao, X. Yin, J. Yang, Enhancing the safety of hydroelectric power generation systems: an intelligent identification of axis orbits based on a nonlinear dynamics method, *Energy*, 2025, **324**, 135864, doi: 10.1016/j.energy.2025.135864.

[3] X. Li, T. Yu, F. Zhang, J. Huang, D. He, F. Chu, Mixed style network based: a novel rotating machinery fault diagnosis method through batch spectral penalization, *Reliability Engineering & System Safety*, 2025, **255**, 110667, doi: 10.1016/j.ress.2024.110667.

[4] D. Zhao, D. Shao, T. Wang, L. Cui, Time-frequency self-similarity enhancement network and its application in wind turbines fault analysis, *Advanced Engineering Informatics*, 2025, **65**, 103322, doi: 10.1016/j.aei.2025.103322. [LinkOut]

[5] X. Li, T. Yu, X. Wang, D. Li, Z. Xie, X. Kong, Fusing joint distribution and adversarial networks: a new transfer learning method for intelligent fault diagnosis, *Applied Acoustics*, 2024, **216**, 109767, doi: 10.1016/j.apacoust.2023.109767.

[6] S. Zhi, H. Wu, H. Shen, T. Wang, H. Fu, Entropy-aided meshing-order modulation analysis for wind turbine planetary gear weak fault detection under variable rotational speed, *Entropy*, 2024, **26**, 409, doi: 10.3390/e26050409.

[7] X. Li, X. Wu, J. Yu, T. Wang, C. Shi, J.-H. Choi, F. Chu, A fault diagnosis data augmentation method integrating multimodal non-Gaussian denoising diffusion generative adversarial network, *Advanced Engineering Informatics*, 2025, **68**, 103776, doi: 10.1016/j.aei.2025.103776.

[8] Q. Li, H. Li, W. Hu, S. Sun, Z. Qin, F. Chu, Transparent operator network: a fully interpretable network incorporating learnable wavelet operator for intelligent fault diagnosis, *IEEE*

Transactions on Industrial Informatics, 2024, **20**, 8628-8638, doi: 10.1109/TII.2024.3366993.
[9] X. Li, X. Wu, T. Wang, Y. Xie, F. Chu, Fault diagnosis method for imbalanced data based on adaptive diffusion models and generative adversarial networks, *Engineering Applications of Artificial Intelligence*, 2025, **147**, 110410, doi: 10.1016/j.engappai.2025.110410.
[10] S. Zhi, Y. Niu, L. Ma, H. Wu, H. Shen, T. Wang, Local entropy selection scaling-extracting chirplet transform for enhanced time-frequency analysis and precise state estimation in reliability-focused fault diagnosis of non-stationary signals, *Eksploatacja i Niezawodność – Maintenance and Reliability*, 2025, 205977, doi: 10.17531/ein/205977.
[11] J. Wang, H. Shao, J. He, L. Liu, J. Ma, B. Liu, A novel interpretable fault diagnosis method using multi-image feature extraction and attention fusion, *Pattern Recognition Letters*, 2025, **189**, 38-47, doi: 10.1016/j.patrec.2025.01.006.
[12] J. Zhao, D. He, Z. Jin, X. Zhang, J. Zhou, A new method for bearing remaining useful life prediction based on dynamic wavelet and physical information constraints, *Expert Systems with Applications*, 2026, **296**, 129023, doi: 10.1016/j.eswa.2025.129023.
[13] H. Shao, Y. Xiao, J. Leng, X. Zhao, B. Liu, Collaborative human-computer fault diagnosis *via* calibrated confidence estimation, *Advanced Engineering Informatics*, 2025, **65**, 103349, doi: 10.1016/j.aei.2025.103349.
[14] J. Guo, Q. He, D. Zhen, F. Gu, A. D. Ball, Multi-sensor data fusion for rotating machinery fault detection using improved cyclic spectral covariance matrix and motor current signal analysis, *Reliability Engineering & System Safety*, 2023, **230**, 108969, doi: 10.1016/j.ress.2022.108969.
[15] J. Wang, H. Shao, S. Yan, B. Liu, C-ECAFormer: a new lightweight fault diagnosis framework towards heavy noise and small samples, *Engineering Applications of Artificial Intelligence*, 2023, **126**, 107031, doi: 10.1016/j.engappai.2023.107031.
[16] X. Yan, D. Jiang, L. Xiang, Y. Xu, Y. Wang, CDTFAFN: a novel coarse-to-fine dual-scale time-frequency attention fusion network for machinery vibro-acoustic fault diagnosis, *Information Fusion*, 2024, **112**, 102554, doi: 10.1016/j.inffus.2024.102554.
[17] X. Li, G. Ran, S. Xiao, K. Su, Q. Li, T. Wang, Z. Qin, F. Chu, A fault diagnosis method for bearings based on multiscale graph convolutional network under non-stationary speed conditions, *Mechanism and Machine Theory*, 2025, **217**, 106267, doi: 10.1016/j.mechmachtheory.2025.106267.
[18] H. Zhou, Q. Huang, C. Zhou, P. He, N. Zhe, H. Wang, Rotating machinery fault diagnosis method based on temporal–spatial vibration feature fusion extraction, *IEEE Sensors Journal*, 2025, **25**, 1184-1197, doi: 10.1109/JSEN.2024.3496776.
[19] J. Sun, X. Gu, J. He, S. Yang, Y. Tu, C. Wu, A robust approach of multi-sensor fusion for fault diagnosis using convolution neural network, *Journal of Dynamics*, *Monitoring and Diagnostics*, 2022, 103-110, doi: 10.37965/jdmd.2022.95.
[20] H. Shao, Y. Ming, Y. Liu, B. Liu, Small sample gearbox fault diagnosis based on improved deep forest in noisy environments, *Nondestructive Testing and Evaluation*, 2025, **40**, 3935-3956, doi: 10.1080/10589759.2024.2404489.
[21] X. Yan, Z. Sun, J. Zhao, Z. Shi, C.-A. Zhang, Fault diagnosis of rotating machinery equipped with multiple sensors using space-time fragments, *Journal of Sound and Vibration*, 2019, **456**, 49-64, doi: 10.1016/j.jsv.2019.05.036.
[22] H. Wang, S. Li, L. Song, L. Cui, A novel convolutional neural network based fault recognition method *via* image fusion of multi-vibration-signals, *Computers in Industry*, 2019, **105**, 182-190, doi: 10.1016/j.compind.2018.12.013.
[23] Y. Zhang, J. Ding, Y. Li, Z. Ren, K. Feng, Multi-modal data cross-domain fusion network for gearbox fault diagnosis under variable operating conditions, *Engineering Applications of Artificial Intelligence*, 2024, **133**, 108236, doi: 10.1016/j.engappai.2024.108236.
[24] Zhao D L, Ma W, Liang W K, *et al.* Data fusion fault diagnosis and simulation of hydroelectric units vibration Zhongguo Dianji Gongcheng Xuebao, *Proceedings of the Chinese Society of Electrical Engineering*, 2005, **25**(20), 137-142.
[25] F. Guo, J. Guo, Intelligent detection algorithm of mechanical vibration fault based on multi-sensor fusion, *International Conference on Artificial Intelligence, Deep Learning and Neural Networks*, Guangzhou, China, IEEE, September 20-22, 2024, 71-77, doi: 10.1109/AIDLNN65358.2024.00019.
[26] D. He, J. Zhao, Z. Jin, C. Huang, C. Yi, J. Wu, DCAGGCN: a novel method for remaining useful life prediction of bearings, *Reliability Engineering & System Safety*, 2025, **260**, 110978, doi: 10.1016/j.ress.2025.110978.
[27] R. Liu, Q. Zhang, D. Lin, W. Zhang, S. X. Ding, Causal intervention graph neural network for fault diagnosis of complex industrial processes, *Reliability Engineering & System Safety*, 2024, **251**, 110328, doi: 10.1016/j.ress.2024.110328.
[28] X. Li, Y. Wang, J. Yao, M. Li, Z. Gao, Multi-sensor fusion fault diagnosis method of wind turbine bearing based on adaptive convergent viewable neural networks, *Reliability Engineering & System Safety*, 2024, **245**, 109980, doi: 10.1016/j.ress.2024.109980.
[29] J. Liu, Z. He, Y. Miao, Causality-based adversarial attacks for robust GNN modelling with application in fault detection, *Reliability Engineering & System Safety*, 2024, **252**, 110464, doi: 10.1016/j.ress.2024.110464.
[30] Z. J. Gale-Day, L. Shub, K. V. Chuang, M. J. Keiser, Proximity graph networks: predicting ligand affinity with message passing neural networks, *Journal of Chemical Information and Modeling*, 2024, **64**, 5439-5450, doi: 10.1021/acs.jcim.4c00311.
[31] X. Yan, Z. Shi, Z. Sun, C.-A. Zhang, Multisensor fusion on hypergraph for fault diagnosis, *IEEE Transactions on Industrial Informatics*, 2024, **20**, 10008-10018, doi: 10.1109/TII.2024.3393137.
[32] J. Xu, H. Ke, Z. Jiang, S. Mo, Z. Chen, W. Gui, OHCA-GCN: a novel graph convolutional network-based fault diagnosis method for complex systems *via* supervised graph construction and optimization, *Advanced Engineering Informatics*, 2024, **61**, 102548, doi: 10.1016/j.aei.2024.102548.

[33] X. Li, S. Xiao, Q. Li, L. Zhu, T. Wang, F. Chu, The bearing multi-sensor fault diagnosis method based on a multi-branch parallel perception network and feature fusion strategy, *Reliability Engineering & System Safety*, 2025, **261**, 111122, doi: 10.1016/j.ress.2025.111122.
[34] S. Zhi, K. Su, J. Yu, X. Li, H. Shen, An unsupervised transfer learning bearing fault diagnosis method based on multi-channel calibrated Transformer with shiftable window, *Structural Health Monitoring*, 2025, 14759217251324671, doi: 10.1177/14759217251324671.
[35] D. Zhao, D. Shao, L. Cui, CTNet: a data-driven time-frequency technique for wind turbines fault diagnosis under time-varying speeds, *ISA Transactions*, 2024, **154**, 335-351, doi: 10.1016/j.isatra.2024.08.029.
[36] Y. Zhao, H. Liao, S. Pan, Y. Zhao, Interpretable multi-graph convolution network integrating spatio-temporal attention and dynamic combination for wind power forecasting, *Expert Systems with Applications*, 2024, **255**, 124766, doi: 10.1016/j.eswa.2024.124766.
[37] J. Hu, W. Li, Y. Zhang, Z. Tian, Cross-domain few-shot fault diagnosis based on meta-learning and domain adversarial graph convolutional network, *Engineering Applications of Artificial Intelligence*, 2024, **136**, 108970, doi: 10.1016/j.engappai.2024.108970.
[38] Q. Xia, Y. Yu, Z. Chang, B. Hui, H. Luo, CPT-SNN: a spiking neural network that can combine the previous timestep, *Neurocomputing*, 2025, **640**, 130253, doi: 10.1016/j.neucom.2025.130253.
[39] Defferrard M, Bresson X, Vandergheynst P. Convolutional neural networks on graphs with fast localized spectral filtering, *Advances in neural information processing systems*, 2016, **29**, doi: 10.48550/arXiv.1606.09375.
[40] Chen T, Kornblith S, Norouzi M, *et al.* A simple framework for contrastive learning of visual representations, *International conference on machine learning PmLR*, 2020, 1597-1607, doi: 10.48550/arXiv.2002.05709.
[41] D. Wang, T. Li, P. Deng, F. Zhang, W. Huang, P. Zhang, J. Liu, A generalized deep learning clustering algorithm based on non-negative matrix factorization, *ACM Transactions on Knowledge Discovery from Data*, 2023, **17**, 1-20, doi: 10.1145/3584862.
[42] M. Belkin, P. Niyogi, Laplacian eigenmaps for dimensionality reduction and data representation, *Neural Computation*, 2003, **15**, 1373-1396, doi: 10.1162/089976603321780317.
[43] S. Y. Han, H. D. Shao, H. K. Jiang, X. Y. Zhang, Intelligent fault diagnosis of aero-engine high-speed bearings using enhanced CNN, *Acta Aeronautica et Astronautica Sinica*, 2022, **43**(9), 625479.
[44] S. Yan, H. Shao, Z. Min, J. Peng, B. Cai, B. Liu, FGDAE: a new machinery anomaly detection method towards complex operating conditions, *Reliability Engineering & System Safety*, 2023, **236**, 109319, doi: 10.1016/j.ress.2023.109319.
[45] J. Zheng, C. Yang, T. Zhang, B. Jiang, X. Fan, X.-M. Wu, H. Shao, PeriodicMFD: a periodic-based framework for multisource fault diagnosis, *IEEE Transactions on Transportation Electrification*, 2025, **11**, 7252-7260, doi: 10.1109/TTE.2024.3525077.
[46] T. Li, Z. Zhao, C. Sun, R. Yan, X. Chen, Hierarchical attention graph convolutional network to fuse multi-sensor signals for remaining useful life prediction, *Reliability Engineering & System Safety*, 2021, **215**, 107878, doi: 10.1016/j.ress.2021.107878.
[47] T. Li, Z. Zhao, C. Sun, R. Yan, X. Chen, Domain adversarial graph convolutional network for fault diagnosis under variable working conditions, *IEEE Transactions on Instrumentation and Measurement*, 2021, **70**, 3515010, doi: 10.1109/TIM.2021.3075016.
[48] L. Hou, H. Yi, Y. Jin, M. Gui, L. Sui, J. Zhang, Y. Chen, Inter-shaft bearing fault diagnosis based on aero-engine system: a benchmarking dataset study, *Journal of Dynamics, Monitoring and Diagnostics*, 2023, 228-242, doi: 10.37965/jdmd.2023.314.