

# Adaptive Dynamic K-NN Spatio-Temporal Graph Convolutional Network for Bearing Fault Diagnosis

Xiuwei Hu\*, Ting Fang, Wenhui Chang, Cunrui Pang, Tonghui Wang

Zhangjiakou Cigarette Factory Co., Ltd., Zhangjiakou, China  
Email: \*hwx\_sdut@163.com

**How to cite this paper:** Hu, X.W., Fang, T., Chang, W.H., Pang, C.R. and Wang, T.H. (2026) Adaptive Dynamic K-NN Spatio-Temporal Graph Convolutional Network for Bearing Fault Diagnosis. *Journal of Computer and Communications*, 14, 102-114.  
<https://doi.org/10.4236/jcc.2026.146008>

**Received:** May 29, 2026

**Accepted:** June 22, 2026

**Published:** June 25, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).  
<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

Rolling bearing fault diagnosis is an important means to ensure the reliable operation of mechanical equipment. Aiming at the problems of poor working condition adaptability, over-smoothing in deep networks, and insufficient mining of fault temporal evolution features existing when the traditional K-nearest neighbor graph convolutional network is applied to bearing fault diagnosis, this paper proposes an adaptive dynamic K-NN spatio-temporal graph convolutional network (AdaDKNN-STGCN). Experimental validations are implemented on the CWRU bearing dataset. The proposed model attains a test accuracy of 99.78% in ten-class classification. It exhibits excellent generalization capability and anti-interference performance under noisy and variable load environments, which provides a reliable solution for intelligent fault diagnosis of rotating machinery bearings.

## Keywords

Rolling Bearing, Fault Diagnosis, Graph Convolutional Network, Adaptive Dynamic KNN, Spatio-Temporal Feature Fusion

## 1. Introduction

Rolling bearings are key vulnerable parts of rotating machinery transmission systems, and their operating conditions largely determine the safety, stability, and service life of the entire mechanical equipment [1]. In industrial applications, equipment shutdowns caused by bearing failures occur frequently. Accordingly, exploring high-precision and robust intelligent diagnosis methods for rolling bearing faults possesses important practical value for equipment health management and predictive maintenance.

Traditional bearing fault diagnosis methods rely on manual experience to extract time-domain indicators including kurtosis, root mean square, and margin index, as well as frequency-domain spectral features [2]. Fault classification is then realized by combining the support vector machine (SVM) and the back propagation (BP) neural network. Such methods are highly dependent on handcrafted features and suffer from poor generalization ability, making them difficult to adapt to complex industrial noise environments [3]. With the development of deep learning, end-to-end fault diagnosis methods based on convolutional neural networks (CNN) and recurrent neural networks (LSTM/GRU) have become the mainstream research direction [4]. CNN performs well in local feature extraction but fails to model the long-range dependencies of signal sequences [5]. Although LSTM is capable of mining temporal information, it ignores the topological correlation among samples and shows limited performance in characterizing weak faults with vibration impulses.

Graph Neural Networks (GNNs) are capable of processing non-Euclidean data. They can convert one-dimensional vibration sequences into graph-structured data to mine correlation information between samples, and have been widely applied to bearing fault diagnosis in recent years [6]. Most existing graph network-based diagnosis methods construct static graph structures following the fixed K-Nearest Neighbor (K-NN) rule, where a fixed number of adjacent nodes is assigned to all sample nodes. Feature learning is then implemented by using basic Graph Convolutional Networks (GCN) and Graph Attention Networks (GAT) [7]-[9]. Although the fixed KNN-GCN method achieves higher diagnosis accuracy than traditional deep learning models, it still suffers from a rigid graph structure, separated spatio-temporal information, and deep-layer over-smoothing.

To address the aforementioned research limitations, this paper proposes AdaD-KNN-STGCN. The main contributions are summarized as follows:

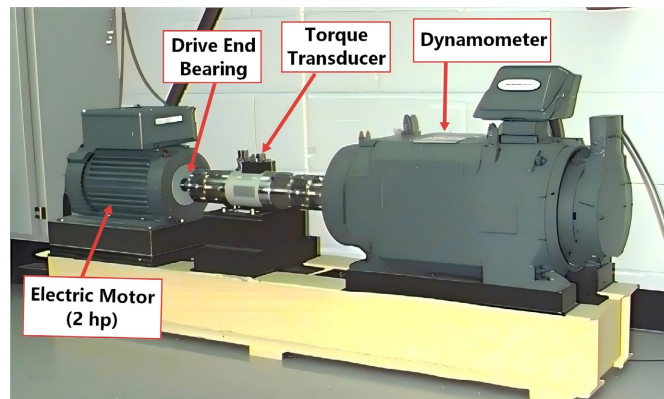
- 1) A dynamic weighted K-NN graph construction algorithm is proposed. It integrates feature similarity and temporal distance to build an adaptive adjacency matrix. Breaking away from the traditional fixed-K strategy, the method realizes adaptive updates of graph topology under different working conditions.
- 2) A spatio-temporal dual-branch graph convolution architecture is constructed. The spatial branch explores topological correlations among nodes, while the temporal branch extracts temporal evolution features, realizing the fusion of complete spatio-temporal information of vibration signals.
- 3) Residual skip connections and a node attention mechanism are introduced, which effectively alleviate the over-smoothing problem of graph convolution and enhance the extraction capability of weak fault features.

## 2. Related Theoretical Foundations

### 2.1. Vibration Signal Characteristics of Rolling Bearings

Rolling bearing faults are mainly divided into four categories: normal state, inner race fault, outer race fault, and rolling element fault. Different damage sizes cor-

respond to various fault severity levels [10]. Bearing vibration signals are non-stationary impact time-series signals, and fault impacts manifest as amplitude fluctuations in the time domain and resonance peak shifts in the frequency domain [11]. **Figure 1** presents the time-frequency domain analysis of vibration signals collected from bearings with ball faults. The characteristics of these signals serve as the basis for the time-frequency fusion features extracted by the model, and also provide physical support for subsequent fault diagnosis based on spatio-temporal graph convolution. Therefore, combining time-domain statistical features and frequency-domain spectral features can fully characterize fault information.



**Figure 1.** Time-domain and frequency-domain analysis of typical fault signals.

## 2.2. Fundamentals of Graph Convolutional Networks

As a dominant graph neural network for non-Euclidean graph-structured data, GCN is capable of capturing node attributes and topological correlations, which remedies the drawback of conventional CNNs limited to regular data [12] [13].

A graph  $G$  is formally defined as  $G = (V, E)$ , where  $V$  represents the set of nodes and  $E$  denotes the set of edges. The core idea of graph convolution is to aggregate neighborhood features and perform nonlinear transformations, so as to fuse each node's own features and its topological neighborhood features, and finally generate node representations containing both attribute and structural information.

## 2.3. Traditional KNN Graph Construction Method

As a typical approach to constructing graphs from unstructured data, KNN graph building screens neighbors based on node feature similarity and creates topological connections [14]-[17]. It converts raw feature data into graph data and is extensively used in topology-free deep learning tasks.

Firstly, we compute the feature similarity across all nodes using Euclidean or cosine distance. Next, each node picks its top-K most similar neighbors [10]. Eventually, undirected edges are created to build the entire graph topology. The calculation formula for Euclidean distance is as follows:

$$d(x_i, x_j) = \sqrt{\sum_{c=1}^C (x_{ic} - x_{jc})^2} \quad (1)$$

In the formula,  $x_i$  and  $x_j$  denote the feature vectors of node  $i$  and node  $j$  respectively, and  $C$  stands for the feature dimension. A smaller distance value indicates higher similarity between node features.

Compared to random and fully-connected graph construction methods, traditional KNN graphing is structurally simple and computationally efficient, and can establish semantically meaningful topologies by capturing feature-level relationships [18]-[21]. However, it relies solely on original feature similarity while ignoring global data distribution and implicit connections. The resulting locally optimal topology will impair the learning performance of graph convolution.

### 3. Dynamic KNN Spatio-Temporal Graph Convolutional Network

#### 3.1. Overall Model Framework

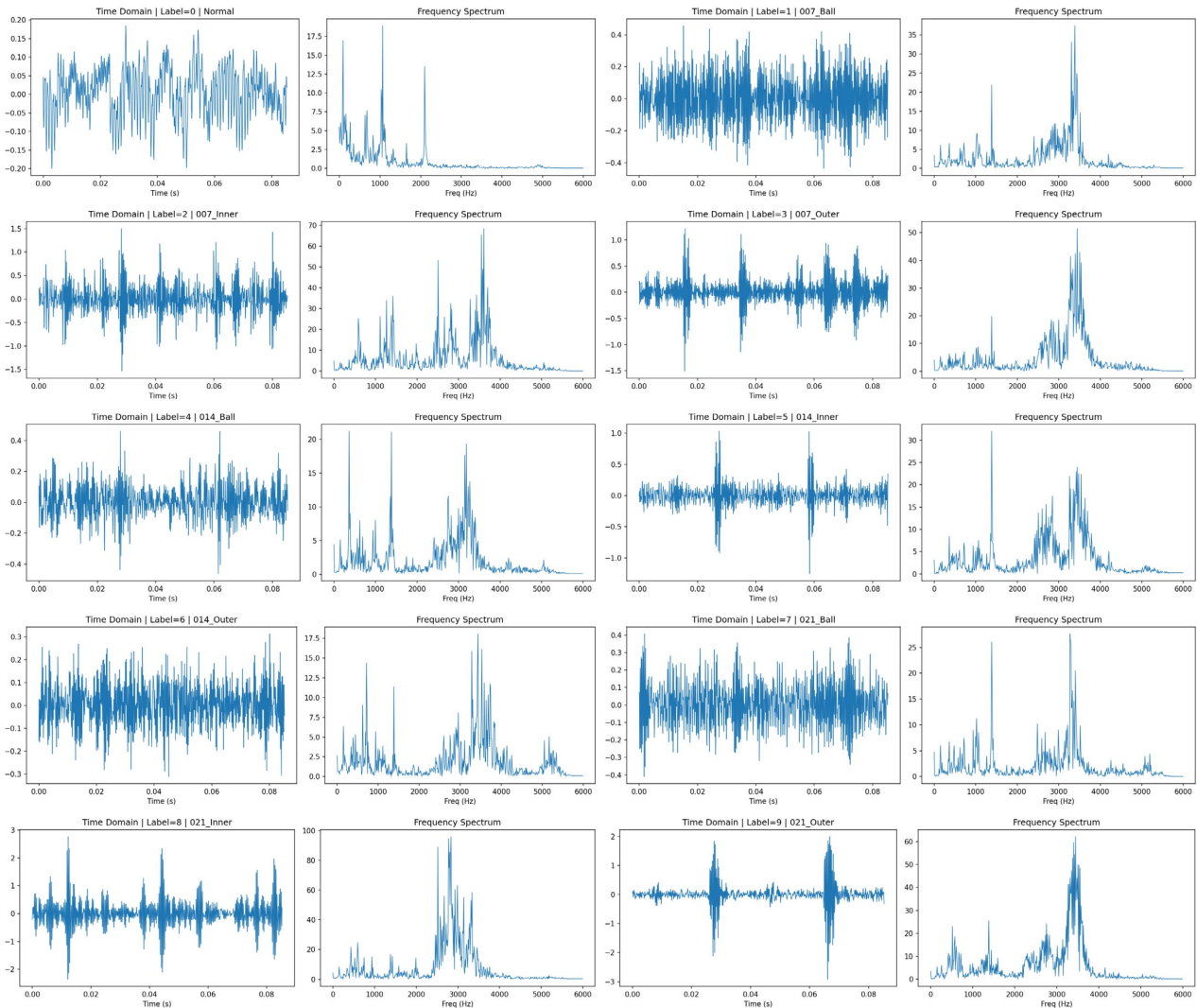


Figure 2. Structure of the proposed model.

The overall architecture of the proposed model is illustrated in **Figure 2**, which consists of five parts: data input layer, spatial-temporal dual-branch graph convolution layer (STDBGCN Layer), residual attention layer (RA Layer), feature fusion layer (FF Layer), and fault classification output layer (FCO Layer).

The raw one-dimensional vibration time-series signals of rolling bearings are fed into the model. The original data are divided into samples via sliding window segmentation, and fused time-domain and frequency-domain features are extracted as node features of the graph. A dynamic weighted K-NN algorithm combining temporal distance and cosine similarity is adopted to generate the adaptive adjacency matrix. The spatial-temporal dual-branch network learns spatial and temporal features in parallel, and residual connections together with the node attention module are used to optimize feature representation. Finally, global pooling and fully connected layers in the fault classification output layer realize the classification of bearing fault patterns.

### 3.2. Vibration Signal Preprocessing and Multi-Domain Node Feature Construction

The original vibration acceleration signals are segmented via non-overlapping sliding windows. The length of each sample is set to  $L$ , and each window segment is regarded as an individual node of the graph network.

We extract time-domain features for each signal segment, including mean value, standard deviation, peak value, kurtosis, waveform indicator, and impulse indicator. Meanwhile, frequency-domain features such as spectral centroid, mean square frequency, and variance frequency are also computed. The concatenation of time-domain and frequency-domain features yields a high-dimensional node feature vector  $x_i \in \mathbb{R}^d$ , which realizes the mapping from raw non-sequential signals to graph node features.

The feature cosine similarity between node  $i$  and node  $j$  is defined as:

$$Sim_{ij} = \frac{x_i \cdot x_j}{\|x_i\| \|x_j\|} \quad (2)$$

The temporal position distance between node  $i$  and node  $j$  is defined as:

$$Dis_{t,ij} = \frac{1}{|t_i - t_j| + \varepsilon} \quad (3)$$

Here,  $\varepsilon$  is a small constant preventing a zero denominator.  $t_i$  and  $t_j$  are sample temporal indices. Linear normalization is performed on node feature similarity and temporal position distance, respectively, to scale them into the range of  $[0, 1]$ .

$$\widetilde{Sim}_{ij} = \frac{Sim_{ij} - \min(Sim_{ab})}{\max(Sim_{ab}) - \min(Sim_{ab})} \quad (4)$$

$$\widetilde{Dis}_{t,ij} = \frac{Dis_{t,ij} - \min(Dis_{t,ab})}{\max(Dis_{t,ab}) - \min(Dis_{t,ab})} \quad (5)$$

where  $\widetilde{Sim}_{ij}$  and  $\widetilde{Dis}_{t,ij}$  denote the normalized feature similarity and temporal distance, respectively. The adjacency weights are obtained after weighted fusion. The overall weighted fusion distance is given by:

$$S_{ij} = \lambda \widetilde{Sim}_{ij} + (1 - \lambda) \cdot \widetilde{Dis}_{t,ij} \quad (6)$$

Here,  $\lambda$  denotes the weight balancing coefficient. To satisfy the requirements of graph convolution operations for the adjacency matrix,  $S_{ij}$  is normalized via the Softmax function. Here,  $k$  represents the indices of all neighboring nodes of node  $i$  which means the summation is conducted over all  $j$  corresponding to node  $i$ .

$$\hat{S}_{ij} = \frac{\exp(S_{ij})}{\sum_k \exp(S_{ik})} \quad (7)$$

Traditional methods adopt a globally fixed  $K$ . In this work, the number of neighbors is adaptively adjusted according to the density of sample features.

$$K_i = K_{base} + \eta \cdot \frac{\sum_j S_{ij}}{N} \quad (8)$$

Here,  $K_{base}$  stands for the base neighbor count,  $\eta$  is the regulation coefficient, and  $N$  is the total node number. Higher feature similarity corresponds to more adaptive neighbors, forming a dynamic topology.

### 3.3. Spatio-Temporal Dual-Branch Graph Convolution Module

The spatial branch adopts a multi-head graph attention network to explore fault topological correlations among nodes. It performs spatial aggregation on the features of dynamic graph nodes and learns the common and distinctive characteristics of samples.

Vibration signals exhibit continuous temporal evolution characteristics. In this paper, gated temporal convolution is introduced to model node sequences in the time dimension, so as to capture the time-varying dependencies of fault impacts and compensate for the lack of temporal information in pure spatial graph networks.

$$X_T = \text{Conv1d}(x_t, W_f) \odot \sigma(\text{Conv1d}(x_t, W_g)) \quad (9)$$

Here,  $x_t \in \mathbb{R}^{L \times d_s}$ .  $L$  is the temporal length,  $d_s$  is the output feature dimension of the spatial branch.  $W_f$  and  $W_g$  are one-dimensional convolution kernels for the feature branch and gating branch, and  $\sigma(\cdot)$  is the Sigmoid activation function.

The spatial and temporal feature vectors are concatenated along the channel dimension to construct the global spatio-temporal feature representation. The fused feature is defined as:

$$X_{ST} = \text{Concat}(X_S, X_T; \text{dim} = \text{channel}) \quad (10)$$

### 3.4. Residual Attention Module

To address the over-smoothing problem of deep graph convolution, we design residual skip connections and a node attention weighting module.

$$H_{out} = H_1 + H_{l-1} + \alpha H_{gcn} \tag{11}$$

Shallow original features are directly delivered to deep networks via skip connections to avoid feature homogenization after multi-layer aggregation. The node attention mechanism amplifies important fault features by weighting, suppresses noisy and redundant information, and enhances the representation of weak faults.

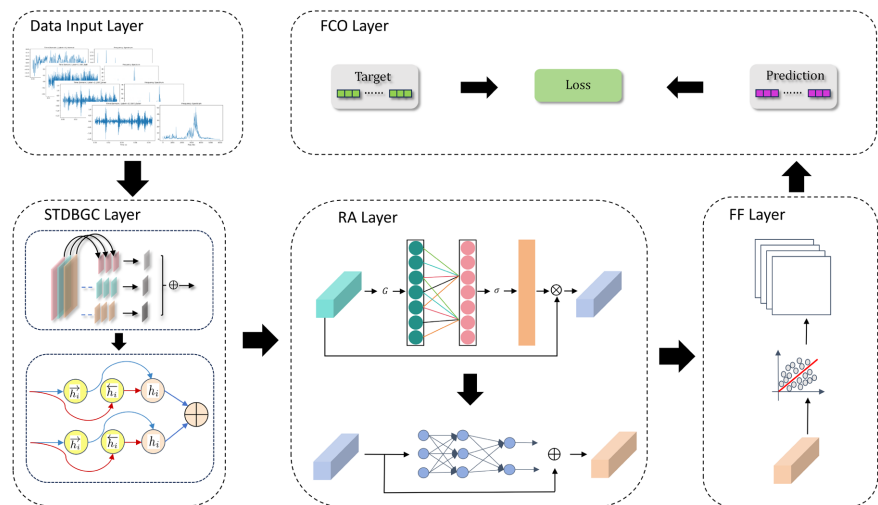
### 3.5. Fault Classification Output Layer

The fused spatio-temporal features are processed by global average pooling and then fed into the fully connected neural network. The Softmax activation function outputs the probability of each fault category to implement fault classification.

## 4. Experiments and Analysis

### 4.1. Datasets

We use the CWRU rolling bearing dataset from Case Western Reserve University. The CWRU dataset consists of vibration time series of various rolling bearing conditions collected from the test rig, as shown in **Figure 3**. The test bearing is the 6205-2RS deep groove ball bearing with a sampling frequency of 12 kHz. Four load conditions are set, namely 0 hp, 1 hp, 2 hp, and 3 hp. The fault types include normal condition (NC), inner race fault (IR), outer race fault (OR), and ball fault (B). Fault defect sizes are 0.1778 mm, 0.3556 mm, and 0.5334 mm. The ten fault modes are shown in **Table 1**. To strictly meet the generalization requirements for industrial applications, the training, validation, and test sets are divided according to experimental working conditions and acquisition batches. Specifically, the data collected under 1 hp and 2 hp loads are used as the training and validation sets, while the data from the independent 3 hp working condition is adopted as the test set. All sample windows segmented from the original vibration signals are assigned to a single dataset only, without distribution across the training and test sets. The dataset is split into a training set (70%), a validation set (20%), and a test set (10%).



**Figure 3.** Test rig for rolling bearings.

**Table 1.** Classification information of rolling bearings.

Fault Name	Label	Fault Location	Bearing Damage Level (mm)	Sample Collection
Normal	0	Normal Bearing	0	3000
007-BALL	1	Ball Fault	0.18	3000
007-IR	2	Inner Race Fault	0.18	3000
007-OUT	3	Outer Race Fault	0.18	3000
014-BALL	4	Ball Fault	0.36	3000
014-IR	5	Inner Race Fault	0.36	3000
014-OUT	6	Outer Race Fault	0.36	3000
021-BALL	7	Ball Fault	0.54	3000
021-IR	8	Inner Race Fault	0.54	3000
021-OUT	9	Outer Race Fault	0.54	3000

## 4.2. Data Preprocessing

Sliding window segmentation is applied to the original vibration signals to construct training samples. The sliding window length is set to 1024 sampling points with a sliding step of 512 sampling points, corresponding to an overlap rate of 50%, which fully preserves the temporal characteristics of fault impulses. Variational Mode Decomposition (VMD) is used for signal denoising and reconstruction. Afterwards, each sample window is standardized to zero mean and unit variance to eliminate dimensional and amplitude interference .

## 4.3. Baseline Models

Four mainstream methods are selected for comparative experiments:

1) SVM: A classic traditional machine learning algorithm. It constructs an optimal hyperplane to achieve sample classification, and performs well on small-scale datasets. It is widely used in traditional bearing fault diagnosis as a benchmark method.

2) LSTM: A typical recurrent neural network. It is designed to capture long-range temporal dependencies of time-series signals and is widely applied to vibration signal analysis and time-series fault classification tasks.

3) KNN-GCN: It first constructs a graph structure based on K-nearest neighbor relationships of samples, then leverages graph convolution to aggregate node features. This method combines traditional distance metric learning with graph neural networks.

4) GAT: It introduces the attention mechanism into graph learning, adaptively assigns different weights to neighboring nodes during feature aggregation, and effectively highlights important associative relationships between nodes.

#### 4.4. Hyperparameters

All baseline models adopt the same dataset partition. The Adam optimizer is used in the experiments with a learning rate of 0.001 and a batch size of 32. The total training epochs are set to 100. CrossEntropyLoss is selected as the loss function, and the basic adjacency number  $K_{base}$  is set to 8.

#### 4.5. Experimental Results

Multi-class fault diagnosis experiments are conducted on the bearing fault dataset. We comprehensively evaluate the overall performance of the proposed AdaD-KNN-STGCN model from multiple perspectives, including model convergence, classification accuracy, confusion matrix, and feature visualization. Meanwhile, comparative analysis is carried out against mainstream algorithms such as SVM, LSTM, KNN-GCN, and GAT.

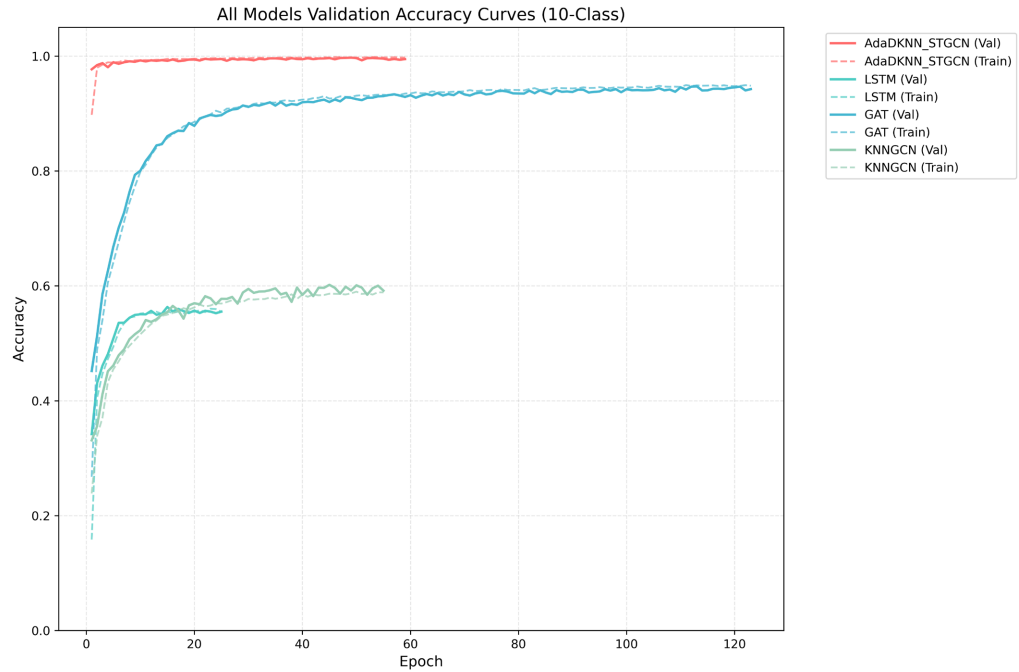
As shown in **Table 2**, the proposed AdaDKNN-STGCN model achieves an accuracy of 99.78% in the 10-class fault diagnosis task, which is remarkably higher than that of the comparative models including SVM, LSTM, KNN-GCN, and GAT. The standard deviation is merely 0.08, indicating the optimal stability of the model. According to the paired t-test for significance, the p-values between all comparative models and the proposed model are less than 0.05. At the 95% confidence level, the performance difference is statistically significant. It is statistically verified that the diagnostic performance of the proposed model is remarkably superior to existing mainstream algorithms. This verifies the effectiveness and superiority of the proposed method for fault diagnosis.

**Table 2.** Table comparison of fault diagnosis accuracy of different models.

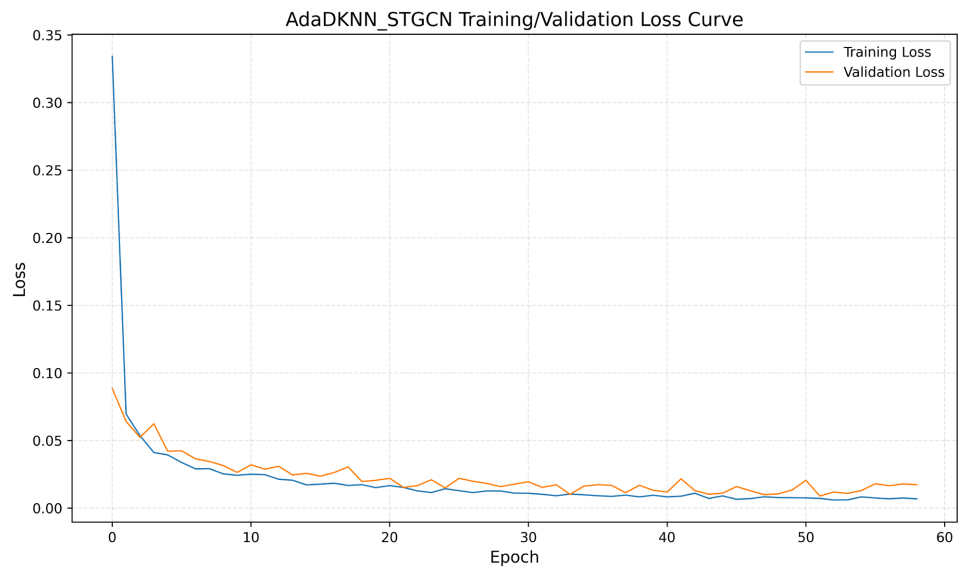
Model	SVM	LSTM	KNN-GCN	GAT	AdaDKNN-STGCN (Ours)
Accuracy (%)	62.33	57.27	60.87	95.67	<b>99.78</b>
Std (5 runs)	0.31	0.35	0.28	0.22	<b>0.08</b>
95% Confidence Interval (CI)	[62.02, 62.64]	[56.99, 57.55]	[60.62, 61.12]	[95.47, 95.87]	[99.72, 99.82]
p-value	p < 0.001	p < 0.001	p < 0.001	p < 0.001	-

The training and validation accuracy curves of different models are presented in **Figure 4**. It can be observed that AdaDKNN-STGCN converges faster. In addition, the training and validation accuracy curves of the model closely overlap, indicating its outstanding generalization ability and effective suppression of overfitting. **Figure 5** illustrates the loss curve of AdaDKNN-STGCN, which further demonstrates stable training performance. The classification details are analyzed via the confusion matrix in **Figure 6**. The results reveal that the proposed model can accurately identify samples under normal conditions as well as bearing samples with different damage severities and fault locations.

To explore the feature learning capability of the model, t-SNE is adopted to reduce the dimensionality of high-dimensional features from the output layer for 2D visualization, as shown in **Figure 7**. The feature distribution shows that samples of the same fault category are highly clustered in the feature space. This proves that AdaDKNN-STGCN can fully extract deep features from vibration signals and effectively distinguish different fault types and damage degrees. The excellent feature representation capability lays a solid foundation for the high diagnosis accuracy of the model.



**Figure 4.** Comparison curves of training and validation accuracy of models.



**Figure 5.** Training and validation curve for loss.

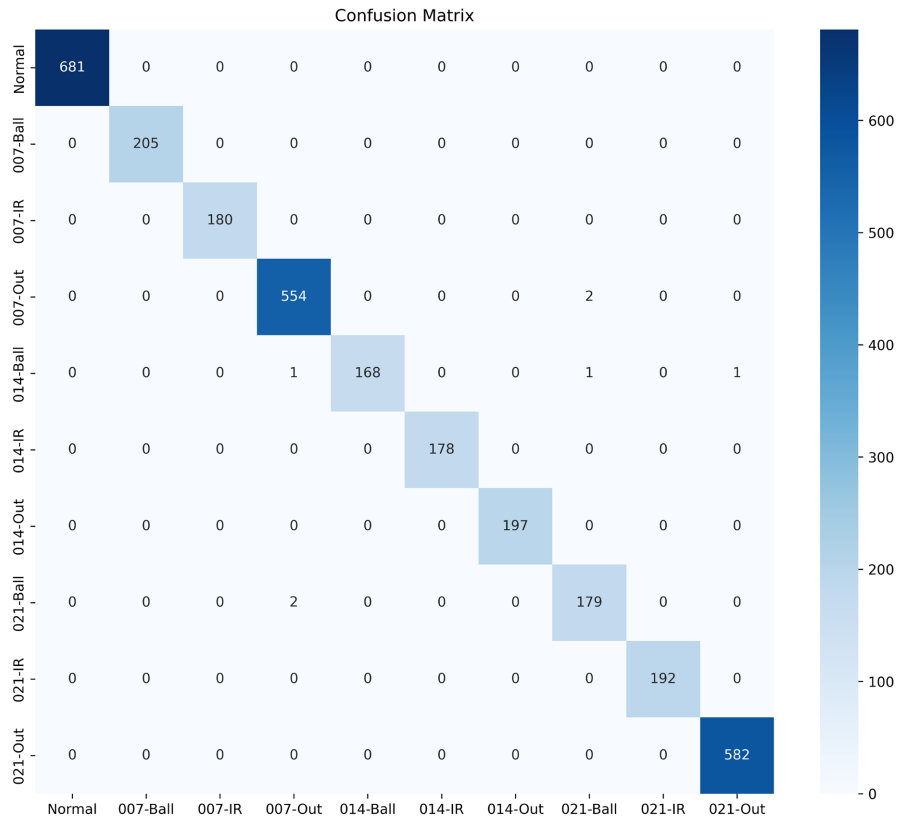


Figure 6. Confusion matrix for each category.

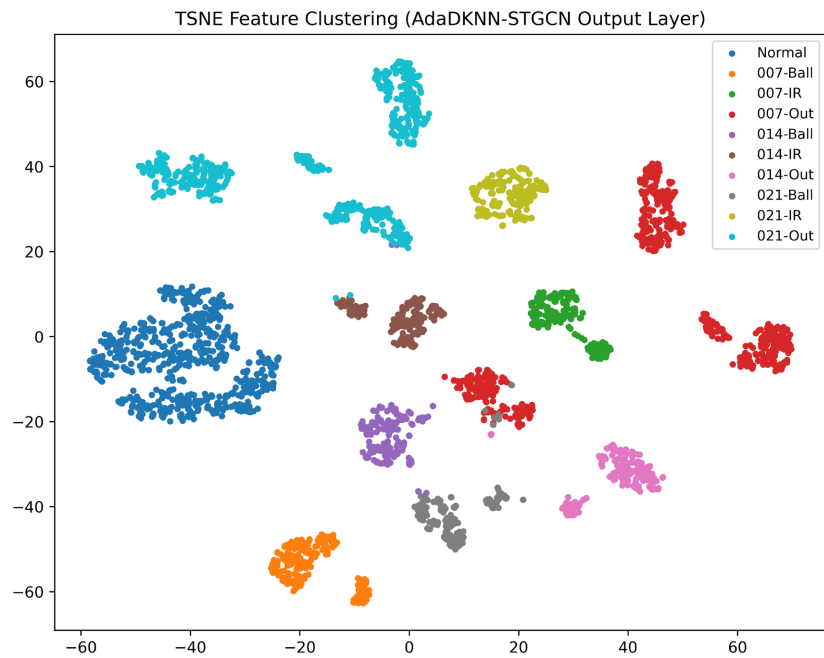


Figure 7. Visualization based on t-SNE algorithm.

## 5. Conclusion

Aiming at the drawbacks of traditional K-nearest neighbor graph convolutional

networks in bearing fault diagnosis, such as poor working condition adaptability, feature over-smoothing, and insufficient mining of temporal features, this paper proposes an AdaDKNN-STGCN model. The model constructs graph nodes by fusing time-frequency features and adopts a spatio-temporal dual-branch structure to simultaneously learn the spatial topology and temporal evolution features of signals. Experimental results demonstrate that the proposed model achieves an accuracy of 99.78% on the 10-classification task and presents excellent generalization and robustness under complex working conditions. This method effectively remedies the deficiencies of traditional algorithms and provides a new technical solution for intelligent fault diagnosis of rotating machinery.

## Acknowledgements

This research was partially supported by the Science and Technology Project of China Tobacco Hebei Industrial Co., Ltd. (Grant No. HBZY2026A016).

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

- [1] Ni, Q., Ji, J.C., Halkon, B., Feng, K. and Nandi, A.K. (2023) Physics-informed Residual Network (PIResNet) for Rolling Element Bearing Fault Diagnostics. *Mechanical Systems and Signal Processing*, **200**, Article 110544. <https://doi.org/10.1016/j.ymssp.2023.110544>
- [2] Geetha, G. and Geethanjali, P. (2024) An Efficient Method for Bearing Fault Diagnosis. *Systems Science & Control Engineering*, **12**, Article 2329264. <https://doi.org/10.1080/21642583.2024.2329264>
- [3] Wang, L. and Zhao, W. (2025) An Ensemble Deep Learning Network Based on 2D Convolutional Neural Network and 1D LSTM with Self-Attention for Bearing Fault Diagnosis. *Applied Soft Computing*, **172**, Article 112889. <https://doi.org/10.1016/j.asoc.2025.112889>
- [4] Wang, J., Mo, Z., Zhang, H. and Miao, Q. (2019) A Deep Learning Method for Bearing Fault Diagnosis Based on Time-Frequency Image. *IEEE Access*, **7**, 42373-42383. <https://doi.org/10.1109/access.2019.2907131>
- [5] Abtane, M., Dahi, K., Martinez, H., Sedki, M., El Kimi, H., Dahhassi, C., *et al.* (2025) Axle Bearing Fault Diagnosis for High-Speed Trains: A Comprehensive Review of Methodologies, Technologies, Challenges and Emerging Trends. *Measurement*, **251**, Article 117098. <https://doi.org/10.1016/j.measurement.2025.117098>
- [6] Zhu, H., He, Z., Wei, J., Wang, J. and Zhou, H. (2021) Bearing Fault Feature Extraction and Fault Diagnosis Method Based on Feature Fusion. *Sensors*, **21**, Article 2524. <https://doi.org/10.3390/s21072524>
- [7] Keshun, Y., Puzhou, W., Peng, H. and Yingkui, G. (2025) A Sound-Vibration Physical-Information Fusion Constraint-Guided Deep Learning Method for Rolling Bearing Fault Diagnosis. *Reliability Engineering & System Safety*, **253**, Article 110556. <https://doi.org/10.1016/j.ress.2024.110556>
- [8] Dong, Z., Jiang, Y., Jiao, W., Zhang, F., Wang, Z., Huang, J., *et al.* (2025) Double Attention-Guided Tree-Inspired Grade Decision Network: A Method for Bearing

- Fault Diagnosis of Unbalanced Samples under Strong Noise Conditions. *Advanced Engineering Informatics*, **64**, Article 103004. <https://doi.org/10.1016/j.aei.2024.103004>
- [9] Chen, Z., Zhu, L., Lu, H., Chen, S., Zhu, F., Liu, S., et al. (2024) Research on Bearing Fault Diagnosis Based on Improved Genetic Algorithm and BP Neural Network. *Scientific Reports*, **14**, Article No. 15527. <https://doi.org/10.1038/s41598-024-66318-0>
- [10] Yu, L., Qu, J., Gao, F. and Tian, Y. (2019) A Novel Hierarchical Algorithm for Bearing Fault Diagnosis Based on Stacked LSTM. *Shock and Vibration*, **2019**, Article 2756284. <https://doi.org/10.1155/2019/2756284>
- [11] Zou, P., Hou, B., Lei, J. and Zhang, Z. (2020) Bearing Fault Diagnosis Method Based on EEMD and LSTM. *International Journal of Computers Communications & Control*, **15**, 1-14. <https://doi.org/10.15837/ijccc.2020.1.3780>
- [12] Ruan, D., Chen, X., Gühmann, C. and Yan, J. (2023) Improvement of Generative Adversarial Network and Its Application in Bearing Fault Diagnosis: A Review. *Lubricants*, **11**, Article 74. <https://doi.org/10.3390/lubricants11020074>
- [13] He, D., Wu, J., Jin, Z., Huang, C., Wei, Z. and Yi, C. (2025) AGFCN: A Bearing Fault Diagnosis Method for High-Speed Train Bogie under Complex Working Conditions. *Reliability Engineering & System Safety*, **258**, Article 110907. <https://doi.org/10.1016/j.res.2025.110907>
- [14] Hou, L., Yi, H., Jin, Y., Gui, M., Sui, L., Zhang, J., et al. (2023) Inter-Shaft Bearing Fault Diagnosis Based on Aero-Engine System: A Benchmarking Dataset Study. *Journal of Dynamics, Monitoring and Diagnostics*, **2**, 228-242. <https://doi.org/10.37965/jdmd.2023.314>
- [15] Hu, W., Xin, G., Wu, J., An, G., Li, Y., Feng, K., et al. (2023) Vibration-Based Bearing Fault Diagnosis of High-Speed Trains: A Literature Review. *High-Speed Railway*, **1**, 219-223. <https://doi.org/10.1016/j.hspr.2023.11.001>
- [16] Liang, B. and Feng, W. (2023) Bearing Fault Diagnosis Based on ICEEMDAN Deep Learning Network. *Processes*, **11**, Article 2440. <https://doi.org/10.3390/pr11082440>
- [17] Tao, L., Liu, H., Ning, G., Cao, W., Huang, B. and Lu, C. (2025) LLM-Based Framework for Bearing Fault Diagnosis. *Mechanical Systems and Signal Processing*, **224**, Article 112127. <https://doi.org/10.1016/j.ymssp.2024.112127>
- [18] Eren, L., Ince, T. and Kiranyaz, S. (2018) A Generic Intelligent Bearing Fault Diagnosis System Using Compact Adaptive 1D CNN Classifier. *Journal of Signal Processing Systems*, **91**, 179-189. <https://doi.org/10.1007/s11265-018-1378-3>
- [19] Yuan, L., Lian, D., Kang, X., Chen, Y. and Zhai, K. (2020) Rolling Bearing Fault Diagnosis Based on Convolutional Neural Network and Support Vector Machine. *IEEE Access*, **8**, 137395-137406. <https://doi.org/10.1109/access.2020.3012053>
- [20] Fu, W., Li, S., Wen, B., Zheng, B., Liao, W. and Tan, C. (2025) A Review of Rolling Bearing Fault Diagnosis: Data Preprocessing and Model Optimization. *Measurement Science and Technology*, **36**, Article 062002. <https://doi.org/10.1088/1361-6501/add7fb>
- [21] Wang, J., Li, T., Yang, Y., Chen, S. and Zhai, W. (2025) DiagLLM: Multimodal Reasoning with Large Language Model for Explainable Bearing Fault Diagnosis. *Science China Information Sciences*, **68**, Article 160103. <https://doi.org/10.1007/s11432-024-4333-7>