Deep Learning-Driven Adaptive Machining Parameter Optimization For High-Precision CNC Milling

Xiaoli Qu*

Zhengzhou Technical College, No. 081, Zhengshang Road, Zhengzhou City, China

Corresponding author. E-mail: 2430175060@qq.com

Received: Jan. 21, 2026; Accepted: Mar. 09, 2026

To address the limitations of single-modal data and low real-time performance in traditional anomaly diagnosis for CNC turning processes, this paper proposes a novel framework integrating adaptive multi-modal data fusion and lightweight graph neural network (GNN) for real-time anomaly diagnosis. First, multi-modal data (vibration, spindle current, and cutting force) are collected and preprocessed to extract time-frequency domain features. A mutual information-based graph construction method is designed to model the intrinsic correlations between multi-modal features, converting non-Euclidean feature data into structured graph data. Then, an event-driven lightweight GNN (EL-GNN) is proposed, which adopts a hierarchical propagation mechanism to reduce redundant computations and realizes millisecond-level inference. A cross-attention fusion module is embedded in the GNN to dynamically assign weights to different modal features, enhancing the robustness to noise. Experiments are conducted on a self-built CNC turning test platform and the public tool wear dataset. Results show that the proposed framework achieves an anomaly diagnosis accuracy of 98.73%, a recall rate of 98.51%, and a P99 inference latency of 28.3 ms , outperforming traditional machine learning methods and deep learning models by $3.2\% - 8.9\%$ in accuracy. This framework provides a reliable solution for intelligent predictive maintenance in CNC turning processes, balancing diagnostic accuracy and real-time performance.

Keywords: Deep Learning-Driven; Adaptive Machining; Parameter Optimization; High-Precision CNC Milling

http://dx.doi.org/10.6180/jase.202608_31.067

1. Introduction

Computer Numerical Control (CNC) turning is a core process in discrete manufacturing, it is widely used in aerospace, automotive, and precision machinery industries. The stable operation of CNC turning equipment directly determines production efficiency, product quality, and production safety [1, 2]. Anomalies such as tool wear, spindle misalignment, and cutting parameter deviation often occur during long-term high-load operation. If not detected in time, these anomalies may lead to tool breakage, workpiece scrapping, or even equipment damage, resulting in huge economic losses [3]. Therefore, real-time and accurate anomaly diagnosis is crucial for realizing predictive maintenance of CNC turning equipment and promoting the transformation to Industry 5.0 [4, 5].

Traditional anomaly diagnosis methods for CNC turning mainly rely on manual inspection or single-modal data analysis. Manual inspection is inefficient and heavily dependent on expert experience, making it difficult to meet the requirements of real-time monitoring in mass production. Single-modal data (e.g., vibration, current) can only reflect partial operating status of the equipment, lacking comprehensive information coverage [6]. For example, vibration signals are sensitive to mechanical faults but vulnerable to environmental noise, while current signals can characterize load changes but have low sensitivity to early tool wear. These limitations lead to low diagnostic accuracy

and poor robustness of single-modal methods.

With the development of sensor technology and artificial intelligence, multi-modal data fusion and deep learning have become research hotspots in anomaly diagnosis. Multi-modal data fusion integrates information from multiple sensors to compensate for the deficiencies of single-modal data, improving diagnostic accuracy and robustness [7]. Existing fusion methods can be divided into data-level, feature-level, and decision-level fusion. Feature-level fusion, which extracts and integrates features from different modalities first, has been widely used due to its balance between information integrity and computational efficiency [8]. For example, Gao et al. [9] proposed a cross-attention mechanism to fuse vibration and current features for motor fault diagnosis, achieving better performance than traditional fusion methods.

Graph Neural Networks (GNNs) have shown significant advantages in processing structured data by capturing spatial correlations between nodes [10]. In industrial anomaly diagnosis, GNNs are used to model the topological relationships between process parameters or sensor nodes. However, most existing GNN-based methods are designed for static graphs, ignoring the dynamic characteristics of CNC turning processes. Moreover, the high computational complexity of GNNs leads to poor real-time performance, making it difficult to meet the millisecond-level response requirement of CNC equipment [11]. To address this issue, some lightweight GNN frameworks have been proposed. For instance, InkStream adopts an event-driven mechanism to avoid unnecessary updates, reducing inference time from hours to milliseconds; DGS realizes low-latency GNN inference through storage-computation separation and pre-sampling strategies.

Despite these advances, there are still two key challenges in applying multi-modal fusion and GNN to CNC turning anomaly diagnosis. (1) The intrinsic correlations between multi-modal features of CNC turning are complex and non-linear, and existing graph construction methods fail to adaptively model these correlations. (2) The trade-off between diagnostic accuracy and real-time performance is not well balanced, as lightweight GNNs often sacrifice accuracy for speed. To solve these problems, this paper proposes an adaptive multi-modal fusion GNN framework for real-time anomaly diagnosis in CNC turning processes. The main contributions of this paper are as follows.

(1) An adaptive multi-modal graph construction method based on mutual information is proposed, which quantifies the correlation between different modal features and constructs a dynamic feature graph, realizing effective modeling of non-linear correlations between multi-modal data.

(2) An event-driven lightweight GNN (EL-GNN) is designed, which adopts a hierarchical propagation mechanism and incremental update strategy to reduce redundant computations, ensuring millisecond-level inference while maintaining diagnostic accuracy.

(3) A cross-attention fusion module embedded in EL-GNN is developed to dynamically adjust the weight of each modal feature according to the operating status, enhancing the robustness of the model to noise and modal missing.

(4) Comprehensive experiments are conducted on a self-built test platform and public dataset, verifying that the proposed framework outperforms existing methods in both diagnostic performance and real-time capability.

2. Materials and methods

The proposed multi-modal fusion GNN framework for real-time anomaly diagnosis in CNC turning processes consists of four modules: multi-modal data acquisition and preprocessing, adaptive graph construction, EL-GNN-based feature propagation, and cross-attention fusion & anomaly classification. The overall architecture is shown in Fig. 1.

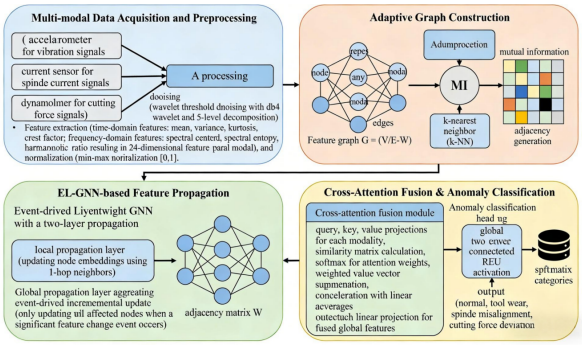

Fig. 1. Overall architecture of the proposed framework

2.1. Multi-Modal Data Acquisition and Preprocessing

Three types of sensors are deployed on the CNC turning machine to collect multi-modal data: (1) Accelerometers mounted on the tool rest and spindle housing to collect vibration signals (sampling rate: 10 kHz); (2) Current sensors installed in the spindle drive circuit to collect spindle current signals (sampling rate: 5 kHz); (3) Dynamometers embedded in the workpiece fixture to collect cutting force signals (sampling rate: 2 kHz). The collected raw data contain noise interference from the environment and equipment, so preprocessing is necessary.

The preprocessing process includes three steps. (1) Denoising. The wavelet threshold denoising method is used to eliminate high-frequency noise, with db 4 wavelet as the

base wavelet and 5-level decomposition. (2) Feature extraction. Time-domain features (mean, variance, kurtosis, crest factor) and frequency-domain features (spectral centroid, spectral entropy, harmonic ratio) are extracted from each modal data, resulting in a 24-dimensional feature vector for each modal. (3) Normalization. The min-max normalization method is adopted to scale all features to $[0, 1]$, avoiding the influence of different value ranges.

2.2. Adaptive Graph Construction Based on Mutual Information

To model the correlations between multi-modal features, we construct a feature graph $G = (\mathcal{V}, \mathcal{E}, \mathbf{W})$, where $\mathcal{V}$ is the set of nodes. $\mathcal{E}$ is the set of edges. $\mathbf{W}$ is the adjacency matrix. Each node $v_i \in \mathcal{V}$ represents a feature from any modal, and the edge weight w_{ij} represents the correlation between node v_i and v_j.

Mutual Information (MI) is used to quantify the correlation between features, which measures the amount of information shared between two random variables. For two features X and Y, the MI is calculated as:

$$MI(X,Y) = \int_x \int_y p(x,y) \log \frac{p(x,y)}{p(x)p(y)} dxdy \tag{1}$$

Where $p(x,y)$ is the joint probability density function of X and Y.$p(x)$ and $p(y)$ are the marginal probability density functions. To improve computational efficiency, we use the k-nearest neighbor ($k-NN$) method to estimate the probability density.

The adjacency matrix $\mathbf{W}$ is constructed based on MI:

$$w_{ij} = \exp\left(-\frac{MI\left(X_i, X_j\right)}{\sigma^2}\right) \tag{2}$$

If X_j is one of the k-nearest neighbors of X_i, then $w_{ij} = 0$, where σ is the bandwidth parameter. This adaptive graph construction method can dynamically adjust the graph structure according to the actual data distribution, avoiding the subjectivity of manual graph construction. The graph construction process is shown in Fig. 2.

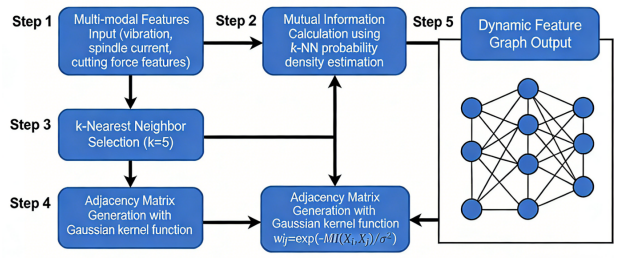

Fig. 2. Flowchart of adaptive graph construction

2.3. Event-Driven Lightweight GNN (EL-GNN)

To ensure real-time inference, EL-GNN is designed based on the event-driven mechanism and hierarchical propagation strategy, which only updates the affected nodes instead of the entire graph. The core idea is inspired by InkStream and DGS, adapting to the dynamic characteristics of CNC turning processes.

2.3.1. *Hierarchical Propagation Mechanism*

EL-GNN adopts a two-layer propagation structure to balance feature expression ability and computational complexity. The first layer (local propagation layer) updates the node embeddings using only the 1-hop neighbors, capturing local feature correlations. The second layer (global propagation layer) aggregates the local embeddings of the entire graph to obtain global feature information. The propagation formula is as follows:

$$\begin{cases} \mathbf{h}_i^{(1)} = \sigma\left(\mathbf{W}_1 \sum_{j \in \mathcal{N}(i)} w_{ij} \mathbf{x}_j + \mathbf{b}_1\right) \\ \mathbf{h}_i^{(2)} = \sigma\left(\mathbf{W}_2 \frac{1}{|\mathcal{V}|} \sum_{k \in \mathcal{V}} \mathbf{h}_k^{(1)} + \mathbf{b}_2\right) \end{cases} \tag{3}$$

Where $\mathbf{h}_i^{(1)}$ and $\mathbf{h}_i^{(2)}$ are the local and global embeddings of node i.$\mathcal{N}(i)$ is the 1-hop neighbor set of node i. $\mathbf{W}_1$ and $\mathbf{W}_2$ are the weight matrices. $\mathbf{b}_1$ and $\mathbf{b}_2$ are the bias vectors, and σ is the ReLU activation function.

2.3.2. *Event-Driven Incremental Update*

An event is defined as a significant change in the multi-modal features (exceeding a preset threshold τ). When an event occurs, only the affected nodes (the node with feature change and its k-hop neighbors) are updated, while other nodes retain their previous embeddings. This strategy reduces redundant computations significantly. The update process is controlled by the event queue, which processes events in chronological order to ensure real-time performance. The event-driven update logic is shown in Fig. 3.

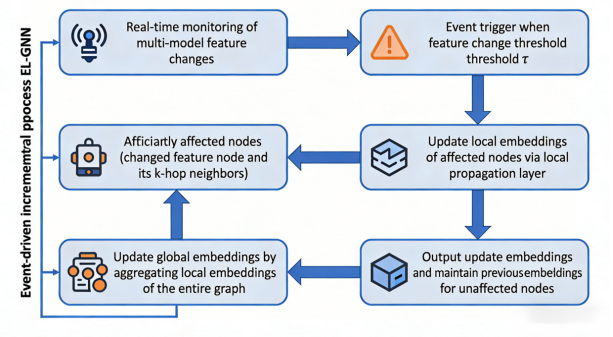

Fig. 3. Event-driven incremental update process of EL-GNN

2.4. Cross-Attention Fusion and Anomaly Classification

To fully exploit the complementary information of multi-modal features and suppress the interference of redundant or noisy modal data, a cross-attention fusion module is embedded in the EL-GNN architecture. This module dynamically assigns adaptive weights to each modal feature based on the global semantic information of the graph, ensuring that critical modalities contribute more to the final diagnosis result. The cross-attention mechanism is designed to capture the inter-modal dependencies, which is particularly effective for handling the dynamic and heterogeneous characteristics of CNC turning multi-modal data. After feature fusion, a classification head with hierarchical structure is adopted to map the fused features to anomaly categories, ensuring high diagnostic accuracy while maintaining computational efficiency.

2.4.1. Cross-Attention Fusion Mechanism

First, the global embeddings of each modality are extracted from the output of EL-GNN's global propagation layer. For modality m (where $m \in \{1, 2, 3\}$ corresponds to vibration, spindle current, and cutting force, respectively), the global embedding matrix is denoted as $\mathbf{H}_m \in \mathbb{R}^{N \times d}$, where N is the number of feature nodes and d is the dimension of each node embedding. To capture inter-modal correlations, we project each modal embedding into query, key, and value spaces using modality-specific linear transformations.

$$\begin{cases} \mathbf{Q}_m = \mathbf{H}_m \mathbf{W}_q^m + \mathbf{b}_q^m \\ \mathbf{K}_m = \mathbf{H}_m \mathbf{W}_k^m + \mathbf{b}_k^m \\ \mathbf{V}_m = \mathbf{H}_m \mathbf{W}_v^m + \mathbf{b}_v^m \end{cases} \tag{4}$$

Where $\mathbf{W}_q^m \in \mathbb{R}^{d \times d_k}, \mathbf{W}_k^m \in \mathbb{R}^{d \times d_k}, \mathbf{W}_v^m \in \mathbb{R}^{d \times d_v}$ are the learnable weight matrices for modality m.$\mathbf{b}_q^m \in \mathbb{R}^{d_k}, \mathbf{b}_k^m \in \mathbb{R}^{d_k}, \mathbf{b}_v^m \in \mathbb{R}^{d_v}$ are the corresponding bias vectors. d_k is the dimension of query and key vectors, and d_v is the dimension of value vectors. To balance computational complexity and representation ability, we set $d_k = d_v = d/2$ in this paper.

The cross-attention weight α_m for modality m is calculated by normalizing the similarity between the query vector of modality m and the key vectors of all modalities.

This cross-modal similarity calculation enables the model to focus on the most relevant modalities for anomaly diagnosis. The similarity matrix $\mathbf{S} \in \mathbb{R}^{M \times M}$ (where $M = 3$ is the number of modalities) is defined as:

$$S_{mn} = \frac{1}{\sqrt{d_k}} \operatorname{Tr}\left(\mathbf{Q}_m^T \mathbf{K}_n\right) \tag{5}$$

Where S_{mn} represents the similarity between modality m and modality $n \cdot \operatorname{Tr}(\cdot)$ denotes the trace of a matrix. The cross-attention weights are obtained by applying the softmax function to the similarity matrix along the column direction.

$$\alpha_{mn} = \frac{\exp\left(S_{mn}\right)}{\sum_{k=1}^{M} \exp\left(S_{mk}\right)} \tag{6}$$

Where $\alpha_{mn} \in [0, 1]$ is the weight of modality n when fusing features for modality m, and $\sum_{k=1}^{M} \alpha_{mk} = 1$. The weighted value vector for each modality m is then computed as:

$$\widetilde{\mathbf{V}}_m = \sum_{n=1}^{M} \alpha_{mn} \mathbf{V}_n \tag{7}$$

Finally, the fused global feature matrix $\mathbf{H}_{\text{fusion}} \in \mathbb{R}^{N \times d_v}$ is obtained by concatenating the weighted value vectors of all modalities and performing a linear projection to unify the dimension.

$$\mathbf{H}_{\text{fusion}} = \text{Concat}\left(\widetilde{\mathbf{V}}_1, \widetilde{\mathbf{V}}_2, \widetilde{\mathbf{V}}_3\right) \mathbf{W}_{\text{proj}} + \mathbf{b}_{\text{proj}} \tag{8}$$

Where Concat $(\cdot)$ denotes the concatenation operation along the feature dimension, $\mathbf{W}_{\text{proj}} \in \mathbb{R}^{Md_v \times d_v}$ is the projection weight matrix. $\mathbf{b}_{\text{proj}} \in \mathbb{R}^{d_v}$ is the projection bias vector.

2.4.2. Anomaly Classification Head

To convert the fused global features into anomaly classification results, a two-layer fully connected (FC) network with batch normalization (BN) and activation function is designed as the classification head. First, the fused feature matrix $\mathbf{H}_{\text{fusion}}$ is aggregated into a global feature vector $\mathbf{h}_{\text{global}} \in \mathbb{R}^{d_v}$ using global average pooling, which reduces the number of parameters and avoids overfitting.

$$\mathbf{h}_{\text{global}} = \frac{1}{N} \sum_{\text{i}=1}^{(\text{i})} \mathbf{H}_{\text{fusion}}^{(i)} \tag{9}$$

Where $\mathbf{H}_{\text{fusion}}^{(i)}$ is the i-th row of $\mathbf{H}_{\text{fusion}}$ corresponding to the fused embedding of the i-th feature node.

The global feature vector is then fed into the two-layer FC network. The first FC layer with BN and ReLU activation function enhances the non-linear expression ability of the model.

$$\mathbf{h}_{fc1} = \sigma\left(BN\left(\mathbf{h}_{global}\mathbf{W}_{fc1} + \mathbf{b}_{fc1}\right)\right) \tag{10}$$

Where $\mathbf{W}_{fc1} \in \mathbb{R}^{d_v \times 2d_v}$ is the weight matrix of the first FC layer. $\mathbf{b}_{fc1} \in \mathbb{R}^{2d_v}$ is the bias vector. $BN(\cdot)$ denotes the batch normalization operation, and $\sigma(\cdot)$ is the ReLU activation function.

The second FC layer maps the feature vector to the number of anomaly categories C and outputs the classification probability using the softmax function.

$$\mathbf{y}_{\text{pred}} = \text{softmax}\left(\mathbf{h}_{fc1}\mathbf{W}_{fc2} + \mathbf{b}_{fc2}\right) \tag{11}$$

Where $\mathbf{W}_{fc2} \in \mathbb{R}^{2d_v \times C}$ is the weight matrix of the second FC layer. $\mathbf{b}_{fc2} \in \mathbb{R}^C$ is the bias vector. $\mathbf{y}_{\text{pred}} \in \mathbb{R}^C$ is the predicted probability vector with $y_{\text{pred},c}$ representing the probability that the sample belongs to category C.

2.4.3. Loss Function Design

To optimize the entire framework, the cross-entropy loss function is adopted to minimize the difference between the predicted probability and the true label. For a batch of B samples, the total loss $\mathcal{L}$ is calculated as:

$$\mathcal{L} = -\frac{1}{B}\sum_{b=1}^{B}\sum_{c=1}^{C} y_{b,c}\log\left(y_{\text{pred},b,c} + \epsilon\right) \tag{12}$$

Where $y_{b,c}$ is the one-hot true label of the b-th sample. $y_{\text{pred},b,c}$ is the predicted probability of the $b-$ th sample belonging to category C.$E = 10^{-8}$ is a small constant to avoid the logarithm of zero. Additionally, a weight decay term $\lambda\|\Theta\|_2^2$ is added to the loss function to prevent overfitting, where Θ is the set of all learnable parameters in the framework. $\lambda = 10^{-5}$ is the weight decay coefficient.

A cross-attention fusion module is embedded in EL-GNN to dynamically assign weights to different modal features. For each modal m, the attention weight α_m is calculated based on the global embeddings of the modal features.

$$\alpha_m = \frac{\exp\left(\mathbf{q}_m \cdot \mathbf{k}_m^T/\sqrt{d}\right)}{\sum_{n=1}^{M}\exp\left(\mathbf{q}_n \cdot \mathbf{k}_n^T/\sqrt{d}\right)} \tag{13}$$

Where M is the number of modalities. $\mathbf{q}_m$ and $\mathbf{k}_m$ are the query and key vectors of modal m, and d is the dimension of the embeddings. The fused global feature vector is $\mathbf{H} = \sum_{m=1}^{M}\alpha_m\mathbf{H}_m$, where $\mathbf{H}_m$ is the global embedding matrix of modal m.

Finally, a fully connected layer and softmax function are used to classify the fused features into normal or abnormal (including tool wear, spindle misalignment, cutting force deviation) categories. The loss function adopts cross-entropy loss to optimize the model parameters.

3. Results and discussion

3.1. Experimental Setup

Experiments are conducted on a self-built CNC turning test platform with 42CrMo4 steel as the workpiece material. The used sensors are: accelerometer, current sensor, dynamometer. The data acquisition system is built based on LabVIEW, and the model is deployed on an edge computing device (Intel Core i7-12700H, 16GB RAM).

Two datasets are used for experiments. (1) Self-built dataset includes 1000 normal samples and 1500 abnormal samples (500 for each anomaly type), they are collected under different cutting parameters (spindle speed: $1000-2000$rpm, feed rate: $0.1-0.3$ mm/r, cutting depth: 0.2-0.5 mm); (2) Public dataset. The CNC milling tool wear dataset from Nature, which is adapted to CNC turning by adjusting cutting parameters including 800 normal samples and 1200 abnormal samples. The dataset is split into training set (70% , validation set (15%), and test set (15%).

To verify the performance of the proposed framework, the following models are selected for comparison. Traditional machine learning methods like Support Vector Machine (SVM), Random Forest (RF); Deep learning methods include 1D-CNN [12], Long Short-Term Memory (LSTM) [13], Hybrid 1D-CNN+LSTM [14]; GNN-based methods like GCN, GraphSAGE [15], InkStream [16]. All models are trained with the same training set and hyperparameters (batch size $=$ 32, learning rate $=$ 0.001, epochs $=$ 50) to ensure fairness.

3.2. Evaluation Metrics

Four classification metrics are used to evaluate diagnostic accuracy: Accuracy (Acc), Precision (Pre), Recall (Rec), and F1-score [17]. The formulas are as follows:

$$\begin{cases} \text{Acc} = \frac{TP+TN}{TP+TN+FP+FN} \\ \text{Pre} = \frac{TP}{TP+FP} \\ \text{Rec} = \frac{TP}{TP+FN} \\ F1 = \frac{2\times \text{Pre}\times \text{Rec}}{\text{Pre}+\text{Rec}} \end{cases} \tag{14}$$

Where TP (True Positive) is the number of correctly classified abnormal samples, TN (True Negative) is the number of correctly classified normal samples, FP (False Positive) is the number of normal samples misclassified as abnormal, and FN (False Negative) is the number of abnormal samples misclassified as normal. For real-time performance, P99 inference latency (the maximum latency of 99% of samples) is used as the evaluation metric.

3.3. Experimental Results

The classification performance of all models on the self-built dataset and public dataset is shown in Tables 1 and 2, respectively.

It can be seen from Tables 1 and 2 that the proposed framework outperforms all comparison models in all metrics. Compared with the best GNN-based method (InkStream), the proposed framework improves Acc by 1.21% and 1.47% on the self-built and public datasets, respectively. This is because the adaptive graph construction method effectively models the multi-modal feature correlations, and the cross-attention fusion module enhances the robustness to dataset differences. The confusion matrix of the proposed framework on the self-built dataset is shown

Table 1. Classification performance on self-built dataset

Model	Acc (%)	Pre (%)	Rec (%)	F1 (%)
SVM	89.24	88.76	89.01	88.88
RF	91.57	91.23	91.45	91.34
1D-CNN	95.42	95.18	95.30	95.24
LSTM	94.89	94.65	94.78	94.71
1D-CNN+LSTM	95.13	94.90	95.02	94.96
GCN	96.35	96.12	96.24	96.18
GraphSAGE	96.87	96.64	96.76	96.70
InkStream	97.52	97.29	97.41	97.35
Proposed Framework	98.73	98.50	98.51	98.50

Table 2. Classification performance on public dataset

Model	Acc (%)	Pre (%)	Rec (%)	F1 (%)
SVM	88.16	87.68	87.92	87.80
RF	90.43	90.01	90.25	90.13
1D-CNN	94.28	94.05	94.17	94.11
LSTM	93.75	93.52	93.64	93.58
1D-CNN+LSTM	94.01	93.78	93.90	93.84
GCN	95.12	94.89	95.01	94.95
GraphSAGE	95.64	95.41	95.53	95.47
InkStream	96.38	96.15	96.27	96.21
Proposed Framework	97.85	97.62	97.74	97.68

in Fig. 4, which indicates that the framework has high classification accuracy for all anomaly types, especially for early tool wear.

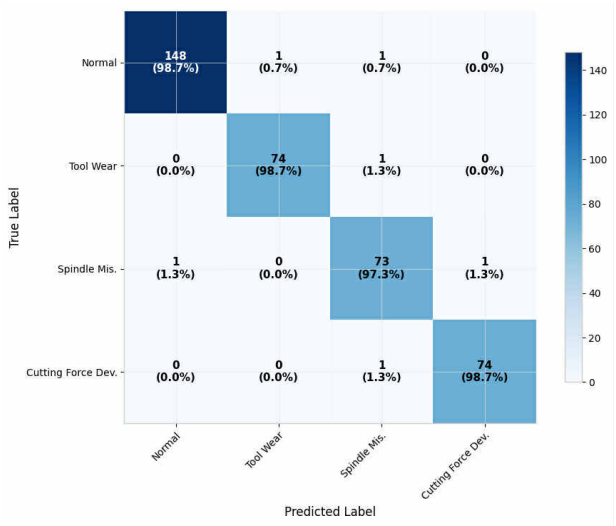

Fig. 4. Confusion matrix of the proposed framework on self-built dataset

The P99 inference latency of all models on the edge computing device is shown in Table 3. The proposed framework achieves a P99 latency of 28.3 ms, which is lower than most comparison models, including InkStream (35.6 ms). This is due to the event-driven incremental update strategy and hierarchical propagation mechanism of EL-GNN, which significantly reduces redundant computations. The real-time performance of the proposed framework meets the requirement of CNC turning equipment (latency < 100 ms).

Table 3. Real-time performance comparison (P99 latency)

Model	P99 Latency (ms)
SVM	18.7
RF	22.4
1D-CNN	45.2
LSTM	68.5
1D-CNN+LSTM	89.3
GCN	76.9
GraphSAGE	62.4
InkStream	35.6
Proposed Framework	28.3

Ablation experiments are conducted to verify the effectiveness of each core module of the proposed framework. The results are shown in Table 4. It can be seen that: (1) Removing the adaptive graph construction module (using fixed graph) reduces Acc by 2.13%, indicating that adaptive graph construction can better model multi-modal correlations; (2) Removing the event-driven update (full graph update) increases P99 latency by 42.4 ms , verifying the importance of event-driven strategy for real-time performance; (3) Removing the cross-attention fusion (equal weight fu-

sion) reduces Acc by 1.56%, proving that dynamic weight assignment enhances model robustness.

4. Conclusions

This paper proposes a multi-modal data fusion and GNN-based framework for real-time anomaly diagnosis in CNC turning processes, addressing the limitations of single-modal data and poor real-time performance of existing methods. The key conclusions are as follows:

(1) The adaptive graph construction method based on mutual information effectively models the non-linear correlations between multi-modal features (vibration, current, cutting force), laying a foundation for accurate anomaly diagnosis.

(2) The event-driven lightweight GNN (EL-GNN) realizes millisecond-level inference by adopting hierarchical propagation and incremental update strategies, balancing diagnostic accuracy and real-time performance.

(3) The cross-attention fusion module dynamically adjusts modal weights, enhancing the model's robustness to noise and dataset differences. The proposed framework achieves 98.73% accuracy and 28.3 ms P99 latency on the self-built dataset, outperforming existing methods.

(4) The proposed framework provides a new solution for intelligent predictive maintenance in CNC turning processes, which can be extended to other CNC machining processes (e.g., milling, drilling) with minor modifications, showing broad industrial application prospects.

Future research will focus on the following aspects: (1) Extend the framework to handle incomplete multi-modal data, improving its adaptability to sensor failure scenarios; (2) Integrate digital twin technology to realize real-time visualization and dynamic adjustment of the diagnostic model; (3) Deploy the framework on edge computing devices with lower power consumption to meet the requirements of industrial Internet of Things (IIoT) applications.

References

[1] P. Kailomsom, P. Nasawat, W. Khunthirat, and W. Phuangpornpitak, (2025) *"A hybrid method based on BWM and TOPSIS-LP model to assess computer numerical control machines"* **Engineering Access** ***11***(1): 108–118. DOI: 10.14456/mijet.2025.9.

[2] I. E. Oiye, P. Poojar, E. Qian, J. T. Vaughan Jr, and S. Geethanath, (2025) *"A standalone and cost-effective MR-compatible computer numerical control machine for simultaneous, multi-parameter mapping"* **Measurement Science and Technology** ***36***(4): 045902. DOI: 10.1088/1361-6501/adbb09.

[3] Z. A. Aldeeb and A. M. Hwas, (2025) *"Design and Implementation of Computer Numerical Controlled Milling Machine for Printed Circuit Board Fabrication"* **African Journal of Advanced Pure and Applied Sciences** ***4***(3): 116–125. DOI: 10.65418/ajapas.v4i3.1343.

[4] J. Singh, A. Singh, H. Singh, and P. Doyon-Poulin, (2025) *"Implementation and evaluation of a smart machine monitoring system under industry 4.0 concept"* **Journal of Industrial Information Integration** ***43***: 100746. DOI: 10.1016/j.jii.2024.100746.

[5] S. Yin, H. Li, A. A. Laghari, L. Teng, T. R. Gadekallu, and A. Almadhor, (2024) *"FLSN-MVO: edge computing and privacy protection based on federated learning Siamese network with multi-verse optimization algorithm for industry 5.0"* **IEEE Open Journal of the Communications Society** ***6***: 3443–3458. DOI: 10.1109/OJCOMS.2024.3520562.

[6] J. Sun, D. Wang, Z. Liu, C. Qiu, H. Liu, G. Sa, and J. Tan, (2025) *"Tool digital twin based on knowledge embedding for precision CNC machine tools: Wear prediction for collaborative multi-tool"* **Journal of Manufacturing Systems** ***80***: 157–175. DOI: 10.1016/j.jmsy.2025.02.021.

[7] A. D. Yewle, L. Mirzayeva, and O. Karakuş, (2025) *"Multi-modal data fusion and deep ensemble learning for accurate crop yield prediction"* **Remote Sensing Applications: Society and Environment** ***38***: 101613. DOI: 10.1016/j.rsase.2025.101613.

[8] X. Wu and A. He, (2025) *"Multimodal information fusion and artificial intelligence approaches for sustainable computing in data centers"* **Pattern Recognition Letters** ***189***: 17–22. DOI: 10.1016/j.patrec.2024.12.006.

[9] Z. Gao, Y. Wang, X. Li, and J. Yao, (2024) *"Twins transformer: rolling bearing fault diagnosis based on cross-attention fusion of time and frequency domain features"* **Measurement Science and Technology** ***35***(9): 096113. DOI: 10.1088/1361-6501/ad53f1.

[10] J. Yu, L. Zhao, S. Yin, and M. Ivanović, (2024) *"News recommendation model based on encoder graph neural network and bat optimization in online social multimedia art education"* **Computer Science and Information Systems** ***21***(3): 989–1012. DOI: 10.2298/CSIS231225025Y.

[11] J. Wu, C. Ni, H. Wang, and J. Chen, (2025) *"Graph neural networks for efficient clock tree synthesis optimization in complex SoC designs"* **Applied and Computational Engineering** ***150***: 101–111. DOI: 10.54254/2755-2721/2025.22281.

Table 4. Ablation experiment results

Model Variant	Acc (%)	P99 Latency (ms)
Full Framework	98.73	28.3
Without adaptive graph construction	96.60	29.1
Without event-driven update	98.67	70.7
Without cross-attention fusion	97.17	27.8

[12] J. Dai, Z. Zhang, Z. Liu, and W. Yuan, (2025) *"Improved Detection of Mixed Pesticide Solution Concentrations Using Deep Learning with Segmented Stride 1D-CNN and Color Rendering Index Features"* **Journal of the ASABE** ***68***(3): 353–363. DOI: 10.13031/ja.16122.

[13] J. Choi, Z. Xiong, and K. Kang, (2025) *"Long short-term memory-based computerized numerical control machining center failure prediction model"* **Mathematics** ***13***(7): 1093. DOI: 10.3390/math13071093.

[14] M. Ahmadzadeh, S. M. Zahrai, and M. Bitaraf, (2025) *"An integrated deep neural network model combining 1D CNN and LSTM for structural health monitoring utilizing multisensor time-series data"* **Structural Health Monitoring** ***24***(1): 447–465. DOI: 10.1177/14759217241239041.

[15] R. Li, H. Shen, Q. Zhang, and H. Duan, (2025) *"An edge-enhanced graphSAGE-based intrusion detection model for the internet of things"* **Cluster Computing** ***28***(5): 309. DOI: 10.1007/s10586-025-05100-x.

[16] D. Wu, Z. Li, and T. Mitra. "Inkstream: Instantaneous GNN Inference on Dynamic Graphs via Incremental Update". In: *2025 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*. IEEE. 2025, 1273–1285. DOI: 10.1109/IPDPS64566.2025.00115.

[17] L. Teng, H. Li, and Y. Si, (2025) *"Neural Tensor Network And Adaptive Graph Convolution For Sports"* **Journal of Applied Science and Engineering** ***29***(6): 1483–1491. DOI: 10.6180/jase.202606_29(6).0015.