Journal of Manufacturing Systems 78 (2025) 294–307

Contents lists available at ScienceDirect

Journal of Manufacturing Systems

journal homepage: www.elsevier.com/locate/jmansys

Technical Paper

Interpretable tool wear monitoring: Architecture with large-scale CNN and adaptive EMD

Yi Sun, Hongliang Song, Hongli Gao *, Jie Li, Shuang Yin

School of Mechanical Engineering, Southwest Jiaotong University, Chengdu 610031, China

ARTICLE INFO

Keywords:
Tool wear monitoring
Architecture interpretable
Large-scale convolutional neural network
Adaptive empirical mode decomposition

ABSTRACT

In manufacturing, tool wear monitoring (TWM) is crucial for ensuring product quality and processing efficiency. Numerous data-driven models based on deep learning have been developed to enhance the accuracy of TWM. However, most models necessitate traditional signal preprocessing and heavily depend on expert knowledge for parameter settings. To address these issues, this study proposes a novel TWM architecture, LCNN-EMD, which combines large-scale convolutional neural networks (LCNN) with adaptive empirical mode decomposition (EMD). By analyzing cutting process mechanisms, it is found that sensor signals can be decoupled into low-frequency noise from the machine's inherent frequencies, mid-frequency features from the tool's chatter, and high-frequency noise from the sensor's quantization. The interpretability of the LCNN-EMD architecture is rooted in its sophisticated capability to analyze and suppress specific frequency features from the input signals. The LCNN adeptly captures low- and mid-frequency features while effectively suppressing high-frequency noise. Concurrently, adaptive EMD dynamically mitigates low-frequency noise, ensuring the preservation of critical mid-frequency features. This dual mechanism not only enhances the accuracy and robustness of LCNN-EMD but also provides it with high interpretability, addressing the limitations of traditional models that depend on extensive expert knowledge and static signal preprocessing. Finally, we conduct tool life experiments and compare the LCNN-EMD with mainstream models for a comprehensive evaluation. Experimental results indicate that LCNN-EMD consistently outperforms the comparative models in terms of accuracy and stability.

1. Introduction

During the tool cutting process, chatter between the tool's back cutting surface and the material being removed is the main source of wear features and is essential for tool wear monitoring (TWM) [1,2]. Chatter signals, reflecting the dynamic interaction between the tool and the material, encompass a frequency range of 500–1500 Hz [3]. However, the analysis of machining signals faces challenges from the machine's inherent frequencies and sensor's quantization noise [4]. Structural vibrations of computer numerical control (CNC) machine, together with resonances from the bedframe and supports, significantly contribute to low-frequency noise, predominantly concentrated within the 0–500 Hz frequency range [5]. Effective identification and removal of low-frequency noise during signal processing are essential to prevent interference with the analysis of tool wear features. In addition, sensors introduce high-frequency noise, typically originating from the electrical features of the sensor, internal interference within the data acquisition system, and electromagnetic interference from the environment, concentrated within the 1000–2000Hz frequency range [1]. High-frequency noise can also significantly degrade the quality of signal analysis, especially when attempting to extract tool wear features from the signal. The impact of noise on machining signal analysis necessitates the adoption of effective noise-filtering and signal-processing strategies [2,6,7].

Mainstream architectures for TWM commonly depend on expert selection of signal features and adjustment of classification algorithms [8,9]. Researchers are developing increasingly sophisticated architectures to reduce reliance on expert experience and augment the level of automation in data processing. Habaibeh et al. [10] proposed an autonomous sensor and signal processing strategy that identifies the most sensitive sensors for machining fault detection, thereby reducing the time and cost associated with these detections. Guo and Yang [11,12], as well as other researchers, have employed machine learning methods to predict tool wear. Specifically, Guo et al. [11] utilized a multi-scale convolutional attention network to integrate multi-source data and enhance prediction performance. Yang et al. [12] predicted

* Corresponding author.
E-mail address: hongli_gao@swjtu.edu.cn (H. Gao).

https://doi.org/10.1016/j.jmsy.2024.12.001
Received 5 July 2024; Received in revised form 8 November 2024; Accepted 2 December 2024
Available online 18 December 2024

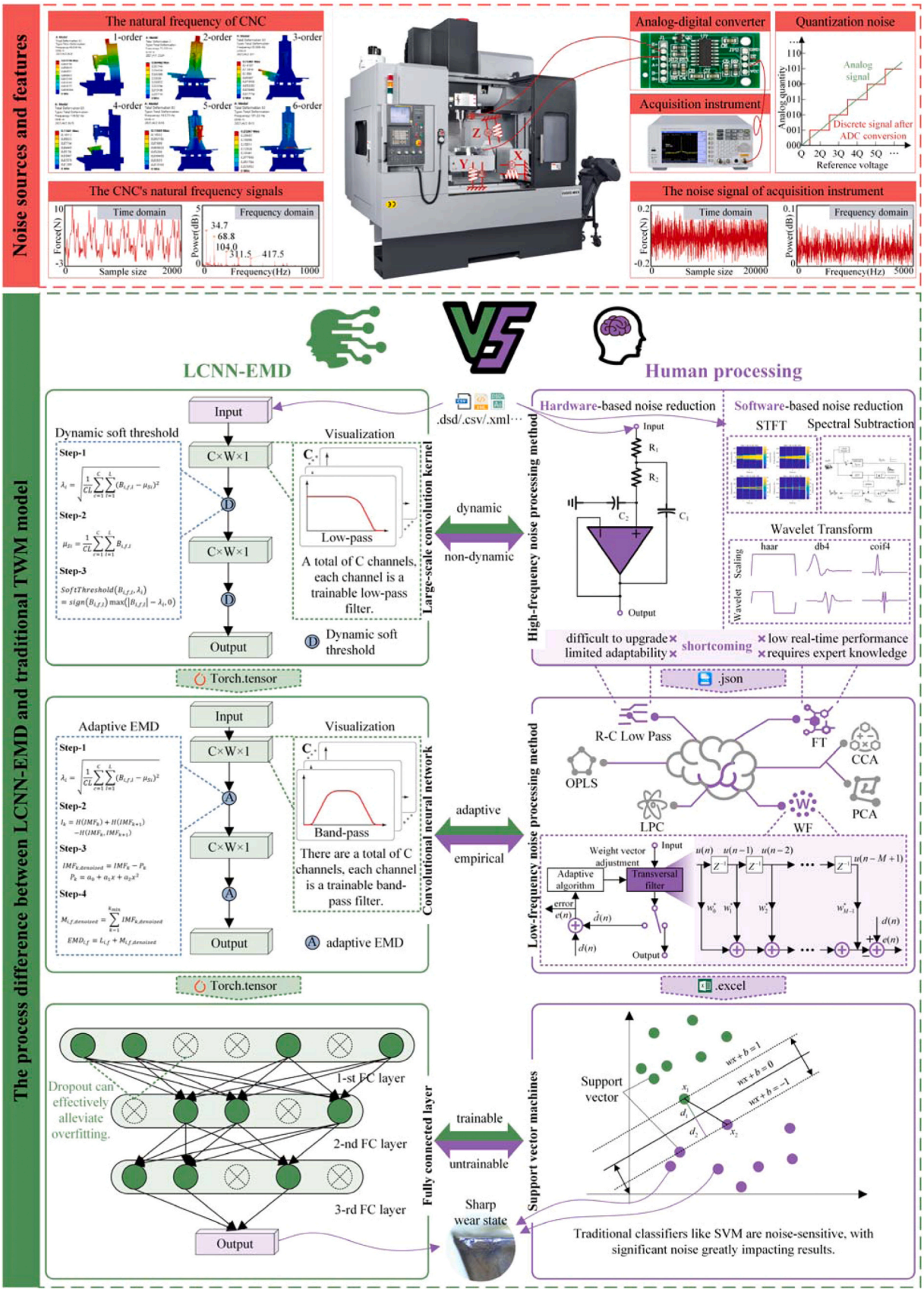

Fig. 1. The sources of noise signals in TWM, as well as the architecture of LCNN-EMD and mainstream models.

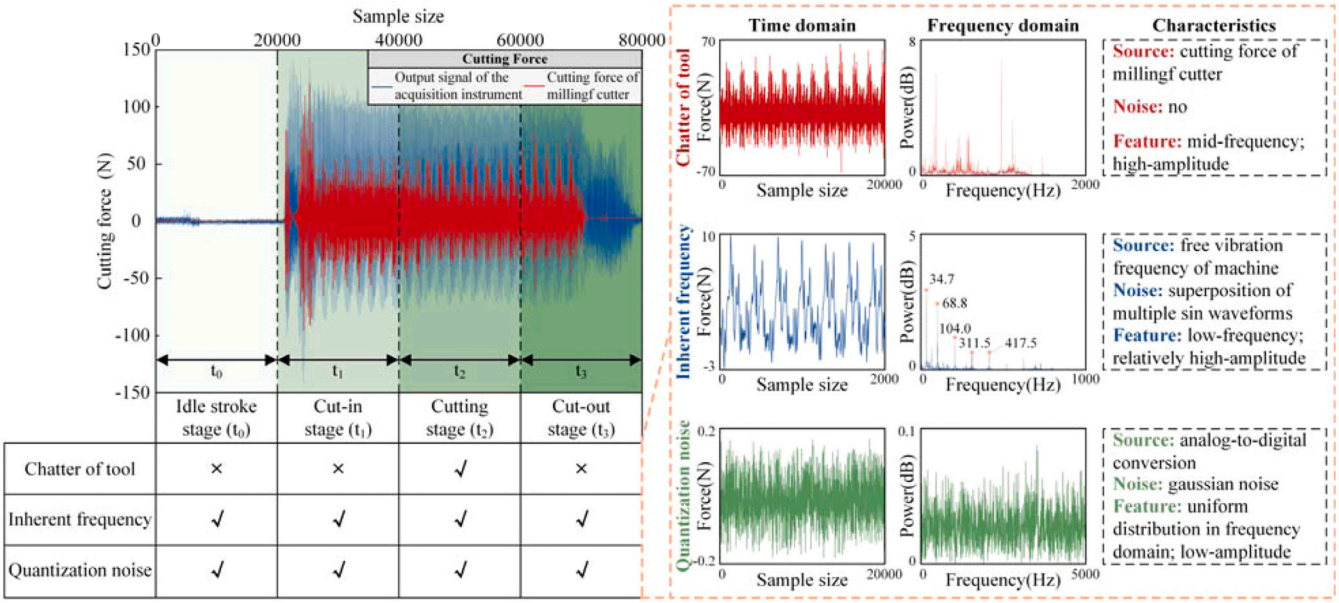

Fig. 2. The stages and characteristics of features and noise in TWM.

the milling force of a single-edged tool based on the theoretical formula of undeformed chip thickness, generating cutting force samples through time shifting and Gaussian distribution to address the considerable dimensionality of the sample set. Building on these studies, Traini et al. [13] developed a data-driven architecture that merges preprocessing and predictive models, utilizing various machine learning algorithms and comparing them to select the optimal results. However, these methods necessitate extensive data preprocessing and computational resources, and carry the risk of overfitting. Zhou et al. [14] proposed a multi-sensor global feature extraction method that selects features from multi-domain data and employs an improved genetic algorithm to find optimal parameters in a bi-objective optimization model, thereby enhancing prediction accuracy. Yet, the convergence speed and stability of the method present unavoidable challenges. It is evident that mainstream TWM architectures comprise multiple distinct yet interdependent components, each fulfilling a specialized role within the overall monitoring framework to ensure precise wear monitoring. Due to the asynchronous nature of these components in terms of data processing speed and capacity, data flow within the architecture is susceptible to delays and blockages, significantly reducing the monitoring system's response speed and efficiency.

The high-frequency noise introduced during the sensor analog-to-digital conversion process in milling poses a significant challenge to the accuracy of TWM [15]. To address this challenge, researchers have developed various noise reduction techniques, including, sparse decomposition [16], wavelet packet decomposition [17], and a combination of acoustic analysis, wavelet denoising, and support vector machines [18]. Zhu et al. [16] focused on processing cutting force noise, extracting useful signals from a complex noise environment through sparse decomposition, yet the method is computationally intensive and the noise reduction efficacy depends on parameter selection. Tran et al. [17] optimized high-frequency noise through wavelet packet decomposition combined with multi-sensor data, but did not compare their method with existing mainstream algorithms. Zhou et al. [18] utilized acoustic analysis and wavelet denoising, in conjunction with support vector machines, achieving an 85 % monitoring accuracy; however, for high-precision manufacturing requirements, the algorithm's accuracy is still inadequate. Furthermore, Li et al. [19] integrated a denoising autoencoder with a deep belief network (DBN) to enhance the robustness and denoising capabilities of their model through sparse representation techniques. However, the computational overhead associated with this model is substantial, making it more applicable for scenarios where real-time processing demands are not stringent. The generalizability of the aforementioned research to a wider array of sensor types and a broader spectrum of noise frequencies necessitates further investigation and validation.

In addition to high-frequency noise, low-frequency noise induced by the machine's inherent frequencies presents a significant challenge in TWM. Garg et al. [20] proposed a simplified approach based on the finite element method to estimate the inherent frequencies and modal shapes of machine structures. However, the simplified model cannot accurately capture the dynamic responses of complex machine structures, leading to imprecise frequency estimations. Additionally, Jia et al. [21] proposed an fault diagnosis model leveraging the multi-parameter optimization capabilities of SDAE to enhance noise suppression within low-frequency domains. However, the model requires intricate parameter configuration, reliant on optimization algorithms to determine the optimal structure, which increases computational complexity and constrain feasibility for real-time applications. Li et al. [22] explored the low-frequency modes of machine structures and their impact on surface formation. Nonetheless, this research necessitates substantial computational resources and expertise. Zhang et al. [23] denoised original signals using wavelet analysis and employed a neuro-fuzzy network to predict tool wear, yet the predictive performance is contingent upon the quality of the training data. Sabareeswaran et al. [24] optimized the machining fixture layout using Genetic Algorithms and Particle Swarm Optimization techniques to maximize the difference between the workpiece's inherent frequency and the tool excitation frequency. The aforementioned research primarily focuses on improving surface quality, but does not adequately consider the potential impact of the machine's inherent frequency on TWM.

Given the ongoing challenges in TWM, the development of an efficient architecture is imperative. To address these issues, we propose a novel architecture, LCNN-EMD, integrating a large-scale convolutional neural network (LCNN) with adaptive empirical mode decomposition (EMD). The detailed steps of LCNN-EMD and its comparison with mainstream models are shown in Fig. 1 [13,18,24–26]. This architecture capitalizes on the robust feature-learning capabilities of LCNN and the advanced non-linear and non-stationary data processing strengths of EMD, significantly enhancing data flow consistency and processing efficiency while diminishing reliance on empirical knowledge. Innovations are as follows:

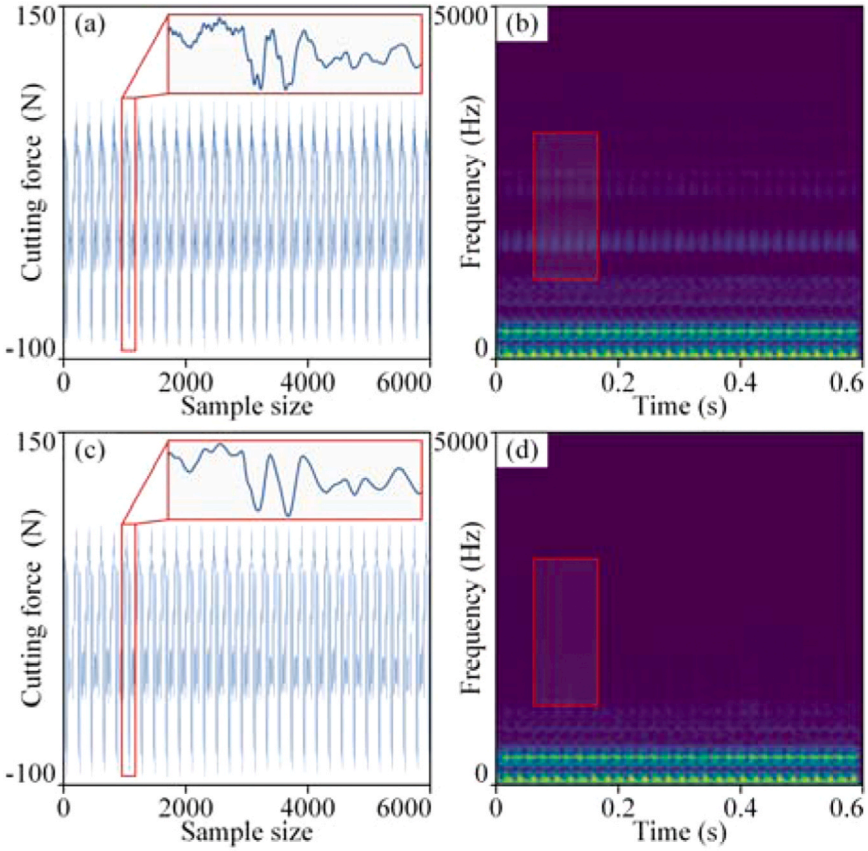

Fig. 3. Time and frequency-domain representations of the first convolution layer outputs for large-scale kernels and small-scale kernels. (a) and (b) are the time and frequency-domain representations of the small-scale kernel's outputs; (c) and (d) are the time and frequency-domain representations of the large-scale kernel's outputs, respectively.

- Addressing the limitations of traditional monitoring architectures: To address the significant rigidity and excessive dependence on prior knowledge of mainstream algorithms, we have developed the LCNN-EMD. LCNN-EMD ensures consistent data flow, minimizes dependency on expert knowledge, and facilitates the progression of TWM towards automation and intelligence.
- Suppressing high-frequency noise: To address the high-frequency noise introduced by sensor quantization, a large-scale convolution kernel based on dynamic soft thresholds is proposed. This kernel, with its wide-field of view, adapts to a higher frequency range and suppresses high-frequency noise from various types of sensors through dynamic threshold.
- Suppressing low-frequency noise: To address the low-frequency noise induced by the machine's inherent frequencies, an adaptive EMD is proposed. The adaptive EMD extracts essential features related to the tool's wear features, effectively identifying and suppressing low-frequency noise induced.

The remainder of this study is organized as follows. The "INTERPRETABILITY OF LCNN-EMD" s section elucidates the functions of each module within the architecture and explains how they interact to enhance the interpretability of LCNN-EMD. The "RESULTS AND ANALYSIS" section provides a detailed description of the dataset, comparative models, and performance analysis. The "ABLATION EXPERIMENT" section explores, through visualization experiments, the specific roles of the large-scale convolutional neural network and adaptive empirical mode decomposition in improving the performance of LCNN-EMD. Finally, the "CONCLUSION" section presents the contributions of this study and outlines directions for future research.

2. Interpretability of LCNN-EMD

2.1. The analysis of tool signals

During the machining process, the signals collected by sensors are complex and information-rich, consisting primarily of three key components: low-frequency noise from machine's inherent frequencies [5], high-frequency noise from sensor's quantization noise [1], and mid-frequency features from the tool's chatter [3], which contain wear features. Fig. 2 shows the stages of occurrence and characteristics of these components, providing a basis for comprehensive assessment of the TWM.

Machine's inherent frequency: low-frequency noise generated by the machine's own vibration modes. These vibrations originate from imbalances in the machine, resonance among components, and wear of parts. As shown by the blue lines on the right side of Fig. 2, their consist of multiple superimposed sine signals, with energy primarily concentrated below 500 Hz (the first and second inherent frequencies are 34.7 Hz and 68.8 Hz, respectively, with the highest amplitude). Identification and filtering of this low-frequency noise are crucial for enhancing the signal-to-noise ratio.

Sensor's quantization noise: high-frequency noise primarily from the electrical properties of the sensors themselves, including resolution loss during the analog-to-digital conversion process and electromagnetic interference from electronic devices. As shown by the green lines on the right side of Fig. 2, their energy amplitude is relatively low and evenly distributed within the 0–2000 Hz frequency band. Quantization noise does not carry direct information about the machining process, yet its presence significantly impacts the accuracy of signal analysis.

Tool's wear features: critically, chatter signals are directly related to

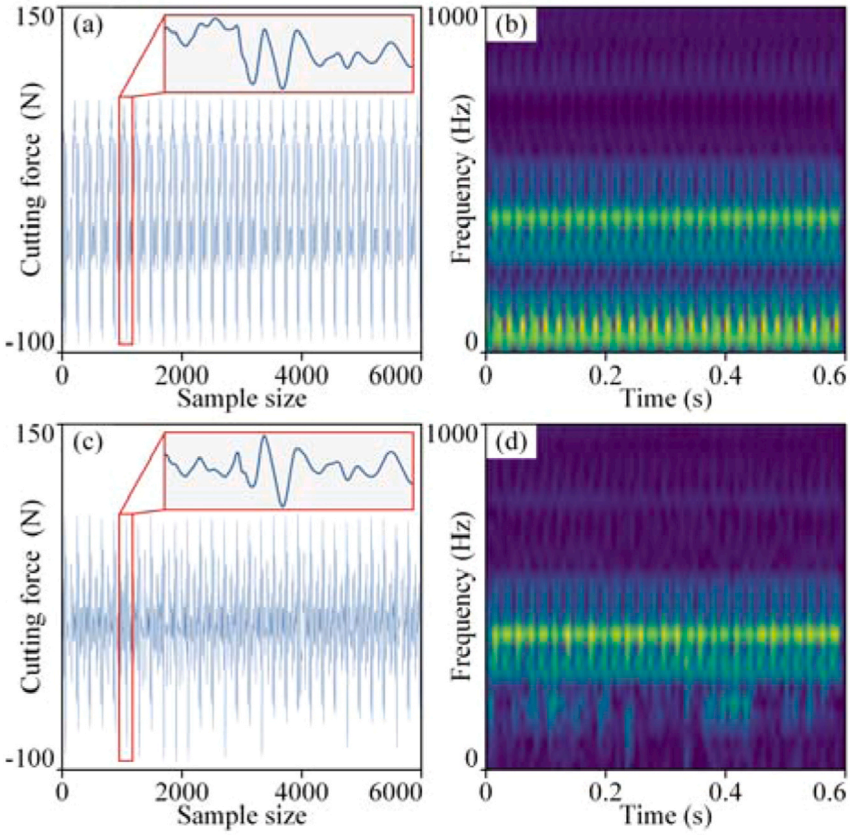

Fig. 4. Time and frequency-domain representations of large-scale kernel and adaptive EMD outputs. (a) and (b) are the time and frequency-domain outputs of the large-scale kernel; (c) and (d) are time and frequency-domain representations after adaptive EMD processing, respectively.

the interaction between the tool's back cutting surface and the material being removed. These chatter signals contain critical information about wear features, changes in cutting forces, and cutting parameters. As shown by the red lines on the right side of Fig. 2, their energy amplitude is high, with the highest energy band near 1000 Hz. Analyzing these signals enables real-time tool wear monitoring, facilitating maintenance decisions.

2.2. The interpretability of LCNN

While CNNs gain widespread adoption in TWM, existing research primarily focuses on parameter optimization and model design, with relatively scant attention paid to the specific influence of kernel size [27]. Kernel size is crucial for the ability of CNNs to extract features, especially in processing signals with complex frequency components. To address this research gap in TWM, this section explores the impact of kernel size on the CNN's feature extraction capabilities.

To investigate this, a two-layer CNN is constructed, where the first layer utilizes two different kernel sizes, a large-scale kernel (128 ×1) and a small-scale kernel (3 ×1), specifically to analyze the effect of kernel size on feature extraction capabilities. Both models are trained 1000 times. By visualizing the output of the first convolutional layer, the impact of kernel size on machining signals is analyzed, as shown in Fig. 3.

Firstly, time-domain features are analyzed. By comparing the localized details of time-domain signals in Fig. 3(a) and (c) (enlarged section within the red box), it is evident that the convolutional layer composed of larger kernels produces smoother waveforms. This contributes to reducing the number of extreme points near signal inflection points, thereby alleviating modal aliasing in subsequent EMD processing. Secondly, frequency-domain features are analyzed. Comparing Fig. 3(b) and (d), it is observed that the left red box in Fig. 3(b) exhibits light blue features, while no frequency-domain features appear in the left red box of Fig. 3(d). Additionally, there is no significant difference in the low-frequency noise below 500 Hz. This indicates that the convolutional layer with large-scale kernels can identify and suppress high-frequency noise around 1500 Hz, while the convolutional layer composed of small-scale kernels is unable to effectively process high-frequency noise.

The aforementioned analysis indicates that large-scale kernels, owing to their larger receptive field area, can effectively capture low- and mid-frequency components, suppress high-frequency noise caused by sensor noise, and consequently enhance feature extraction significantly.

2.3. The interpretability of adaptive EMD

Traditional EMD methods rely on manually selecting intrinsic mode functions (IMFs) to suppress noise, a process that is both time-consuming and experience-dependent [28]. To address this challenge, this section explores the feasibility of adaptive EMD.

Adaptive EMD dynamically adjusts the decomposition process, focusing on extracting features related to tool wear, and effectively suppressing low-frequency noise introduced by the machine's inherent frequencies. Adaptive EMD decomposes time series signals into multiple IMFs, each capturing different frequency components of the signal. Subsequently, the mutual information entropy between IMF is calculated sequentially. The two IMFs with the highest information entropy serve as the demarcation between low- and mid-frequency features of the signal segment. Then, the noise of the low-frequency component is suppressed by the exponential weighted moving average filter, and the trend of the residual component is suppressed by the second-order polynomial. Finally, the processed signals are reassembled, achieving

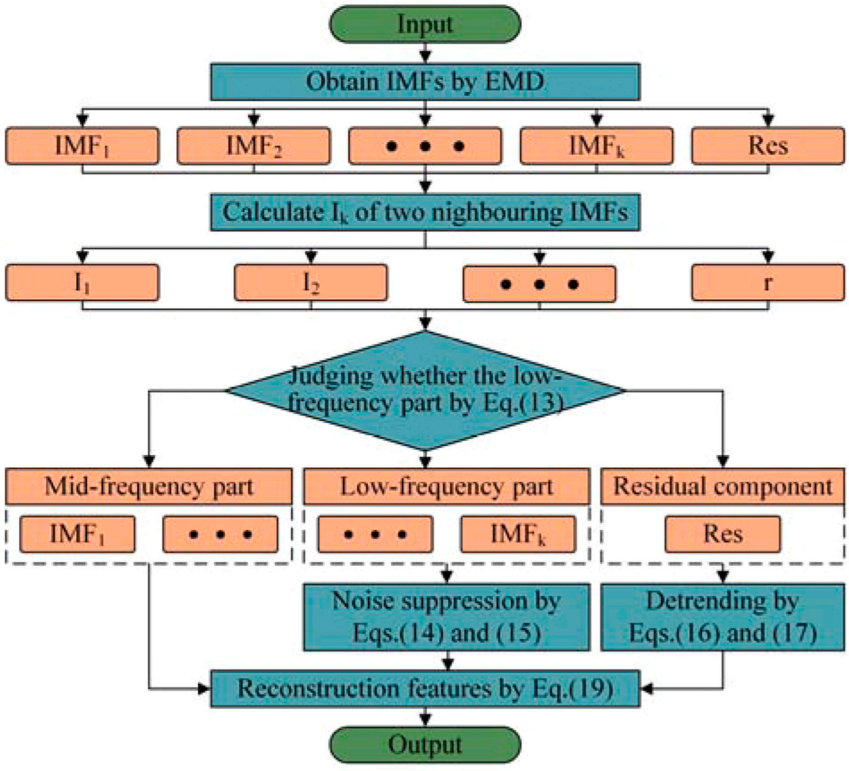

Fig. 5. The process of adaptive EMD.

the suppression of low-frequency noise. Unlike predefined filter parameters, adaptive EMD dynamically adjusts based on the signal's inherent characteristics.

The time-frequency domain characteristics of the signal after adaptive EMD processing are shown in Fig. 4. By comparing signals before and after processing with adaptive EMD, the suppression effect of low-frequency noise can be clearly observed. In time-domain feature analysis, comparing Fig. 3(c) and Fig. 4(a) shows that adaptive EMD effectively eliminates signal trends, making the signals more stable and easier to analyze. In frequency-domain feature analysis, comparing Fig. 3(d) and Fig. 4(b) reveals that noise below 500 Hz is significantly suppressed, while mid-frequency features above 500 Hz are preserved. Consequently, adaptive EMD dynamically adjusts processing parameters based on the signal's actual characteristics, not only improving the efficiency of low-frequency noise suppression but also retaining essential mid-frequency features, thereby enhancing the precision of TWM.

3. Proposed method

3.1. LCNN

LCNN extracts features from local regions of the input using different convolutional kernels, with each kernel acting as an independent feature filter. The convolution process is as follows:

$$Z_{i,f,l} = \sum_{c=0}^{C-1} \sum_{j=0}^{k-1} X_{i,c,l\times S+j} * W_{f,c} + b_f \tag{1}$$

where, $X_{(N,C,L)}$ is the training dataset, $X_{i,f,l}$ is the output value of the f-th convolutional kernel at position l for the i-th sample. N is the number of training samples ($i \in N$), C is the number of channels, and L is to the length of each training sample ($l \in L$). F is the number of convolution kernels ($f \in F$), k is the size of the convolution kernels, S is the step length, and $*$ is a one-dimensional convolution operation. Additionally, $Z_{i,f,l}$ is the output value of the f-th convolutional kernel at position l for the i-th sample, $W_{f,c}$ is the weights with the c-th channel of the f-th convolutional kernel, and b_f is the bias term for the f-th convolutional kernel.

Then, max pooling is employed for subsampling, aiming to reduce the spatial dimensions of the signals while extracting significant features and mitigating overfitting. The subsampling procedure is as follows:

$$P_{i,f,l} = \max_{m=0}^{D-1} Z_{i,f,l\times S_{pool}+m} \tag{2}$$

$$L_{pool\_out} = \left[\frac{L_{pool\_in} - D}{S_{pool}}\right] + 1 \tag{3}$$

where, max is the operation of selecting the maximum value within the pooling window, D is the dimension of the pooling window, S_{pool} is the stride, and $[\cdot]$ is the floor function. L_{pool_in} and L_{pool_out} are the input and output lengths of the max pooling layer, respectively.

The sigmoid and tanh can lead to the vanishing gradient, whereas the ReLU can render a portion of neurons inactive. In LCNN-EMD, the implementation of Leaky ReLU introduces non-linearity, enhancing convergence speed while extracting sparse features. The process of Leaky ReLU is as follows:

$$A_{i,f,l} = \begin{cases} P_{i,f,l} & if P_{i,f,l} > 0 \\ \alpha \times P_{i,f,l} & if P_{i,f,l} \le 0 \end{cases} \tag{4}$$

where, α is the leakage coefficient, employed to control the small gradient value for negative input values.

LCNN-EMD employs batch normalization to normalize the activated features, ensuring that the mean of each feature is close to 0 and the variance is close to 1. This process helps reduce the sensitivity of LCNN-EMD to amplitude variations in the training dataset $X_{(N,C,L)}$, thereby enhancing the model's generalization capability across different cutting parameters. The procedure for batch normalization is described by the following equation:

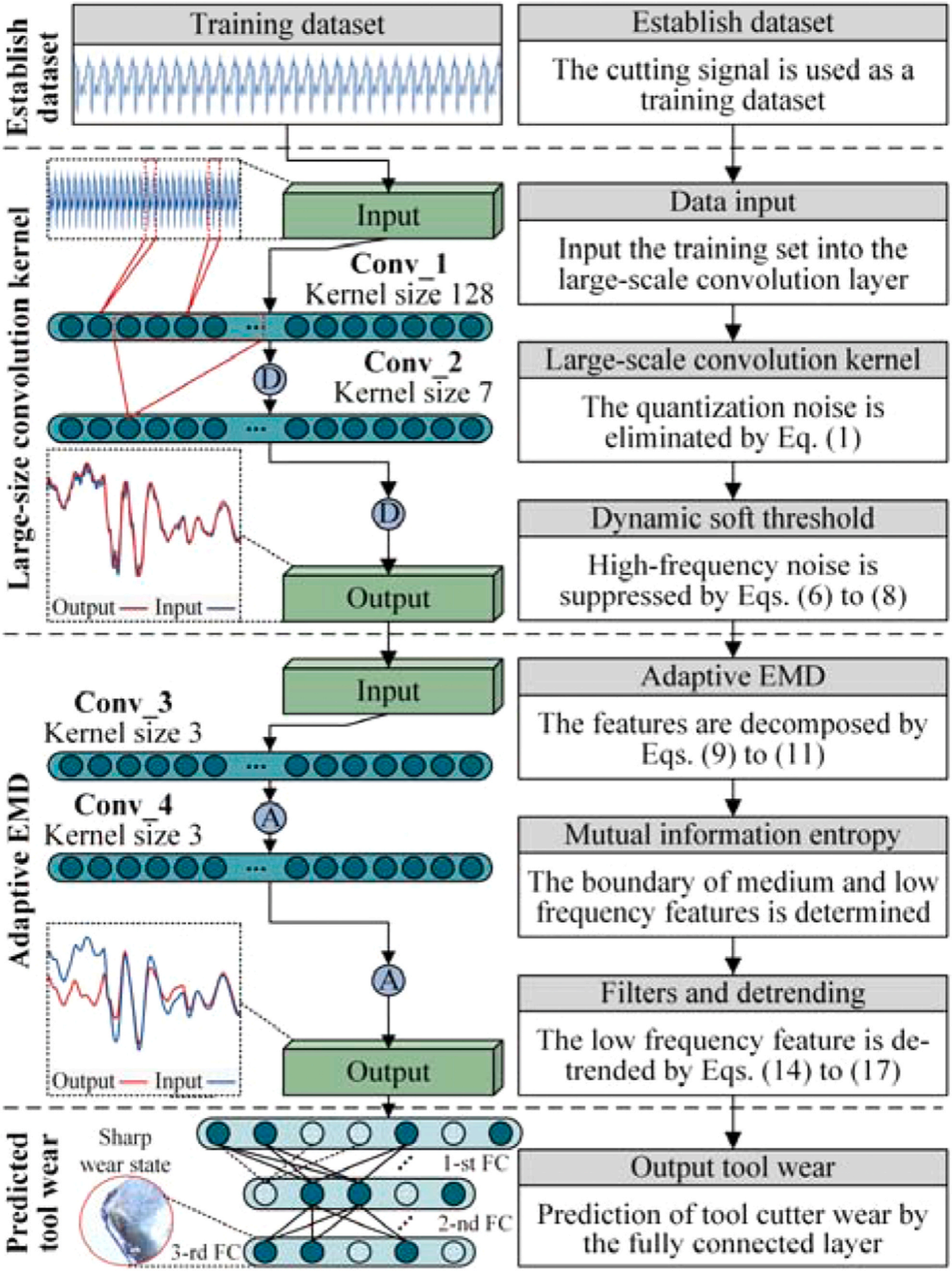

Fig. 6. The main components and processes of LCNN-EMD.

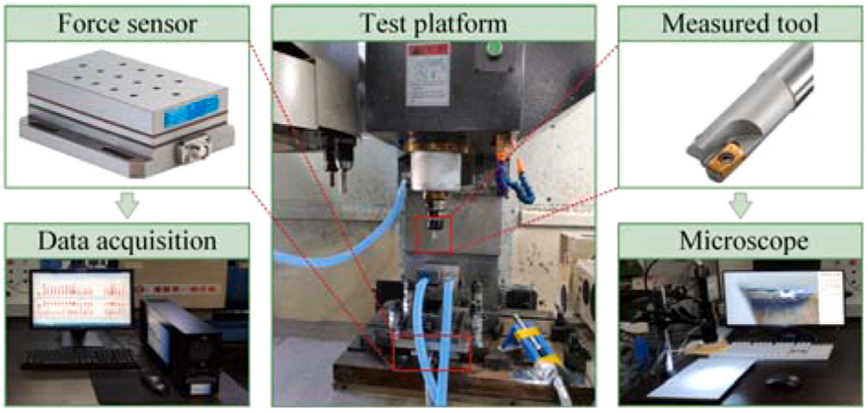

Fig. 7. Tool wear test platform.

Table 1
Experimental parameters of tool wear.

Parameters	Model/Value
CNC machine	Vertical Machining Center VMC850
Measured cutting tool	Duracarb APMT1135 Carbide Tool
Tool holder	INTINIZEL−300R−10-1T-C10-120L
Cutting force sensor	KISTLER−9257B
Signal acquisition instrument	DEWETRON DEWE3021
Charge amplifier	KISTLER−5015A
Tool wear measuring device	BC4800 Industrial Microscope
Test material	Steel 45
Cooling condition	Dry Cutting

$$B_{i,f,l} = \gamma_f \frac{A_{i,f,l} - \mu_f}{\sqrt{\sigma_f^2 + \epsilon}} + \beta_f \tag{5}$$

where, γ_f is the scaling parameter and β_f is the shifting parameter, both of which are used to adjust the normalized values. μ_f and σ_f^2 are the mean and variance of the f-th convolutional kernel calculated during training, respectively. ϵ is a small constant added for numerical stability.

In LCNN-EMD, while max pooling, Leaky ReLU, and batch normalization effectively optimize the scale and distribution of features, these methods primarily focus on feature scaling and distribution without directly addressing noise-related issues. To mitigate this, a dynamic soft thresholding is proposed. Unlike static thresholding methods, the dynamic soft thresholding approach derives its threshold, λ_i, from the statistical distribution of features within each sample. Consequently, this method is insensitive to the amplitude variations of the sample while maintaining high sensitivity to high-frequency noise. The process is described by the following equations:

$$\lambda_i = \sqrt{\frac{1}{CL}\sum_{c=1}^{C}\sum_{l=1}^{L}(B_{i,f,l} - \mu_{Si})^2} \tag{6}$$

$$\mu_{Si} = \frac{1}{CL}\sum_{c=1}^{C}\sum_{l=1}^{L} B_{i,f,l} \tag{7}$$

$$SoftThreshold(B_{i,f,l}, \lambda_i) = sign(B_{i,f,l})\max(|B_{i,f,l}| - \lambda_i, 0) \tag{8}$$

where, $sign(\cdot)$ is the sign function and $|\cdot|$ is the absolute value function.

3.2. Adaptive EMD

Adaptive EMD provides an effective means in TWM for processing and analyzing complex machining signals. To identify between low- and mid-frequency components, EMD is rewritten as:

$$EMD_{i,f,l} = L_{i,f,l} + M_{i,f,l} \tag{9}$$

$$L_{i,f,l} = \sum_{k=k_{\min}}^{K-1} IMF_k + r \tag{10}$$

$$M_{i,f,l} = \sum_{k=1}^{k_{\min}} IMF_k \tag{11}$$

where $M_{i,f,l}$ is the combination of IMFs of the mid-frequency components; $L_{i,f,l}$ is the combination of IMFs of the low-frequency components and residual component; $k_{\min}$ is the IMF's label at the boundary between mid- and low-frequency components; IMF_k is the IMF with different frequencies obtained by EMD; r is the residual component, which is the average trend of the input. Mutual information entropy is used to measure the statistical dependence between two random variables [29], and its expression is:

$$\begin{aligned} I_k &= H(IMF_k) + H(IMF_{k+1}) - H(IMF_k, IMF_{k+1}) \\ &= \sum_{k=1}^{K-1} p(IMF_k, IMF_{k+1}) \log\left(\frac{p(IMF_k, IMF_{k+1})}{p(IMF_k)p(IMF_{k+1})}\right) \end{aligned} \tag{12}$$

where $p(IMF_k)$ and $p(IMF_{k+1})$ are the probability distributions of IMF_k and IMF_{k+1}, respectively. $p(IMF_k, IMF_{k+1})$ is a joint probability distribution.

The mutual information entropy relationship between the two IMFs is used to identify the boundary between the mid- and low-frequency components. The following search objective function is obtained:

$$k_{\min} = \min_{0 \le k < K} [I_k] \tag{13}$$

where $\min_{0 \le k < K}[\cdot]$ is a function that returns the index of the minimum

Table 2
Details of the tool wear dataset.

Field	Number	Sampling points	Spindle speed/ (r/min)	Feed rate / (mm/s)	Back cutting depth/ (mm)	Workpiece materials	Workpiece size/ (mm^3)
Training dataset	600	4096	4500	6.67	1.5	45 steel	100 × 100 × 200
Test dataset	600	4096	4000	5	1	45 steel	100 × 100 × 200

Frequency domain visualization of one-dimensional large-scale convolution kernel weights. From a to p, each represents one of the 16 channels.

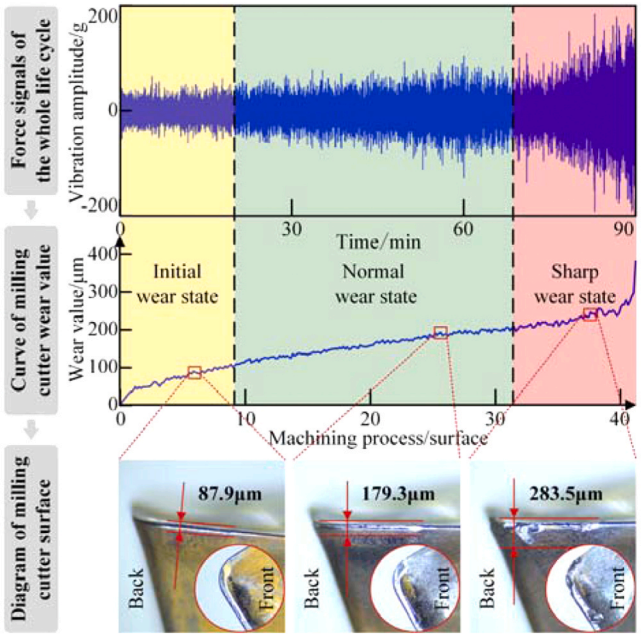

Fig. 8. Signals and surface profiles of the tool at different wear stages.

value from 0 to K.

An exponential weighted moving average filter is applied to the low-frequency components for smoothing, with the smoothing coefficient α_l dynamically adjusted. The process is as follows:

$$IMF_{k,denoised} = \sum_{j=2}^{L} \alpha_l \times IMF_{k,j} + (1-\alpha_l) \times IMF_{k,j-1} \tag{14}$$

$$\alpha_l = \sum_{j=2}^{L} \frac{IMF_{k,j} - IMF_{k,j-1}}{\max(IMF_k)} \tag{15}$$

The trend of the residual component is removed using second-order polynomial fitting. The process is as follows:

$$r_{denoised} = r - \Delta r \tag{16}$$

$$\Delta r = a_0 + a_1 x + a_2 x^2 \tag{17}$$

Finally, the features are reconstructed using denoised low-frequency components $L_{i,f,l,denoised}$, residual component $r_{denoised}$ and unchanged mid-frequency components $M_{i,f,l}$. The process is as follows:

$$L_{i,f,l,denoised} = \sum_{k=k_{\min}}^{K-1} IMF_{k,denoised} + r_{denoised} \tag{18}$$

$$EMD_{i,f} = L_{i,f,l,denoised} + M_{i,f,l} \tag{19}$$

Adaptive EMD ensures effective noise reduction of low-frequency components while preserving the integrity of mid-frequency components, resulting in a reconstructed signal of higher quality. The process is shown in Fig. 5.

Table 3
Overview of comparison models.

Model	Supervised model		Semi-supervised model
	Sequence	Spatial	
RNN [32]	√	×	×
CNN [31]	×	√	×
U2PL-TCM [30]	×	√	×
DBN-EMD [33]	×	√	×
RNN-EMD [32]	√	×	√
LCNN-EMD	×	√	×

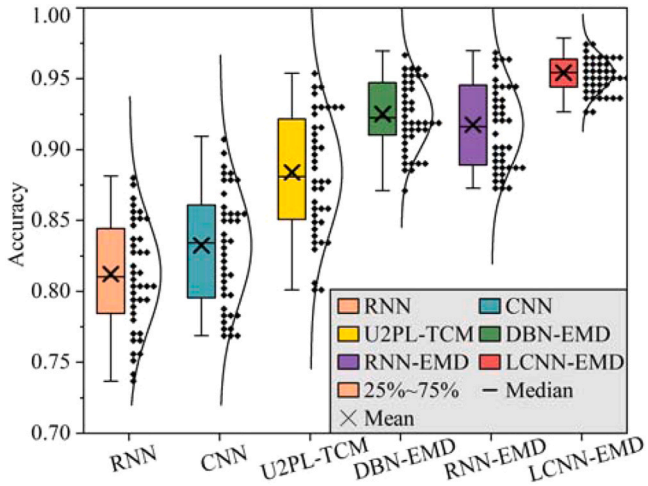

Fig. 9. The accuracy of each model.

3.3. Process of LCNN-EMD

The LCNN-EMD is an architecture specifically designed for TWM, aimed at automatically identifying and suppressing noise in machining signals by integrating multiple key components while ensuring efficient data flow within the architecture. Fig. 6 shows the main components and processes of LCNN-EMD.

4. Results and analysis

4.1. Establishment of the dataset

The LCNN-EMD is validated using signals obtained from a cutting force sensor (KISTLER-9257B). Signal acquisition is performed on a vertical machining center (VMC850) with three channels at a sampling frequency of 10 kHz, as depicted in Fig. 7. The wear values of the tool edges are measured using a GP-300C industrial microscope. The detailed experimental setup is shown in Table 1. Considering the properties of the material and the performance of the machine tool, a strategy of "low cutting depth, high spindle speed, and high feed rate" is adopted, with specifics provided in Table 2. The LCNN-EMD model is trained and tested on a 64-bit Windows 10 operating system, utilizing PyTorch 2.0.0 as the primary tool. The hardware specifications include an Intel® Core™ i7 14700KF processor, a GeForce RTX 2080 Ti graphics card, and 48 GB of 6000 MHz DDR5 memory.

In accordance with the GB/T 16460–2016 standard, the tool is considered to have failed when the uniform wear thickness exceeds 350 μm, at which point the experiment is terminated. Consequently, the wear states of the tool during failure are categorized into three types: initial wear state (0–100 μm), normal wear state (100–200 μm), and severe wear state (greater than 200 μm). The profiles of the front and back face of the tool at various wear stages are shown in Fig. 8.

4.2. Comparison model

Based on the labeling of samples, the field of tool wear monitoring is primarily divided into supervised and semi-supervised models. Additionally, supervised models can be further categorized into sequence and spatial models. To comprehensively evaluate the performance of LCNN-EMD, a series of representative models were selected for comparison, including recurrent neural network (RNN), CNN, unreliable pseudo-label network (U2PL-TCM), DBN-EMD, and RNN-EMD, as shown in Table 3 [30–33]. Notably, both DBN-EMD and RNN-EMD employ

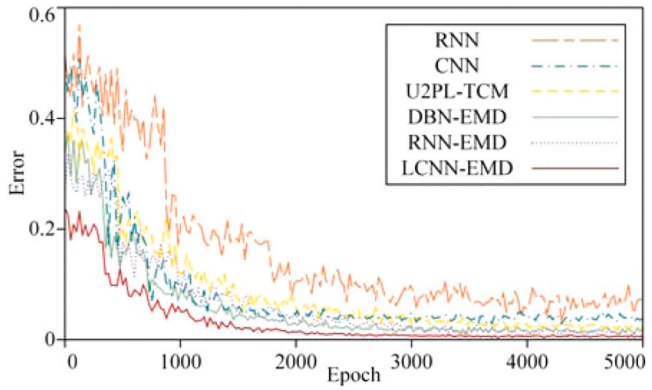

Fig. 10. The training losses of each model.

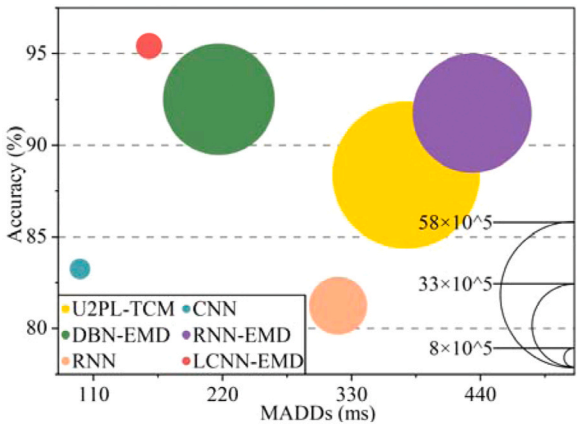

Fig. 11. The timeliness of each model.

adaptive EMD, which enhances their capability to handle non-stationary and non-linear characteristics. By evaluating the performance of these models in practical TWM tasks, a comprehensive understanding of each model's strengths, weaknesses, and applicable scenarios is achieved.

4.3. Results analysis

To thoroughly evaluate the performance of various models, this study designed 40 rounds of repetitive experiments, with results displayed in Fig. 9 using Boxplots. In each round, every model was initialized with random parameters and underwent 5000 epochs.

To begin with, the accuracy distribution of each model is analyzed. The interquartile ranges for RNN, CNN, U2PL-TCM, and RNN-EMD are 0.063, 0.069, 0.073, and 0.057, respectively. These values indicate a wide distribution of accuracy across repeated experiments, suggesting lower stability for these models. Specifically, due to its complex architecture, U2PL-TCM does not fully realize its potential on limited datasets, resulting in the most dispersed accuracy distribution. Conversely, LCNN-EMD shows the most compact distribution, with an interquartile range of 0.021, reflecting superior stability.

Following the distribution analysis, the mean accuracy of each model is analyzed. It is evident that integrating adaptive EMD with deep learning markedly improves the accuracy of models such as DBN-EMD, RNN-EMD, and LCNN-EMD. This integrated approach effectively enhances the models' capabilities to process complex signals, resulting in superior performance in TWM. Among these, LCNN-EMD stands out with the highest mean accuracy of 0.955. This remarkable performance is largely due to LCNN's proficiency in capturing local signal features, coupled with the EMD method's adaptive decomposition, which efficiently extracts essential wear characteristics.

4.4. Stability analysis

Stability is a critical indicator for evaluating model performance in TWM. This section evaluates the models' performance in terms of convergence efficiency and stability. Fig. 10 shows the losses throughout the training process for each model.

During the initial stages of training, both RNN and CNN displayed high loss values, averaging around 0.5. As training epochs increased, CNN's loss gradually decreased and stabilized, converging at 0.21. In contrast, RNN's loss remained relatively unstable, even after 5000 epochs, fluctuating around 0.37. This difference highlights CNN's more stable convergence, attributed to its straightforward feature extraction mechanism and reduced sensitivity to gradient-related. In contrast, the RNN, while attempting to capture the rapid local changes in cutting force signals induced by tool wear, compromised its overall temporal comprehension, thus preventing it from reaching an optimal convergence point.

Among all evaluated models, LCNN-EMD exhibited the lowest loss, with values converging to 0.03. Additionally, LCNN-EMD exhibited the highest stability, achieving convergence after approximately 1700 epochs. This performance is attributed to the synergy of LCNN and EMD, which makes LCNN-EMD highly effective in extracting key features relevant to TWM.

4.5. Timeliness analysis

In this section, the timeliness of six different models is evaluated by comparing parameters, accuracy, and the multiply-add divide operations (MADDs) time for a single sample, as shown in Fig. 11.

The RNN shows a maximum MADDs time of 318.5 ms and a parameter count of 2.282 × 10^6. This is attributable to the inherent characteristics of RNNs, which necessitate substantial computational resources to process information at each time step, directly leading to longer MADDs time and a higher number of parameters. While EMD integration raises RNN accuracy to 91.76%, it also increases MADDs time and parameters. Conversely, the CNN achieves the shortest MADDs time of 98.4 ms and a parameter count of 8.15 × 10^5, benefiting from its optimized layer structure and parameter sharing features.

Overall, LCNN-EMD demonstrates dual advantages in efficiency and performance, with an MADDs time of 157.3 ms, an accuracy of 95.42%, and a parameter count of 1.037 × 10^6. The exceptional performance of LCNN-EMD is attributed to several key design features: the inherent feature extraction efficiency of LCNN, which eliminates the need for excessively deep networks; the dynamic nature of adaptive EMD, ensuring the model adapts to different signals without requiring extensive training parameters; and the incorporation of batch normalization, max pooling, and Leaky ReLU activation functions, which mitigate overfitting while rapidly achieving feature dimensionality reduction with minimal computational overhead.

4.6. Sensitivity to the size of the dataset

The performance of deep learning models is closely linked to the size of the dataset. This section evaluates the models' performance across various dataset sizes, specifically with sample sizes of 150, 300, 450, 600, 750, and 900, as shown in Fig. 12.

Models utilizing adaptive EMD, such as DBN-EMD, RNN-EMD, and LCNN-EMD, exhibit less significant declines in accuracy and less dispersion in result distribution with smaller sample sizes. This indicates that adaptive EMD effectively suppresses low-frequency noise with a limited sample size, mitigating the model's tendency to overfit noise due to insufficient training dataset, thereby enhancing the model's generalization capability. However, constrained by the limitations of deep learning architectures, these models exhibit poor convergence in accuracy as the number of samples increases. In contrast, LCNN-EMD

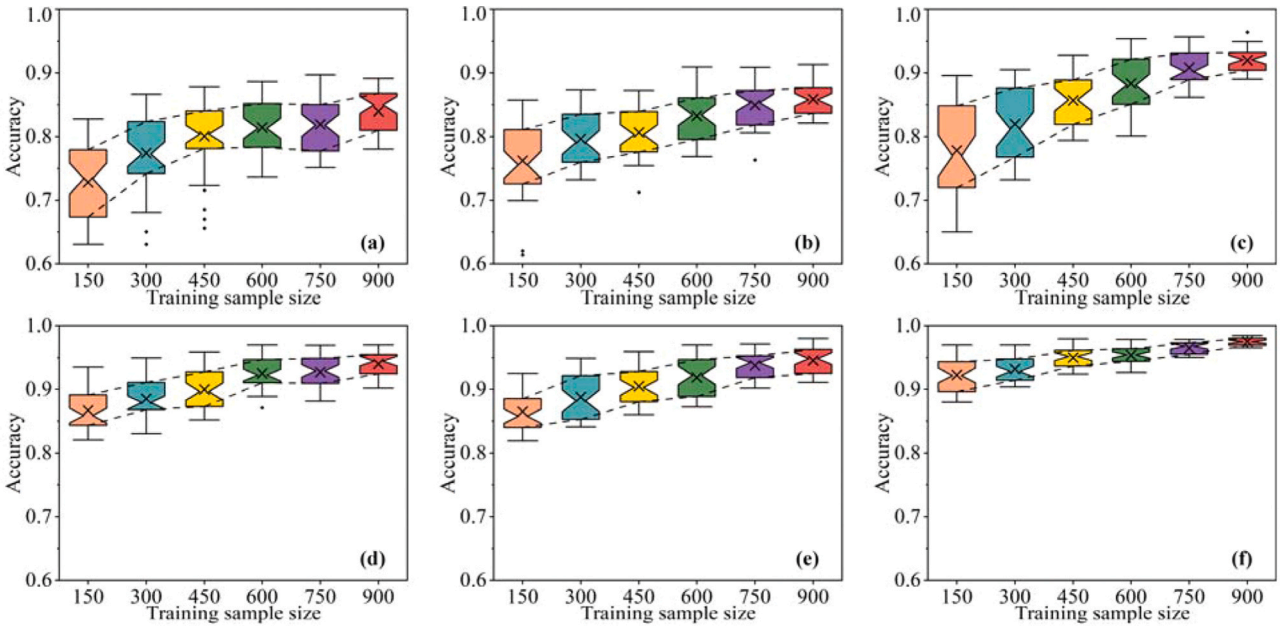

Fig. 12. The accuracy of each model in different training dataset sizes. From a to f: RNN, CNN, U2PL-TCM, DBN-EMD, RNN-EMD, and LCNN-EMD.In the case of limited sample sizes, particularly when the sample size is only 150, the accuracy of U2PL-TCM is merely 0.769, as it fails to learn sufficient feature representations. This inadequacy negatively impacts its accuracy and leads to a widely dispersed result distribution. The large number of trainable parameters in U2PL-TCM makes the quantity of training samples a critical factor in limiting its performance, thereby affecting the model's training adequacy. For RNN and CNN, the distribution of accuracy results shows significant dispersion, including multiple outliers, which further highlights their vulnerability to small datasets. This phenomenon indicates that relying solely on deep learning models is insufficient for accurately extracting key features related to tool wear, especially with limited sample sizes.

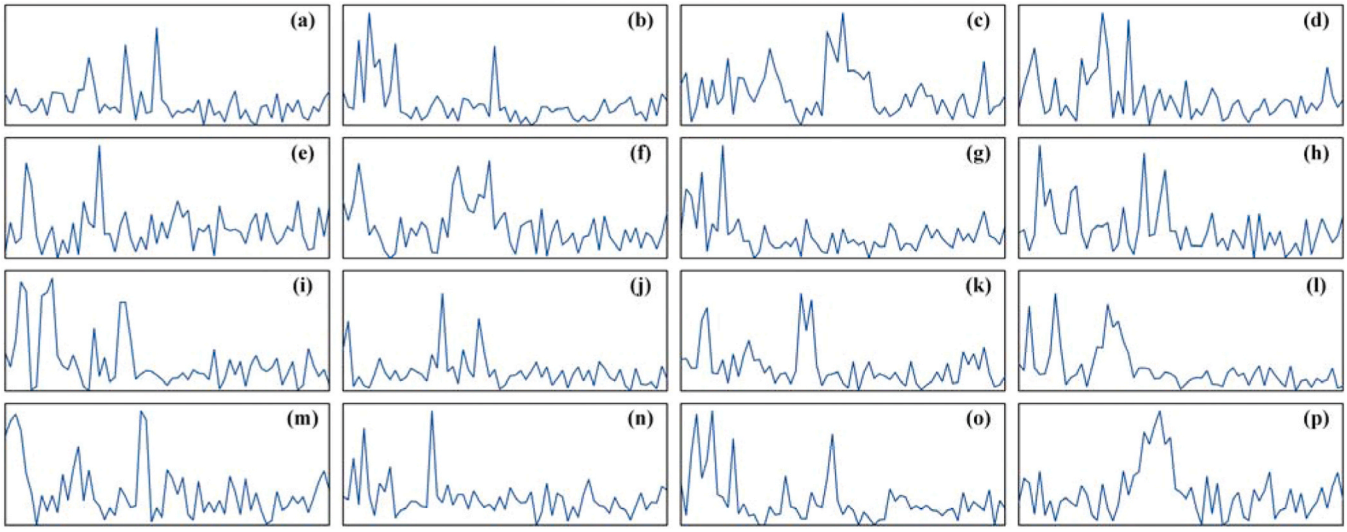

Fig. 13. Frequency domain visualization of one-dimensional large-scale convolution kernel weights. From a to p, each represents one of the 16 channels.

consistently demonstrates more compact accuracy distributions across datasets of varying scales. Despite its complex architecture, LCNN-EMD efficiently identifies and extracts critical features necessary for TWM, ensuring accuracy while reducing dependency on the number of samples.

5. Ablation experiment

5.1. The effect analysis of LCNN

The design and application of large-scale convolution kernels play a crucial role in enhancing the performance of LCNN-EMD. This section provides a visualization analysis of the frequency domain features of one-dimensional large-scale convolution kernel weights, as shown in Fig. 13. From Fig. 13, it can be observed that these kernel functions effectively filter signals in the frequency domain, allowing mid- and low-frequency features to pass through while suppressing high-frequency noise. This capability aids in extracting key tool wear features, thereby enhancing the robustness of LCNN-EMD.

In the large-scale convolution layer, each channel functions as an independent feature detector, capturing specific input frequency features. For example, channels such as those shown in Fig. 14 (c), (f), and (k) are more suitable for the mid-frequency features of the signal, while channels like those in Fig. 14 (b), (i), and (g) prioritize the low-frequency features of the signal. This combined effect enables LCNN-EMD to better adapt to signal variations across different frequency

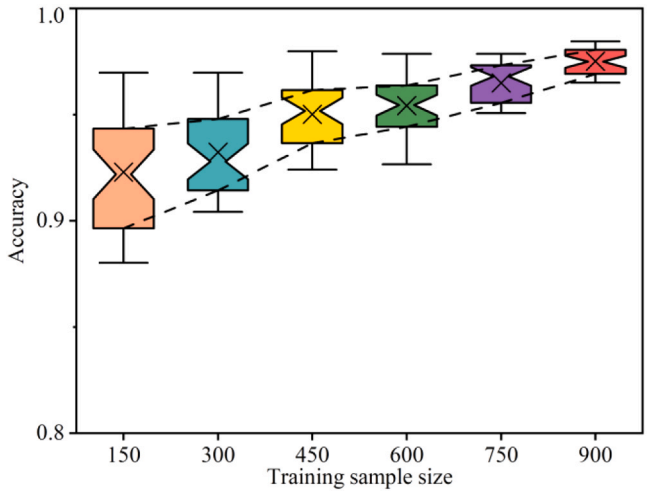

Fig. 14. The accuracy of LCNN-EMD in different training dataset sizes.

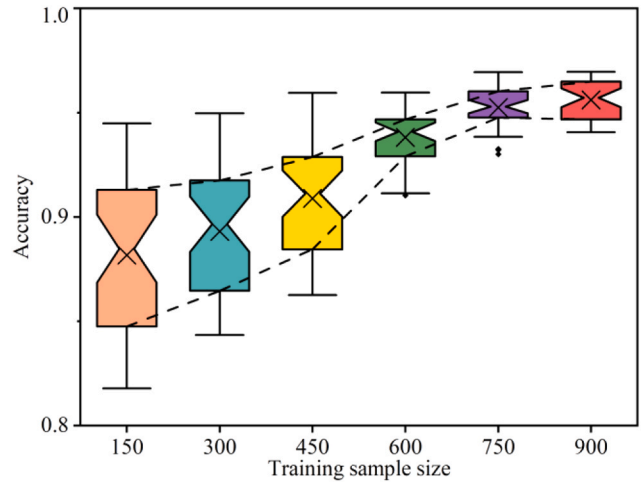

Fig. 15. The accuracy of LCNN in different training dataset sizes.

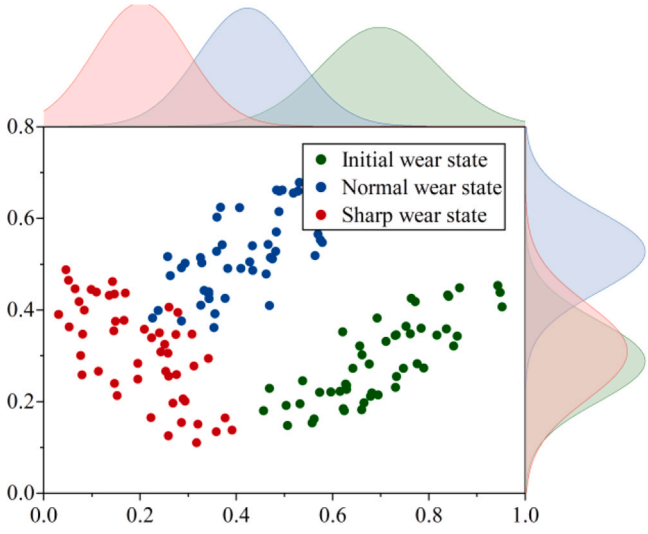

Fig. 16. Visualization of the output features from the last fully connected layer of the LCNN-EMD.

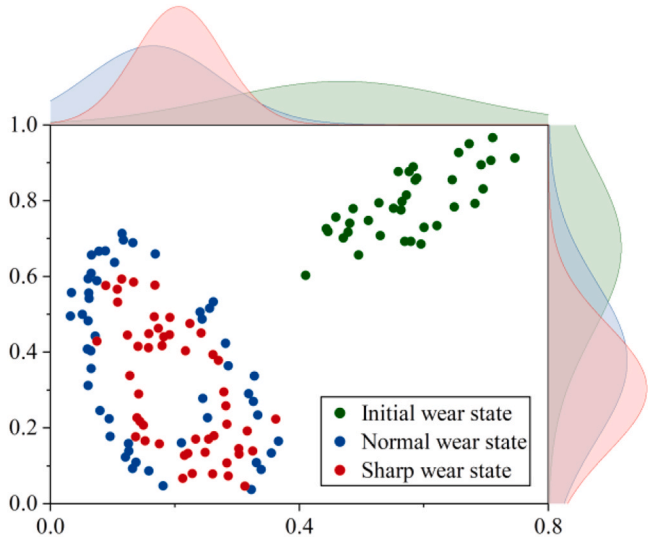

Fig. 17. Visualization of the output features from the last fully connected layer of the LCNN.

ranges, enhancing its signal analysis capability and thereby improving its accuracy in recognizing complex signals. Thus, large-scale convolution kernels play a critical role in LCNN-EMD, providing an effective approach for TWM.

5.2. *The effect analysis of adaptive EMD*

In LCNN-EMD, adaptive EMD is one of the key factors to ensure its robustness. To analyze the impact of adaptive EMD on the robustness of the LCNN-EMD and its underlying mechanisms, we conducted 40 repeated trainings of the LCNN-EMD and LCNN on datasets with varying numbers of training samples. The experimental results, as shown in Fig. 15 and Fig. 16, consistently demonstrate that LCNN-EMD outperforms standard LCNN in both accuracy and result concentration.

Notably, on a small dataset of 150 samples, LCNN-EMD's average accuracy exceeds that of the standard LCNN by 6.52%. This performance improvement is attributed to LCNN-EMD's reliance not only on a fixed set of IMFs but also on the adaptive reconstruction of IMFs, effectively capturing the time-frequency domain features of the signal and transforming them into feature vectors rich with tool wear information.

Visualize the features output by the fully connected layer and reduce dimensionality using T-SNE, as shown in Fig. 16 and Fig. 17. It is evident that in the LCNN, samples from the normal and the sharp wear stage are mixed in the feature space, with significant overlap in both the X and Y axes. Such overlap can interfere with the classification performance of the fully connected layer, ultimately affecting the LCNN's accuracy. In contrast, LCNN-EMD shows clearer sample distribution boundaries in the feature space. Although there is some overlap between the normal wear stage and the sharp wear stage along the Y-axis, the overall distinction in sample distribution is more pronounced, mitigating the feature overlap problem. This clear separation enables LCNN-EMD to better predict tool wear.

5.3. *Effect of the number of dynamic soft thresholding and adaptive EMD in LCNN-EMD*

By selectively omitting specific dynamic soft thresholding and adaptive EMD, we analyzed the resulting changes in model performance. Specifically, we removed the dynamic soft thresholding from the second layer and the adaptive EMD from the fourth layer of LCNN-EMD, creating a modified model referred to as LCNN-EMD-ablation. Both LCNN-EMD and LCNN-EMD-ablation were trained with gaussian noise added to the dataset at SNR levels ranging from 0 to 10 dB, and each

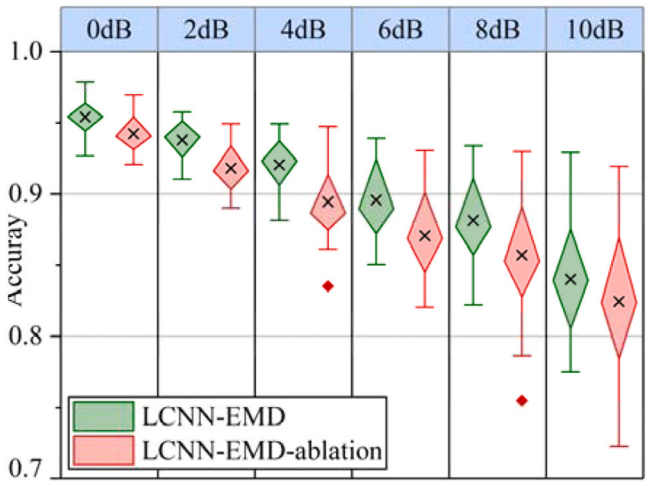

Fig. 18. Performance of each model under different SNR.

model underwent 40 training repetitions. The resulting performance comparisons are shown in Fig. 18.

As the Gaussian noise increased within the SNR range of 0 to 10 dB, the difficulty of isolating useful signal features intensified. In a noise-free environment (0 dB), the accuracy advantage of LCNN-EMD over LCNN-EMD-ablation was minimal, at only 0.012. However, under a 6 dB SNR condition, the accuracy gap widened to 0.021, indicating that the roles of dynamic soft thresholding and adaptive EMD become increasingly significant in maintaining model accuracy as noise levels rise. Notably, LCNN-EMD-ablation exhibited greater sensitivity to noisy environments, with a sharp performance decline and the appearance of outliers particularly at SNRs of 6 dB and 4 dB. These anomalies suggest that the reduction of dynamic soft thresholding and adaptive EMD diminishes the model's robustness against noise interference, rendering LCNN-EMD-ablation less capable of maintaining stable feature extraction.

The dynamic soft thresholding adaptively adjusts thresholds to balance the proportion of signal to noise, effectively suppressing high-frequency noise—especially impactful after the initial convolutional layers, where it substantially reduces noise influence during feature extraction. Conversely, the adaptive EMD gradually filters out low-frequency noise through multi-level decomposition of IMFs. The synergistic effect of dynamic soft thresholding and adaptive EMD ensures that the network can dynamically adjust and eliminate noise at various levels, thereby enhancing the network's noise robustness and the accuracy of feature extraction.

6. Conclusion

In this study, we have proposed the LCNN-EMD architecture for TWM, which combines LCNN with adaptive EMD. This novel architecture addresses the limitations of traditional models that rely heavily on expert knowledge and significantly enhances data flow consistency and processing efficiency. Using visual analysis, we demonstrated that the design of LCNN and dynamic soft thresholding targets high-frequency noise suppression, while adaptive EMD is designed to suppress low-frequency noise and retain mid-frequency features. Through tool wear experiments, it is shown that LCNN-EMD consistently outperforms other mainstream models in terms of accuracy and stability, with significant improvements in tool wear prediction and noise suppression.

It is evident that the trend in TWM technology is towards more automated and intelligent systems, reducing reliance on manual input and expert knowledge. The development of the LCNN-EMD architecture is a significant step in this direction, providing a more reliable and effective architecture for monitoring tool wear. However, in order to further improve the maturity and application range of this technology, further research work is still needed. First, to enhance the robustness and generalizability of LCNN-EMD, future studies should focus on cross-validating the model using datasets from different machining environments, tool types, and a wider range of operating conditions. Additionally, exploring the integration of LCNN-EMD with edge computing can aid in deploying the model in environments with limited access to high-performance computing resources. This will facilitate real-time monitoring and decision-making at the production site.

CRediT authorship contribution statement

Shuang Yin: Visualization. **Jie Li:** Investigation. **Hongli Gao:** Project administration, Funding acquisition. **Hongliang Song:** Software, Formal analysis, Data curation. **Yi Sun:** Writing – review & editing, Writing – original draft, Validation, Software, Investigation.

Declaration of Competing Interest

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

Acknowledgements

This work is supported by Sichuan Science and TechnologyMajor Project (2022ZDZX0044).

References

[1] Li ZX, Liu R, Wu DZ. Data-driven smart manufacturing: tool wear monitoring with audio signals and machine learning. J Manuf Process 2019;48:66–76.
[2] García-Pérez A, Ziegenbein A, Schmidt E, Shamsafar F, Fernández-Valdivielso A, Llorente-Rodríguez R, et al. CNN-based in situ tool wear detection: a study on model training and data in inserts. J Manuf Syst 2023;68:85–98.
[3] Lamraoui M, Thomas M, Badaoui MEl. Cyclostationarity approach for monitoring chatter and tool wear in high speed milling. Mech Syst Signal Pract 2014;44:177–98.
[4] Zhu KP, Li X, Li SS, Lin X. Physics-informed hidden markov model for tool wear monitoring. J Manuf Syst 2024;72:308–22.
[5] del Olmo A, de Lacalle LNL, de Pissón GM, Pérez-Salinas C, Ealo JA, Sastoque L, et al. Tool wear monitoring of high-speed broaching process with carbide tools to reduce production errors. Mech Syst Signal Pract 2022;172.
[6] Guo JM, Zhang JW, Wu D, Gai YH, Chen K. A local manipulation path replanning algorithm on deformable linear objects for collisions resulted from model deviation. J Manuf Syst 2022;65:362–77.
[7] Huang ZW, Li WD, Zhu JM, Wang LH. Cross-domain tool wear condition monitoring via residual attention hybrid adaptation network. J Manuf Syst 2024;72:406–23.
[8] Li RY, Wei PN, Liu XL, Li CL, Ni J, Zhao WK, et al. Cutting tool wear state recognition based on a channel-space attention mechanism. J Manuf Syst 2023;69:135–49.
[9] Wei YP, Wu DZ. Model-based real-time prediction of surface roughness in fused deposition modeling with graph convolutional network-based error correction. J Manuf Syst 2023;71:286–97.
[10] Al-Habaibeh A, Gindy N. Self-learning algorithm for automated design of condition monitoring systems for milling operations. Int J Adv Manuf Tech 2001;18:448–59.
[11] Guo L, Yu YX, Gao HL, Feng TT, Liu YK. Online remaining useful life prediction of milling cutters based on multisource data and feature learning. IEEE Trans Ind Inf 2022;18:5199–208.
[12] Yang YF, Hao BJ, Hao XQ, Li L, Chen N, Xu T, et al. A novel tool (single-flute) condition monitoring method for end milling process based on intelligent processing of milling force data by machine learning algorithms. Int J Precis Eng Man 2020;21:2159–71.
[13] Traini E, Bruno G, Lombardi F. Tool condition monitoring framework for predictive maintenance: a case study on milling process. Int J Prod Res 2021;59:7179–93.
[14] Zhou YQ, Xue W. A multisensor fusion method for tool condition monitoring in milling. Sensors 2018;18.
[15] Yang Q, Mishra D, Awasthi U, Bollas GM, Pattipati KR. Tool wear and remaining useful life estimation in precision machining using interacting multiple model. J Manuf Syst 2024;74:367–86.
[16] Zhu KP, Vogel-Heuser B. Sparse representation and its applications in micro-milling condition monitoring: noise separation and tool condition monitoring. Int J Adv Manuf Tech 2014;70:185–99.
[17] Tran MQ, Liu MK, Elsisi M. Effective multi-sensor data fusion for chatter detection in milling process. ISA Trans 2022;125:514–27.

[18] Zhou CA, Guo K, Sun J. Sound singularity analysis for milling tool condition monitoring towards sustainable manufacturing. Mech Syst Signal Pract 2021;157.
[19] Li S, Mao JD, Gong X, Li ZY. A novel lidar signal noise reduction algorithm based on improved deep belief network. Phys Scr 2023;98.
[20] Garg J, Garg SB. A simplified methodology for finding the natural frequencies and mode shapes of the machine tool structures. Mater Today Proc 2021;38:99–104.
[21] Jia N, Cheng Y, Liu YY, Tian YY. Intelligent fault diagnosis of rotating machines based on wavelet time-frequency diagram and optimized stacked denoising auto-encoder. IEEE Sens J 2022;22:17139–50.
[22] Li XP, Nie HF. Analysis on dynamic characteristics of CNC machine tool and experiment with influence of rolling guide joint considered. Adv Mater Res 2012; 443-444:745.
[23] Zhang CJ, Yao XF, Zhang JM, Jin H. Tool condition monitoring and remaining useful life prognostic based on a wireless sensor in dry milling operations. Sensors 2016;16.
[24] Sabareeswaran M, Padmanaban KP, Sundararaman KA. Comparison of evolutionary techniques for the optimization of machining fixture layout under dynamic conditions. Proc Inst Mech Eng C J Mech 2018;232:2145–58.
[25] Zhou YQ, Sun WF. Tool wear condition monitoring in milling process based on current sensors. IEEE Access 2020;8:95491–502.
[26] Gao HN, Shen HD, Yu L, Yinling W, Li RY, Nazir B. Milling chatter detection system based on multi-sensor signal fusion. IEEE Sens J 2021;21:25243–51.
[27] Wang DL, Wu Q. On the selection of hyperparameters in convolutional neural networks. Int Conf Comput Sci Comput Intell (CSCI 2021) 2021;2021:1728–31.
[28] Isavand J, Kasaei A, Peplow A, Wang XF, Yan JH. A reduced-order machine-learning-based method for fault recognition in tool condition monitoring. Measurement 2024;224.
[29] Jiang M, Munawar MA, Reidemeister T, Ward PAS. Efficient fault detection and diagnosis in complex software systems with information-theoretic monitoring. IEEE Trans Depend Secur 2011;8:510–22.
[30] Sun Y, He JG, Gao HL, Song HL, Guo L. A new semi-supervised tool-wear monitoring method using unreliable Pseudo-labels. Measurement 2024;226.
[31] Qin B, Wang YQ, Liu K, Jiang SW, Luo Q. A novel online tool condition monitoring method for milling titanium alloy with consideration of tool wear law. Mech Syst Signal Pract 2023;199.
[32] XC Cao, B Yao, BQ Chen, WP He, SQ Guo, K Chen. Intelligent tool condition monitoring based on multi-scale convolutional recurrent neural network, IEICE Trans Inf Syst, E106d; 2023: p. 644–52.
[33] Wang DS, Hong RJ, Lin XC. A method for predicting hobbing tool wear based on CNC real-time monitoring data and deep learning. Precis Eng 2021;72:847–57.