FACTA UNIVERSITATIS
Series: **Mechanical Engineering**
https://doi.org/10.22190/FUME240804047G

Original scientific paper

A MONITORING METHOD OF MILLING CHATTER BASED ON OPTIMIZED HYBRID NEURAL NETWORK WITH ATTENTION MECHANISM

Haining Gao[1], Hongdan Shen[1], Caixu Yue[2], Rongyi Li[2], Steven Y. Liang[3], Yinlin Wang[1], Wenfu Liu[1], Yong Yang[1]

[1]Huanghuai University, Zhumadian, China
[2]Harbin University of Science and Technology, Harbin, China
[3]Georgia Institute of Technology, Atlanta, USA

Abstract. *Machining chatter is a self-excited vibration between the cutting tool and the workpiece, which can reduce surface quality and tool life, and even endanger the safety of operators in severe cases. Considering that milling chatter has multi-scale features and the debugging of neural network hyperparameters heavily relies on experience, a milling chatter monitoring method based on an optimized hybrid neural network with an attention mechanism (MISSA-MSCNN-BiLSTM-ATM) is proposed. Firstly, the harmonic of the spindle rotation frequency is filtered out using the spindle rotation frequency removal technique (SFT). Then, an improved sparrow search algorithm (MISSA) is proposed based on multiple strategies including improved circle chaotic mapping, golden sine strategy, and enhanced Lévy flight. Subsequently, MISSA is utilized to optimize the hyperparameters of the milling chatter classification hybrid neural network model, combining multi-scale convolutional neural networks (MSCNN), bidirectional long short-term memory (BiLSTM), and attention mechanism (ATM). In numerical simulations with CEC2005 complex functions, MISSA demonstrates better optimization accuracy, stability, and shorter computation time compared to other intelligent algorithms. Compared with other milling chatter classification models, the proposed method exhibits significant improvements in accuracy and stability.*

Key words: *Chatter monitoring, Attention mechanism, Hybrid neural network, Optimization algorithm*

Received: August 04, 2024 / Accepted November 10, 2024
Corresponding author: Haining Gao
College of Energy Engineering, Huanghuai University, No. 76 Kaiyuan Avenue, Yicheng District, Zhumadian, China, 463000
E-mail: 20191908@huanghuai.edu.cn

1. INTRODUCTION

Aerospace components, such as turbine blades, are characterized by thin-walled structures, complex geometries, and are predominantly manufactured from difficult-to-machine materials such as titanium alloys. Chatter frequently occurs during the milling process, resulting in deteriorated surface quality and accelerated tool wear. Recent research has demonstrated that chatter monitoring is crucial for ensuring both the integrity of the cutting system and workpiece quality in aerospace manufacturing processes.

Chatter monitoring consists of four primary components: signal acquisition, signal processing, feature extraction, and pattern recognition. Various signal types employed for chatter monitoring include vibration signals [1-3], cutting force signals [4-7], current signals [8-10], sound signals [11-12], and multi-sensor information [13-15]. Among these, acceleration signals are particularly advantageous due to their non-intrusive nature, high reliability, and sufficient response bandwidth [4]. However, experimental data acquisition through sensors is characterized by limited processing conditions, high costs, and significant time requirements. To address these limitations, scholars have generated simulated cutting force signals covering a wide range of machining parameters based on dynamic milling process models [16], or used generative adversarial networks to generate rich training data [17]. Current multi-sensor fusion approaches for chatter detection lacks comprehensive theoretical frameworks for sensor selection, sampling strategies, and fusion mechanisms. Consequently, these methods predominantly rely on empirical parameter settings and case-specific validation, creating significant challenges in establishing unified evaluation criteria and generalized solutions.

Sensor data acquisition in machining processes encompasses multiple signal components, including spindle rotation frequency with its harmonics, chatter frequencies, local outliers, and measurement noise. In chatter monitoring systems, the presence of local outliers and noise can result in both false alarms and missed detections [18]. While stable components, particularly the spindle rotation frequency and its associated harmonics, do not contribute to chatter phenomena, their removal can significantly enhance chatter monitoring sensitivity [19]. To address these issues, scholars have conducted extensive research. Albertelli et al. [18] used order tracking and synchronous averaging methods to eliminate noise and cutting periodicity effects. Yan et al. [19] removed periodic components related to spindle rotation using a designed time-frequency filter. Wan et al. [20] used a matrix notch filter to filter out the harmonics of the spindle rotation frequency and color noise components. Matthew et al. [21] used an adaptive threshold method to reduce the impact of noise, rotation frequency, and harmonics. Subsequently, a wavelet based adaptive filter was constructed to reduce the impact of noise [4]. Kuo et al. [17] employed fractional-order convolutional kernels to eliminate local outliers in training data. Li et al. [22] added an artificial offset after filtering out stable frequency components, avoiding zero division operations. Wan et al. [23] used their proposed variable forgetting factor recursive least squares algorithm to filter out chatter-independent components composed of environmental noise and periodic components.

Feature extraction methodologies in machining processes can be categorized into three primary domains: time-domain, frequency-domain, and time-frequency domain approaches. Time-domain feature extraction offers computational efficiency and straightforward implementation. These primarily comprise dimensionless features, including statistical measures such as mean [10,19], standard deviation [10,19], root mean square [4,5,19,24], shape factor [19,23,25], and kurtosis [4,19,23,25]. In the frequency domain, features are derived through discrete Fourier transform analysis and non-parametric power spectral

density estimation. Key metrics in this domain include spectral mean [19,25] and spectral centroid frequency [19,25]. Time-frequency domain methods have gained significant prominence due to their capability to simultaneously capture temporal and spectral characteristics. These methods encompass short-time Fourier transform [1,10,21], continuous wavelet transform [1,26], wavelet packet transform [9,14,27,28], variational mode decomposition (VMD) [29,30,31], and multi-synchronous compression transform [4,19,31]. Jauhari et al. [30] introduced Bayesian optimization to the VMD parameter selection process, achieving automatic optimization of decomposition parameters. In addition to the above features, some special features have also been applied to chatter monitoring and achieved good results. Thaler et al. [24] found that deterministic features extracted from recurrence plots can effectively monitor chatter. Whether the recurrence plot can accurately reflect the dynamic characteristics of the cutting system depends on the setting of its hyperparameters. To address this issue, Chen et al. [32] proposed an adaptive particle swarm optimization algorithm to calculate the hyperparameters and automatically obtain the recurrence plot.

Threshold methods or intelligent classification algorithms are often used to identify chatter. Many chatter indicators have been proposed to detect chatter, such as entropy [14, 23,30], energy ratio [19], and statistical indicators [18,28,29,33]. However, most existing indicators are directly or indirectly related to machining parameters, leading to changes in the indicators with changes in machining parameters and making early chatter detection difficult. Albertelli et al. [18] developed an automated chatter threshold determination method based on statistical process control theory. Lu et al. [29] selected the correlation coefficient between chatter-sensitive intrinsic mode functions as a chatter indicator. This indicator is not easily affected by machining parameters. Zhao et al. [33] accumulated the information content in the maximum information entropy features calculated from the sequence to obtain a chatter indicator. The threshold based on two risk levels overcomes the limitations of a single threshold and improves the reliability of monitoring. Yang et al. [28] established a fractional-order energy entropy probability model for chatter detection, overcoming the difficulty of manually adjusting thresholds.

The most widely used intelligent classification algorithms are support vector machines [15,23,25,29], logistic regression [34], and random forest [19,14,34]. Wan et al. [25] improved the support vector machine using the adaptive boosting algorithm, which improved the accuracy of chatter monitoring. Deep learning has demonstrated powerful feature learning and classification capabilities, attracting widespread attention and gradually being applied to the field of chatter detection. Sener et al. [26] proposed a chatter detection method based on a deep convolutional neural network. This method introduced cutting parameters such as cutting depth and spindle speed into the DCNN model, effectively improving the test accuracy. Transfer learning approaches [34,35,36] have been successfully applied to chatter detection to leverage knowledge from source domains to target domains. Kounta et al. [35] proposed a transfer learning framework based on pre-trained VGG16 and ResNet50 models, which achieves high-precision detection of chatter in industrial machining processes through the extraction and classification of FFT image features.

The SFT is employed to filter out the spindle rotation frequency and its multiples. Then, by combining time-frequency domain methods, time-domain features, frequency-domain features, and time-frequency domain features related to the machining state are obtained. Subsequently, the MISSA algorithm is derived by optimizing the sparrow search algorithm through a combination of strategies including improved circle chaos mapping, golden sine strategy, and enhanced Levy flight. A MSCNN module integrated with spatial attention

mechanism is constructed to adaptively extract critical spatial feature information at different scales. Next, BiLSTM is utilized to handle long-term dependencies in sequences and capture temporal feature information. Then, by integrating the multi-head attention mechanism (MSA), the contribution degree of extracted features is effectively grasped, and the features are emphasized and strengthened, leading to the MSCNN-BiLSTM-ATM chatter identification model. Finally, the MISSA algorithm is used to optimize the hyperparameters in the model, thereby improving the accuracy of chatter recognition.

2. Signal Preprocessing and Feature Extraction

2.1 Signal Preprocessing

The periodic components (spindle rotation frequency and its multiples) generated by spindle rotation and intermittent contact between the tool and workpiece are the main components of the machine tool processing acquisition signals. However, when chatter occurs in the system, non-periodic components caused by regenerative effects will be generated, and these components will gradually approach the natural frequency of the machine tool. But in the early stages of chatter, the frequency amplitude is very small and often submerged in the rotation frequency, making it impossible to detect it from the signal in time. For example, in the initial stage of chatter, the energy proportion of periodic components at low frequencies still accounts for more than 80% [24], greatly increasing the difficulty of monitoring the initial stage of chatter. Therefore, in order to achieve high-precision and efficient identification in the initial stage of chatter, it is particularly important to remove the spindle rotation frequency and its multiples.

Compared to traditional notch filters that only remove a single frequency, the proposed SFT can eliminate multiple related frequencies, effectively highlighting the initial chatter frequency submerged in the frequency spectrum. Firstly, the spindle rotation frequency and its multiples that need to be removed are identified. Secondly, a Fast Fourier Transform (FFT) is applied to the collected signal, and the sections of the spectrum corresponding to the spindle rotation frequency and its multiples are removed. Finally, an inverse Fast Fourier Transform (IFFT) is conducted on the defrequenced spectrum. The specific algorithm steps are outlined below.

Step 1: Perform FFT on time series $\{x\{k\}, k=1,2,\ldots,M\}$ to obtain spectrum $X(i)$ as shown in Eq. (1). The length of the spectrum is $M=length(X)$

$$X(i)=\sum_{k=1}^{M}x(k)W_M^{ki} \quad k=1,2,\cdots,M \tag{1}$$

where, j is an imaginary unit and $W_M^{ki}=\exp\left(-j\frac{2\pi}{M}\right)$.

Step 2: Set the spindle rotation frequency f_z that needs to be removed. Its doubling frequency is cf_z ($c=1,2,\ldots,N$). The calculation of spindle rotation frequency is shown in Eq. (2).

$$f_z=n_s/60 \tag{2}$$

where, n_s is the spindle speed.

Step 3: Considering the fluctuation of FFT transformation frequency, introduce c_1 floating constant with a value of 0.5. Set the spectral amplitude of the interval corresponding to frequency (f_z-c_1,f_z+c_1) in the spectrum to zero, and the specific calculation is shown in Eq. (3).

$$\bar{X}(i)=\begin{cases} 0 & cf_z-c_1\le i\le cf_z+c_1 \\ X(i) & (c-1)f_z+c_1\le i\le cf_z-c_1 \end{cases} \tag{3}$$

Step 4: If i a is greater than $c\cdot f_z+c_1$ and less than $N/2$, make $c=c+1$ and repeat steps 2 to step 4. Otherwise, proceed to the next step.

Step 5: Because the spectrum is symmetric about the center, let $X(M+1-i)=X(i)$, $i=1,2,\ldots,M/2$, to obtain the complete spectrum after removing the frequency.

Step 6: Perform IFFT on the spectrum to obtain the time-series after frequency reduction, and calculate as shown in Eq. (4).

$$x(k)=\sum_{i=1}^{M}\bar{X}(i)W_M^{-ki} \tag{4}$$

2.2 Feature Extraction

Comprehensively and accurately obtaining the state characteristics during the machining process is the prerequisite to precise identification of milling chatter. Currently, time-domain methods, frequency-domain methods, and time-frequency domain methods are commonly used to obtain state characteristics during the machining process. The vibration signals, cutting force signals, and torque signals of the milling process have nonlinear and non-stationary characteristics [13]. These signals contain a large amount of state information from the machining process and are a comprehensive reflection of various excitation signals. Since time-domain and frequency-domain features are only calculated from one domain of the data, important features with high resolution are discarded, while time-frequency domain methods such as wavelet analysis can solve the above problems. In this paper, time-domain features, frequency-domain features, and time-frequency domain features are selected as state features, as shown in Table 1.

3. MISSA

The traditional sparrow search algorithm (SSA) has the problem of being easily trapped in local optima. This article optimizes the sparrow search algorithm by improving the circle chaotic mapping, golden sine strategy, and enhanced Lévy flight strategy to improve the optimization accuracy.

3.1 Improving the Circle Chaotic Mapping

The SSA population is initialized using the improved circle chaotic mapping to ensure a uniform distribution, as follows:

$$x_{n+1}=\mathrm{mod}(3.85\cdot x_n+0.4-\frac{0.7}{3.85\cdot\pi}\sin(3.85\cdot\pi\cdot x_n),1) \tag{5}$$

where, x_{n+1} is the value of introducing Singer chaotic mapping, n is the dimension of the solution.

Table 1 Time frequency domain characteristics

Category	Feature	Expression	Feature	Expression
Time domain	Peak to peak value (PTP)	$X_1 = \max\|x_i\| - \min\|x_i\|$	Root mean square (RMS)	$X_2 = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}$
	Absolute mean (ABS)	$X_3 = \frac{1}{N}\sum_{i=1}^{N}\|x_i\|$	standard deviation (STD)	$X_4 = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\|x_i\| - X_3\right)^2}$
	Kurtosis (KUR)	$X_5 = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\|x_i\| - X_3}{X_4}\right)^4$	Skewness (SKE)	$X_6 = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\|x_i\| - X_3}{X_4}\right)^3$
	Peak factor (PAF)	$X_7 = \frac{\max\|x_i\|}{X_2}$	Pulse factor (PUF)	$X_8 = \frac{\max\|x_i\|}{X_3}$
	Waveform factor (WAF)	$X_9 = \frac{X_2}{X_3}$	Clearance factor (CF)	$X_{10} = \frac{\max(\|x_i\|)}{X_2^2}$
	Mean of frequency spectrum (MFA)	$X_{11} = \frac{1}{N}\sum_{i=1}^{N} P(f_i)$	Gravity frequency of frequency spectrum (GFFS)	$X_{12} = \frac{\sum_{i=1}^{N}(f_i \cdot P(f_i))}{\sum_{i=1}^{N} P(f_i)}$
Frequency domain	Mean square frequency of frequency spectrum (MSFS)	$X_{13} = \frac{\sum_{i=1}^{N}(f_i \cdot f_i \cdot P(f_i))}{\sum_{i=1}^{N} P(f_i)}$	Root mean square frequency of frequency spectrum (RMSFS)	$X_{14} = \sqrt{X_{13}}$
	Frequency spectrum variance (FSV)	$X_{15} = \sum_{i=1}^{N}((f_i - FC)^2 \cdot P(f_i)) \Big/ \sum_{i=1}^{N} P(f_i)$		
Time - frequency	Wavelet energy entropy	$X_{16} = \sum_{i=1}^{N}\left\|x_{(j,i)}^{p}\right\|^2$		

3.2 Golden Sine Strategy

Add the golden sine strategy when updating the discoverer's position, and the update formula is shown in Eq. (6).

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^{t} \times \left|\sin(r_1)\right| + r_2 \times \sin(r_1) \times \left| x_1 \times X_{best}^{t} - x_2 \times X_{i,j}^{t} \right|, R_2 < ST \\ X_{i,j}^{t} + Q \cdot L, R_2 > ST \end{cases} \tag{6}$$

with

$$\begin{cases} x_1 = -\pi + (1 - ((\sqrt{5} - 1)/2)) \times 2\pi \\ x_2 = -\pi + ((\sqrt{5} - 1)/2) \times 2\pi \end{cases} \tag{7}$$

where, t is the current iteration number, $r_1 \in [0,2\pi]$, $r_2 \in [0,\pi]$, respectively determine the distance and direction of individual position iteration movement, x_1 and x_2 are the gold coefficients.

3.3 Enhanced Lévy Flight Strategy

The Lévy flight step size factor is usually a fixed value. When its value is set to a larger number, although the global search ability is enhanced, a high-precision solution cannot be obtained. When its value is set to a smaller number, more iterations are required to find the theoretical value of the algorithm, resulting in reduced efficiency. To address this issue, a Lévy flight mechanism with enhanced search is proposed as follows.

The calculation of the dynamic step size factor is given below:

$$\alpha(t) = \frac{t}{T_{\max}} \sinh(1 - \frac{t}{T_{\max}}) \times r \tag{8}$$

where, t is the current iteration number, T_{max} is the maximum iteration number, r is the adjustment parameter. By plotting the function curve of Eq. (8), we can observe that the range of $\alpha(t)$ is [0, 0.261]. Multiple experimental results have shown that when r=3.86, the variation condition of $\alpha(t)$ at [0,1] can be met, and the enhanced Levi flight strategy can find the optimal value.

Although the new positions generated through Lévy flight search can escape from local optima, it cannot be guaranteed that the updated positions are evenly distributed around the optimal position. Therefore, a Gaussian random distribution function is added as a constraint factor in the Lévy mechanism to make the particle population more evenly distributed in the exploration space. The definition of the constraint factor is shown in the following equation:

$$\gamma = N(0, \delta) \tag{9}$$

where, δ is a scale parameter with a value of t/T_{max}.

The final enhanced Levi's flight expression is shown below:

$$Levy = \gamma \cdot \alpha(t) \cdot \frac{\mu \cdot \sigma}{|\nu|^{\frac{1}{\beta}}} \tag{10}$$

with

$$\sigma = \left[\frac{\Gamma(1+\beta) \times \sin(\frac{\pi \cdot \beta}{2})}{\Gamma(\frac{1+\beta}{2}) \times \beta \times 2^{(\frac{\beta-1}{2})}} \right]^{\frac{1}{\beta}} \tag{11}$$

where, β is 1.5, μ and ν follow a normal distribution.

The position update of the joiner is shown in Eq. (12).

$$X_{i,j}^{t+1} = \begin{cases} Q \cdot \exp\left(\frac{X_{worst} - X_{i,j}^{t}}{i^2}\right), i > \frac{N}{2} \\ X_p^{t+1} + \left| X_{i,j}^{t} - X_p^{t+1} \right| A^{+} \cdot L \cdot Levy, otherwise \end{cases} \tag{12}$$

4. CHATTER MONITORING MODEL

4.1 MSCNN

CNN is a feedforward neural network inspired by biology that is mainly composed of convolutional layers, pooling layers, and fully connected layers. The convolutional layer mainly extracts implicit feature information from the data through convolutional operations and then uses activation functions to introduce nonlinear factors into the features, enhancing the expressive ability of the features. The operation process of the convolutional layer is shown in Eq. (13).

$$X_{i,j}^{l+1} = \sum_{j=1}^{L} \sum_{i=1}^{m} (X_{i,j}^{l} \times w_{i,j}^{l}) + b \tag{13}$$

where, $X_{l+1i,j}$ is the feature of layer l+1 after convolution calculation, $w_{i,j}^{l}$ is a convolutional kernel, $X_{i,j}^{l}$ is the j-th feature value of the i-th feature map, b is the bias term, and L is the size of the convolutional kernel.

The time-domain vibration and cutting force signals in cutting machining often show the characteristics of multiple time scales. When the cutting chatter occurs, its characteristics usually also show the nature of multi-scale. To effectively extract the milling chatter component and enhance the recognition model effect, this paper designs a model based on the Inception module and optimizes its model structure. The feature extraction ability under different scales is enhanced, as shown in Fig. 1.

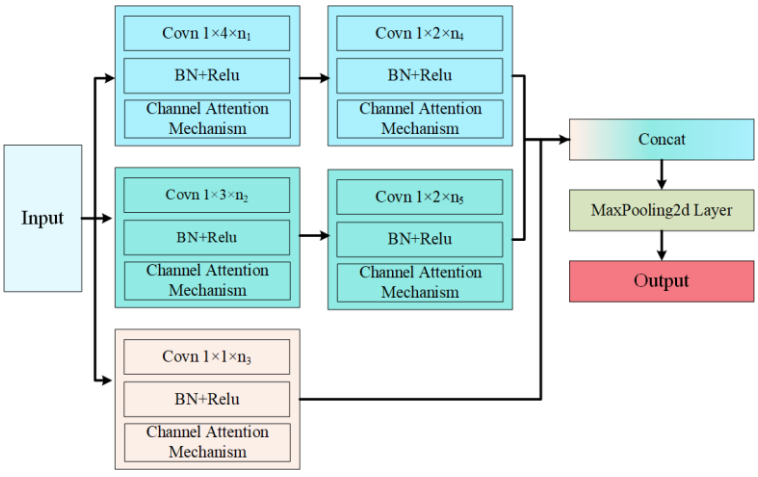


Fig. 1 MSCNN structure

The size of the convolution kernel in each layer of the multi-scale convolutional neural network is fixed, and the number of channels is a parameter to be optimized. A channel attention mechanism is added to each convolutional neural network module to generate weights for each channel and continuously optimize and update the network parameters during the training process. The multi-scale convolutional neural network can continuously optimize the processing status data to obtain the importance of different features, strengthen key features based on their importance, suppress interference information, and improve the diagnostic efficiency and effectiveness of the model.

4.2 BiLSTM

BiLSTM fully considers the relationship between the past and future of time series data and can effectively extract the temporal characteristics of signals for time series data with high correlation and strong periodic changes.

The data enters the BiLSTM through the input layer, and a value is obtained through forward calculation of LSTM, while another value is also obtained through backward calculation of LSTM. Therefore, the value in the hidden layer is determined by these two values. The specific calculation process is shown in Eq. (14).

$$\begin{cases} \vec{h}_t = f(\vec{w}_x \cdot X_t + \vec{h}_{t-1} \cdot \vec{w}_h + \vec{b}_n) \\ \overleftarrow{h}_t = f(\overleftarrow{w}_x \cdot X_t + \overleftarrow{h}_{t-1} \cdot \overleftarrow{w}_h + \overleftarrow{b}_n) \\ Y_t = f(\vec{w}_y \cdot \vec{h}_t + \overleftarrow{w}_y \cdot \overleftarrow{h}_t + b_y) \end{cases} \tag{14}$$

where, f is the activation function, w is the weight and bias term, $\vec{h}_t$ is the output of the forward layer, $\overleftarrow{h}_t$ is the output of the backward layer, and Y_t is the final output.

4.3 Attention Mechanism

4.3.1 Channel Attention Mechanism

The channel attention mechanism mainly consists of squeeze and excitation operations, as shown in Fig. 2. Firstly, the feature map undergoes the squeeze operation, compressing the two-dimensional feature map into a real number through global pooling to obtain global features. Then, the excitation operation is performed on the global features, and a weight value is generated by passing the obtained global features through two fully connected layers and a sigmoid function. Finally, the weight value is multiplied by the original feature map to obtain the final feature. The calculation process is as follows.

$$z_c = F_{sq}(\mu_c) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} \mu_c(i, j) \tag{15}$$

$$s = F_{ex}(z, W) = \sigma_{sigmoid}(W_2 \cdot \sigma_{\mathrm{Re}LU}(W_1 z)) \tag{16}$$

$$\tilde{x}_c = F_{scale}(u_c, s_c) = s_c \cdot u_c \tag{17}$$

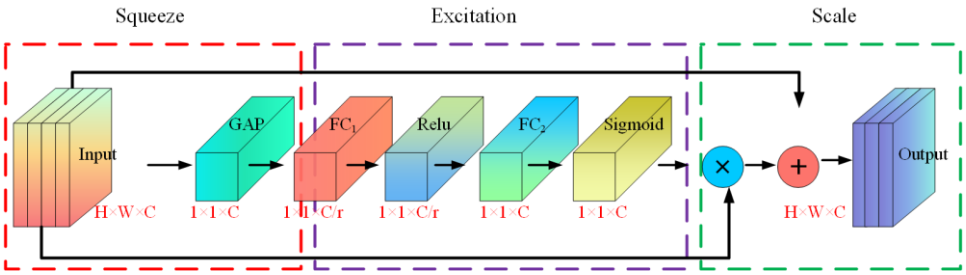


Fig. 2 Channel attention mechanism

4.3.2 Multihead Self-Attention (MSA)

The multi-head attention mechanism is composed of multiple self-attention structures and is used to simultaneously process the same feature information. Its output is the concatenation of multiple self-attention results. The formula is as follows.

$$MultiHead(Q,K,V) = Concat(head_1,\cdots,head_h)\cdot W^o \tag{18}$$

$$head_i = Attention(QW_i^Q, KW_i^K, VW_i^V) \tag{19}$$

where, W_i^Q, W_i^K and W_i^V are the weights of the mapping matrix, and W^o is the output weight matrix.

4.4 MISSA-MSSNN-BiLSTM-MSA Chatter Monitoring Model

The structure of the MSCNN-BiLSTM-ATM chatter monitoring model constructed in this paper is shown in Fig. 3.

There are many hyperparameters in the MSCNN-BiLSTM-ATM model, such as learning rate, regularization parameter, number of neural network layers, number of convolutional layers, and batch size, which will affect the training effect of the model. Therefore, the MISSA algorithm is introduced to optimize the hyperparameters in the model so as to improve its accuracy.

The accuracy of the validation set samples is used as the fitness function max$f$ for the MISSA algorithm to optimize the hyperparameters of the MSCNN-BiLSTM-ATM model.

$$\begin{cases} \max f = \dfrac{\sum_{i=1}^{N}(Y_i - Y_{vaild})\times 100}{N} \\ \mathrm{numF}_j \in [\mathrm{numF}_{\min}, \mathrm{numF}_{\max}] \\ lr_j \in [lr_{\min}, lr_{\max}] \\ L_2R_j \in [L_2R_{\min}, L_2R_{\max}] \\ LayerS_j \in [LayerS_{\min}, LayerS_{\max}] \end{cases} \tag{20}$$

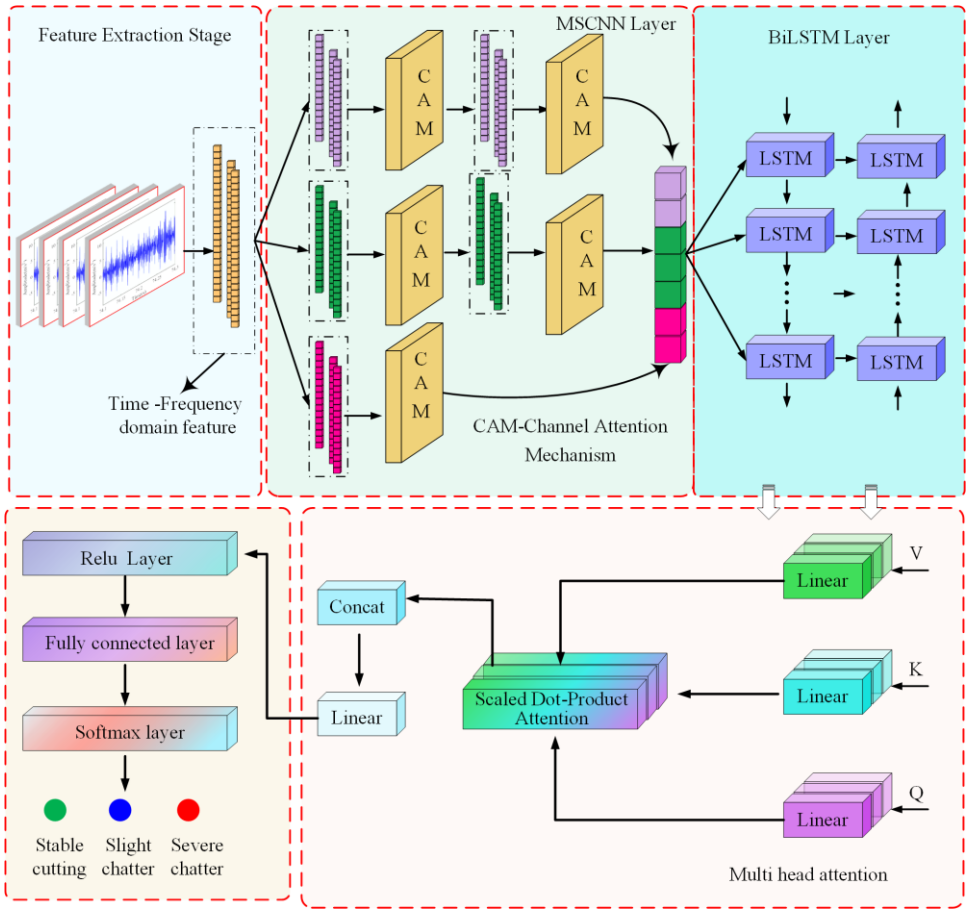
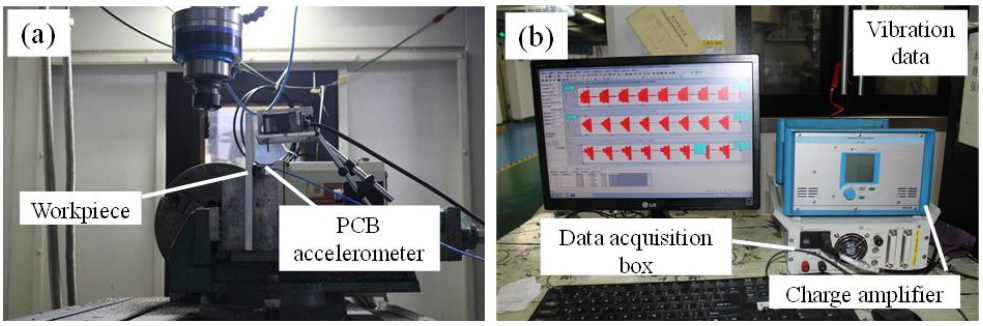


Fig. 3 MSCNN-BiLSTM- ATM structure

5. MILLING EXPERIMENT

5.1 Experimental Setup

The VDL-1000E three-axis CNC machine tool produced by Dalian Machine Tool is used. The tool is a flat-bottomed milling cutter with four teeth and a diameter of 10mm. The workpiece material is TC4 titanium alloy, and the workpiece size is 200×200×5mm. A PCB accelerometer with a sensitivity of 10.42mv/g is used, and the Donghua DH5922 acquisition system is used to collect the acceleration signals during the machining process. The sampling frequency is 5000Hz. The milling mode is down milling and dry cutting. The experimental site is shown in Fig. 4. The simulation experiment uses the Windows10 (64-bit) operating system, with hardware platform parameters of Intel(R) Core(TM) i9-12900k CPU, NVIDIA GeForce RTX3080, 3.2 GHz clock speed, and 32 GB memory. The computing environment is Matlab R2023b.



(a) Experimental site (b) Vibration data acquisition

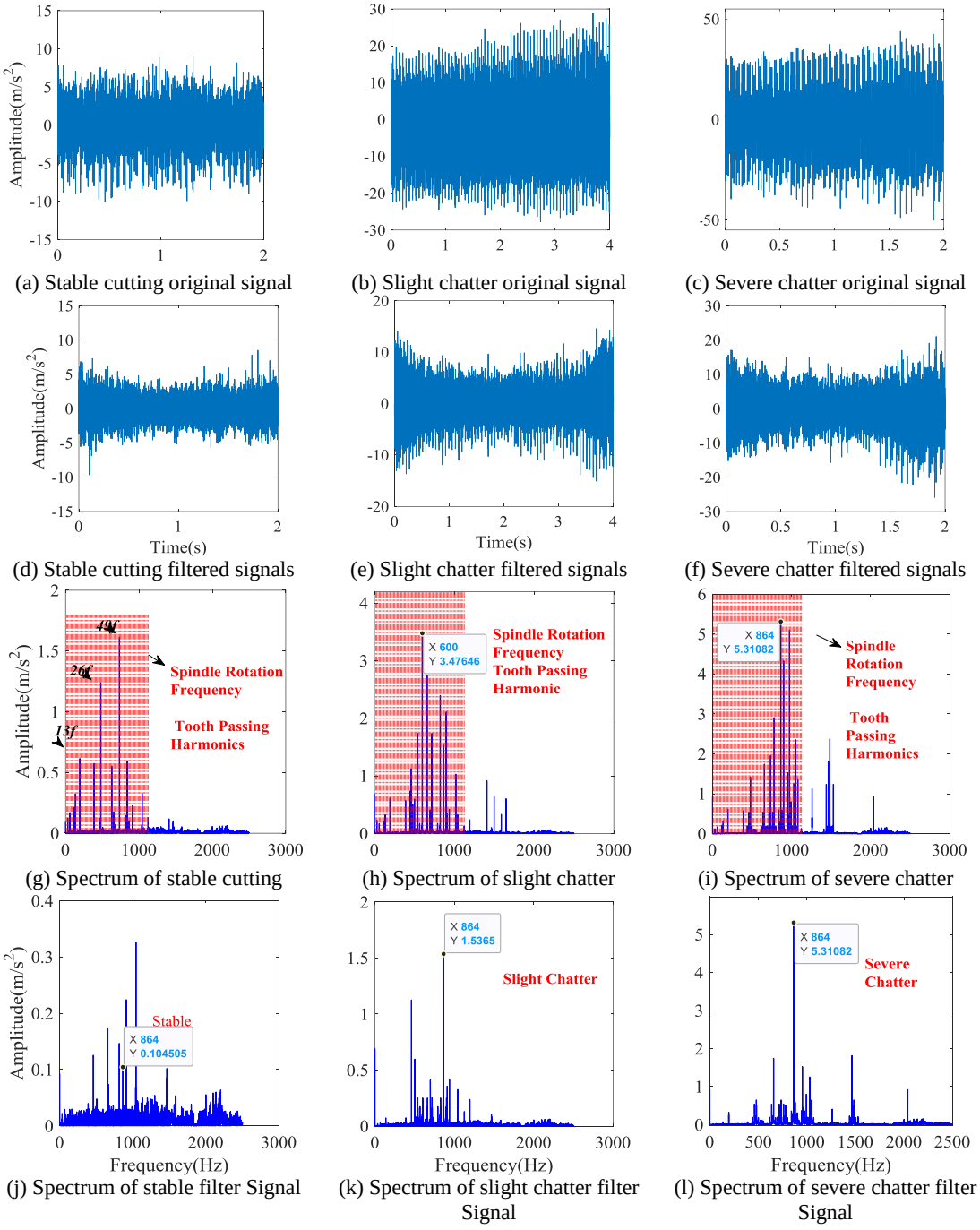
Fig. 4 Experimental setup

5.2 Experimental Parameter Settings

The modal test was conducted using a force hammer with a sensitivity of 3.41 pC/N to obtain the natural frequency of the titanium alloy thin-walled component, which was 864 Hz. The cutting parameters and machining conditions selected based on the stability lobe diagram drawn in reference [13] are shown in Table 2. The radial cutting depth is 0.5 mm, and the feed rate is 0.1 mm/tooth.

Table 2 Machining state with different conditions

Cutting parameters			Machining state	Cutting parameters			Machining state
No.	Rotating speed (r/min)	Axial depth (mm)		No.	Rotating speed (r/min)	Axial depth (mm)	
1	750	6.8	Stable	13	1050	7.2	Slight Chatter
2	750	7.2		14	1050	7.5	
3	900	5.9		15	1200	5.1	
4	900	6.3		16	1200	5.4	
5	1050	6.0		17	750	8.2	Severe Chatter
6	1050	6.4		18	750	8.5	
7	1200	4.3		19	900	7.6	
8	1200	4.7		20	900	7.8	
9	750	7.9	Slight Chatter	21	1050	8.0	
10	750	8.0		22	1050	8.2	
11	900	7.0		23	1200	5.8	
12	900	7.1		24	1200	6.5	



(a) Stable cutting original signal (b) Slight chatter original signal (c) Severe chatter original signal

(d) Stable cutting filtered signals (e) Slight chatter filtered signals (f) Severe chatter filtered signals

(g) Spectrum of stable cutting (h) Spectrum of slight chatter (i) Spectrum of severe chatter

(j) Spectrum of stable filter Signal (k) Spectrum of slight chatter filter Signal (l) Spectrum of severe chatter filter Signal

Fig. 5 Comparison of data preprocessing results

6. Result Analysis

6.1 Data preprocessing results

The SFT technique is adopted to filter the time-frequency signals obtained from milling, removing the signal components related to the rotational speed and obtaining other disturbance components. The filtering results of the acceleration signals in the X direction under the conditions of N3, N12, and N20 are shown in Fig. 5. Fig. 5(a), (b) and (c) display the original vibration signals. Fig. 5(d), (e) and (f) display the filtered vibration signals. Fig. 5(g), (h) and (i) display the FFT spectra of the original vibration signals. Fig. 5(j), (k) and (l) display the FFT spectra of the filtered vibration signals.

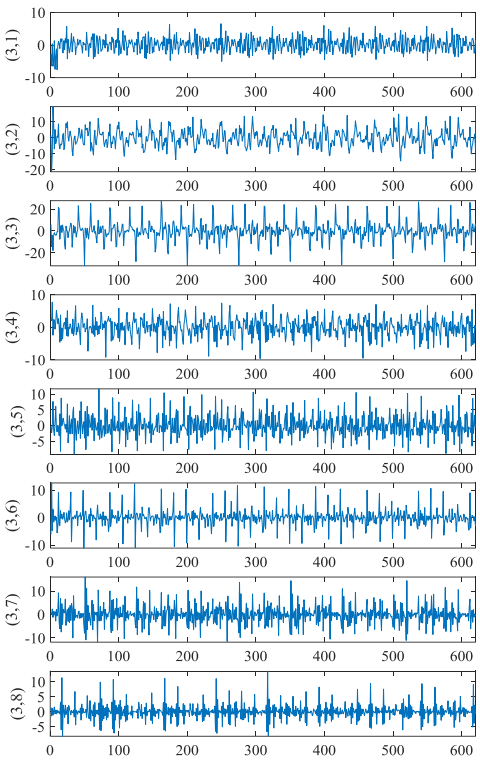
As can be seen by comparing Fig. 5(a), (b), and (c), the machining vibration gradually increases with unstable machining conditions. As shown in Fig. 5(d), (e) and (f), the amplitude of the filtered time-domain signals decreases to varying degrees. Fig. 5(g), (h), and (i) indicate that the spectra of the original stable cutting signals and slight chatter are mainly concentrated around the spindle rotational frequency and its multiples. The maximum spectral value of severe chatter is concentrated near the natural frequency of the cutting system. As shown in Fig. 5(j), (k), and (l), the SFT technique can effectively eliminate periodic components. After removing the periodic components, the chatter frequency becomes more prominent.

The Harr wavelet is selected as the basic wavelet function for wavelet packet decomposition. The original signal is decomposed into 8 frequency bands through 3-level wavelet packet decomposition. The wavelet packet decomposition results of the acceleration signal in the X direction under the N12 condition before and after signal processing are shown in Fig. 6.

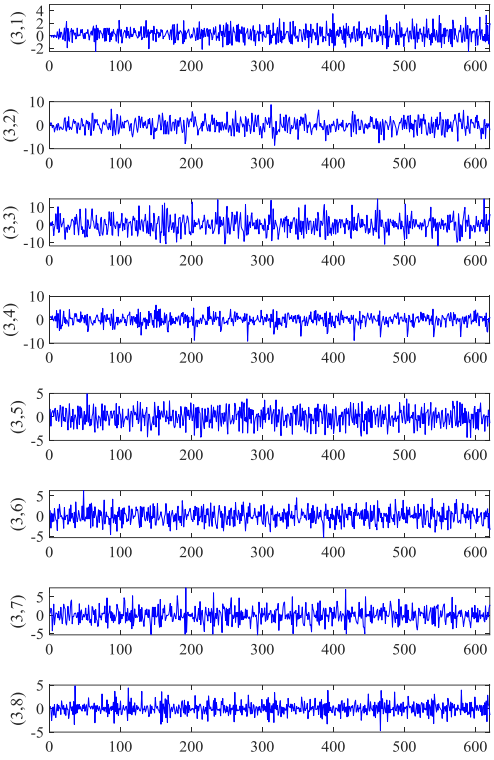
As can be seen from Fig. 6, the wavelet packet coefficients of each frequency band are inconsistent. The energy of the original signal is mainly concentrated in the frequency range of (625, 937.5], accounting for 44.94%. The energy of the original signal after removing the main frequency and its multiples is still concentrated in the frequency range of (625, 937.5], accounting for 62.20%. Removing the main frequency and its multiples from the signal enhances the significance of chatter characteristics.

The time-frequency domain analysis is conducted on the vibration signals with the main frequency and its multiples removed to obtain the time-frequency domain characteristics under different working conditions (See table 2 in detail). The box plot of some time-frequency domain characteristics is shown in Fig. 7.

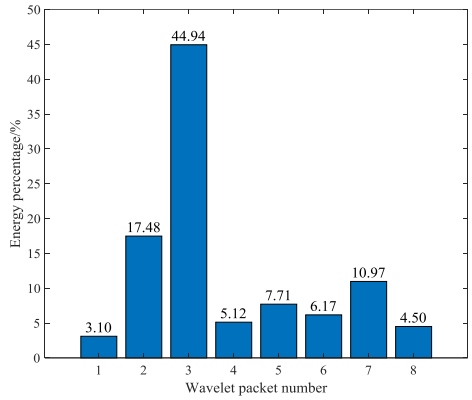
It can be seen from Fig. 7 that the five indicators of STD, RMS, ABS, SKE, MFA, SKE and KUR basically follow the same pattern, with their mean values gradually increasing as chatter occurs. The data for stable cutting and slight chatter are relatively more concentrated, while the data for severe chatter are more scattered. The two indicators of FSV and GFFS follow basically the same pattern, with their mean values gradually decreasing as chatter occurs. The data for slight chatter are relatively more scattered. The values of WAF, PTP and PAF under slight chatter and severe chatter conditions are basically similar, and are all larger than those under stable milling conditions. Therefore, based on the extracted signal characteristics and distribution patterns, it is theoretically possible to identify these three milling conditions through machine learning.



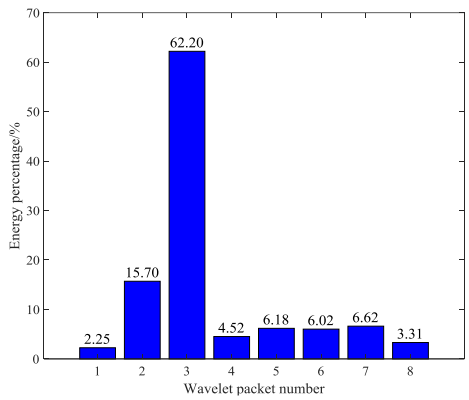
(a) Original signal wavelet packet coefficients



b) Filtered wavelet packet coefficients



(c) Proportion of original signal energy



(d) Proportion of filtered signal energy

Fig. 6 Comparison of results before and after signal processing

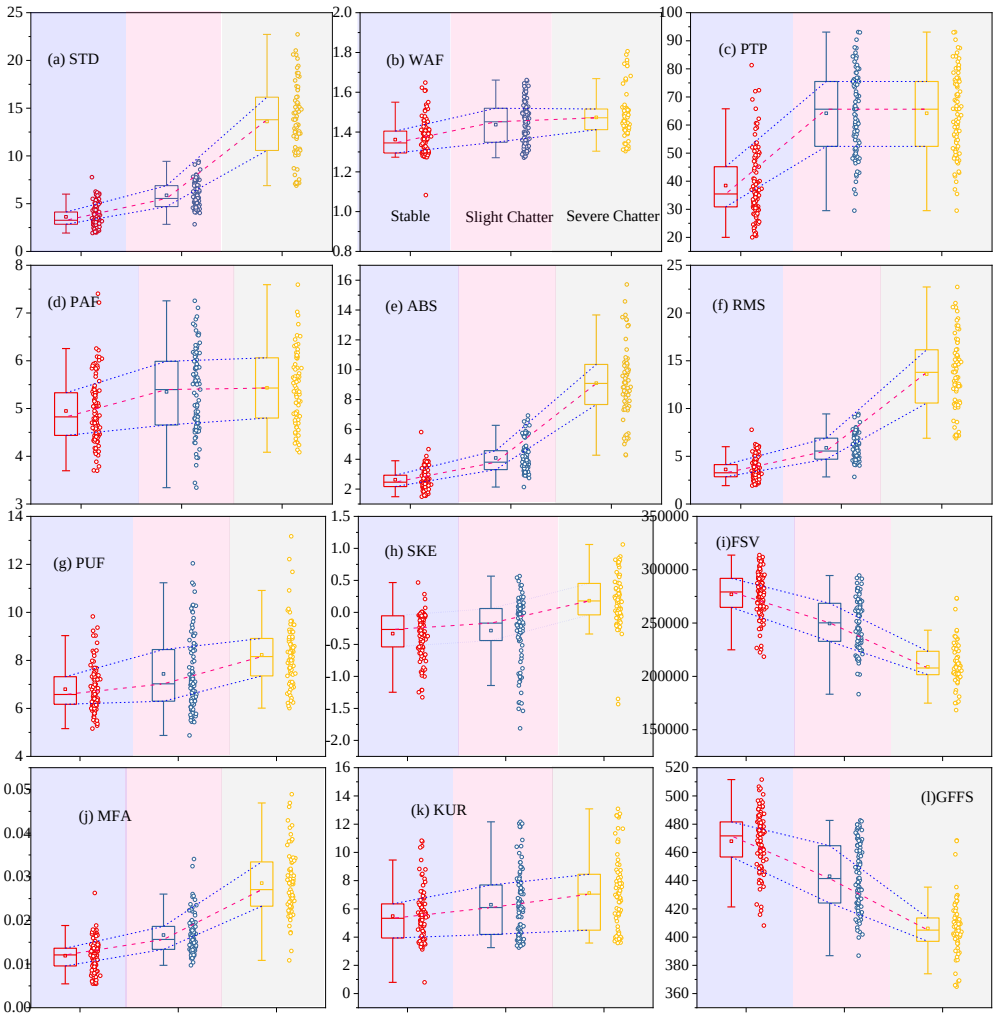


Fig. 7 Distribution pattern of time-frequency domain characteristics

6.2 Performance Analysis of MISSA

To validate the optimization performance of the MISSA algorithm, simulations were conducted on benchmark functions in comparison with Sparrow Search Algorithm (SSA) [37], Improved Grey Wolf Optimizer (IGWO) [38], Improved Sparrow Search Algorithm (ISSA5) [39], Aquila Optimizer (AVOA) [40], and Snake Optimization Algorithm (SO) [41]. This study selected F1, F2, F4, F8, F17, F18, and F22 from the CEC2005 test suite as the benchmark functions. The population size was set to 30, and the maximum number of iterations was 500. The dimensions of F1, F2, F4, and F8 were 30, while the dimensions of F17 and F18 were 2, and the dimension of F22 was 4. Each experiment was independently conducted 30 times, and three indicators, namely the optimal value (OPV), mean, and standard deviation (STD), were used to evaluate the solution accuracy and

stability of the proposed algorithm. The comparison results are presented in Table 3. The computational time of each algorithm is shown in Fig. 8.

Table 3 Comparison of algorithm accuracy and stability

Function	Index	SSA	ISSA5	IGWO	AVOA	SO	MISSA
F_1	OPV	1.34E-153	1.52E-182	5.61E-63	0	1.72E-195	0
	Mean	6.46E-84	2.106E-180	3.80E-61	0	2.34E-192	0
	STD	1.94E-83	0	5.98E-61	0	0	0
	Rank	4	3	5	**1**	2	**1**
F_4	OPV	3.5E-80	6.32E-78	5.39E-13	0	3.31E-87	6.75E-148
	Mean	2.27E-41	3.52E-76	1.512E-11	3.59E-298	2.94E-85	3.74E-130
	STD	6.73E-41	3.52E-76	2.19E-11	0	3.80E-85	1.12E-129
	Rank	5	4	6	1	3	**2**
F_8	OPV	-1063.29	-938.03	-980.2	-1063.29	-1062.52	-1063.36
	Mean	-1063.29	-870.55	-831.51	-1063.29	-1034.05	-1063.30
	STD	6.46E-4	38.32	80.83	0	34.60	2.278E-2
	Rank	3	6	5	2	4	**1**
F_{17}	OPV	397.89E-3	397.89E-3	397.89E-3	397.89E-3	397.89E-3	397.89E-3
	Mean	397.89E-3	397.89E-3	397.89E-3	397.89E-3	397.89E-3	397.89E-3
	STD	0	0	0	0	0	0
	Rank	**1**	**1**	**1**	**1**	**1**	**1**
F_{18}	OPV	3	3	3	3	3	3
	Mean	3	3	3	3	3	3
	STD	0	0	0	9.53E-8	0	0
	Rank	**1**	**1**	**1**	**1**	**1**	**1**
F_{22}	OPV	-10.40	-10.40	-10.40	-10.40	-10.40	-10.40
	Mean	-10.40293	-10.40294	-10.40294	-10.40294	-10.40133	-10.40294
	STD	4.34E-5	0	0	0	2.62E-3	0
	Rank	2	**1**	**1**	**1**	3	**1**

According to the simulation experimental results in Table 3, the MISSA algorithm demonstrates the best performance in terms of optimization stability and solution accuracy on the benchmark test functions F1, F4, F8, F17, F18, and F22. On the F1 test function, the MISSA algorithm has the same optimization capability as the AVOA algorithm, but its average computational time is only 40.41% of the AVOA algorithm (as shown in Fig. 8). On the F17 test function, all six algorithms exhibit the same optimization performance. The average computation time of the MISSA algorithm is only 25.03%, 78.16%, 5.71%, 38.15%, and 98.17% of the SSA, ISSA5, IGWO, AVOA, and SO algorithms, respectively. On the F18 test function, all six algorithms also exhibit the same optimization performance. The average computation time of the MISSA algorithm is only 26.83%, 76.84%, 5.32%,

35.01%, and 90.68% of the SSA, ISSA5, IGWO, AVOA, and SO algorithms, respectively. On the F22 benchmark test function, the MISSA algorithm has the same optimization capability as the ISSA5, IGWO, and AVOA algorithms. The average computation time of the MISSA algorithm is only 59.47%, 8.52%, and 53.18% of the ISSA5, IGWO, and AVOA algorithms, respectively. On the F4 test function, although the MISSA algorithm ranks second, its optimization capability is not significantly different from the AVOA algorithm. The average computation time of the MISSA algorithm is only 50.98% of that of the AVOA algorithm. Based on the comprehensive comparison of the experimental results, it can be concluded that the MISSA algorithm has higher convergence accuracy, better stability, and shorter computation time, which significantly outperforms the other five algorithms.

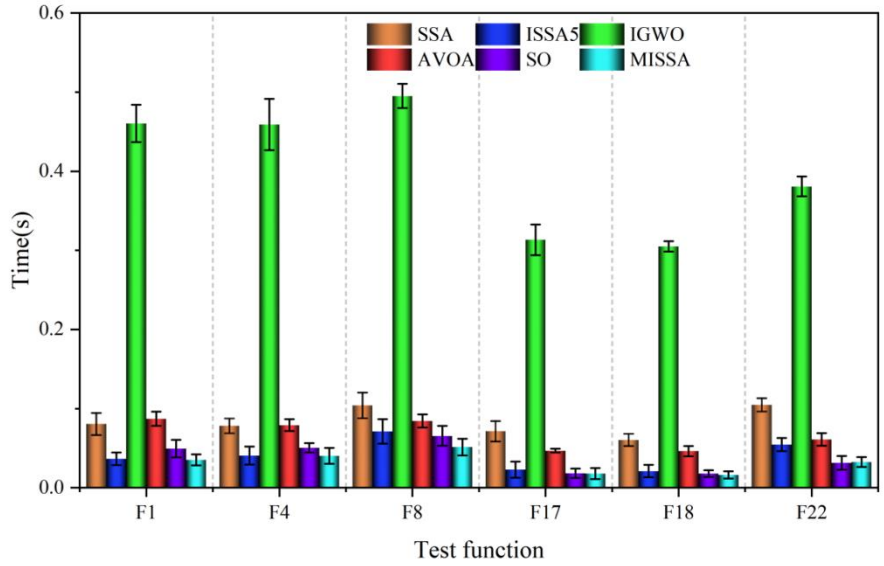
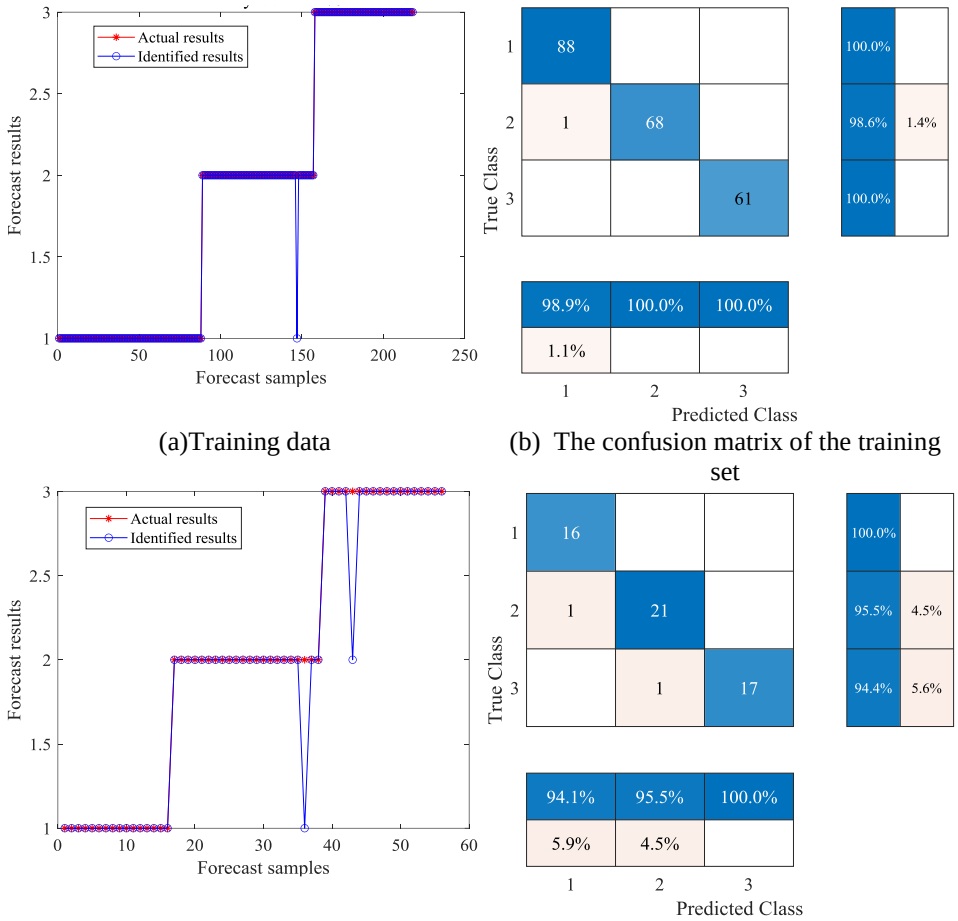


Fig. 8 Computational time of different algorithms

6.3 Detection of Milling Chatter

With a population size of 30 and a population iteration count of 15, the MISSA algorithm was used to solve Eq. (20). The optimal learning rate for the MSCNN network was found to be 0.043949. The numbers of channels for the convolutional kernels were determined to be 16, 18, 16, 32, and 32, respectively. The regularization parameter was set to 6.7562×10-4, and the number of hidden layer neurons in the BiLSTM was set to 31. All data were divided into a training set and a test set, with proportions of 80% and 20%, respectively. The recognition results for the training set and test set under the three states are shown in Fig. 9.

According to Fig. 9(a) and (b), the recognition accuracy of the test set is 99.5413%, with only one slightly chatter data being recognized as stable cutting. From Fig. 9(c) and (d), the recognition accuracy of the test set is 96.4286%, with one slightly chatter being recognized as stable cutting and one severe chatter being recognized as slightly chatter.



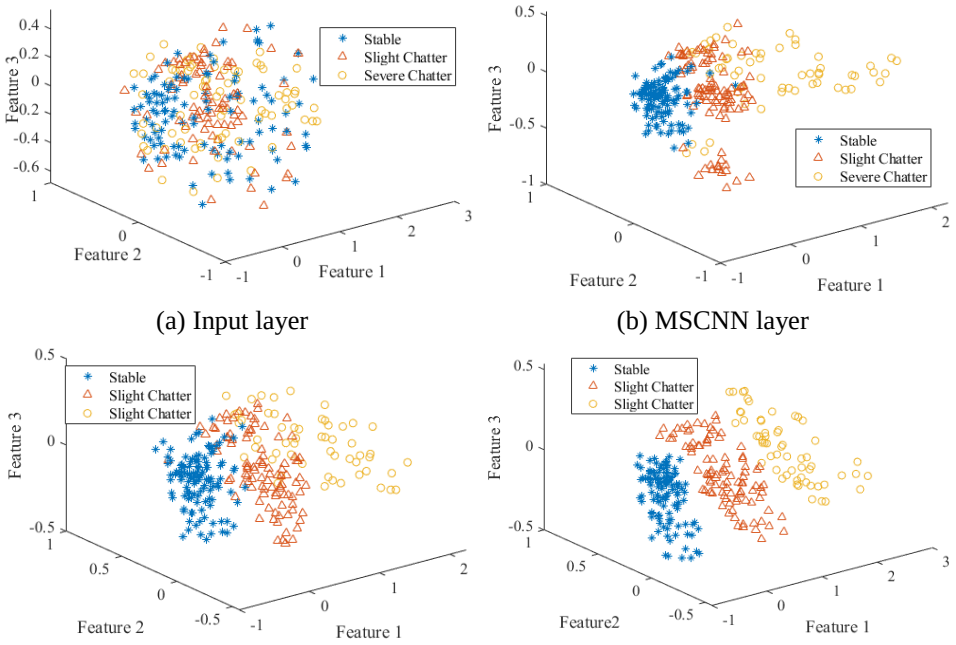
(a)Training data (b) The confusion matrix of the training set

(c)Test data (d) The confusion matrix of the test set

Fig. 9 Identification results of different processing states

The t-SNE algorithm was used to visualize the features extracted from different network layers of the MISSA-MSCNN-BiLSTM-ATM model, and the results are shown in Fig. 10. Fig. 10(a) demonstrates the t-SNE visualization effect of the test set data samples at the input layer of the model. It can be seen that various data samples are mixed with each other, resulting in poor clustering effects. Fig. 10(b) presents the visualization results after feature extraction by the MSCNN layer. Although the data is basically separated, there are still many data points mixed together, indicating poor convergence. Fig. 10(c) shows the visualization results after feature extraction by the BiLSTM layer. It can be observed that data samples of the same class have basically converged together, but there are still a few data points mixed in. Fig. 10(d) is the visualization result after feature extraction by the Atten layer. It is evident that all data samples have been completely classified into three

categories, corresponding to the three processing states. Therefore, the MISSA-MSCNN-BiLSTM-ATM model can effectively classify milling states.



(a) Input layer (b) MSCNN layer

(c) BiLSTM layer (d) MSA layer

Fig. 10 Visualization results of classification model

To validate the reliability and superiority of the model, comparisons were made with DCNN [26], Adaboost SVM [25], MSST-RF [19], and ILR-DNN [7]. Accuracy, precision, recall, F1 score, and specificity were used as performance indicators to evaluate the state classification model. To avoid errors, the average results of ten repeated experiments were taken as the final evaluation results. The comparison results for the training set are shown in Fig. 11, and the comparison results for the test set are shown in Fig. 12.

According to Fig. 11, the proposed method achieved an average accuracy of 99.48%, which was a 2.1% increase compared to DCNN, a 1.57% increase compared to Adaboost SVM, a 2.62% increase compared to MSST-RF, and a 1.05% increase compared to ILR-DNN. Additionally, the average precision of the proposed method was 99.59%, which represented a 1.83% improvement over DCNN, a 1.24% improvement over Adaboost SVM, a 2.43% improvement over MSST-RF, and a 1.2% improvement over ILR-DNN. Furthermore, the average recall rate of the proposed method was 99.43%, an increase of 2.3% compared to DCNN, 1.69% compared to Adaboost SVM, 2.73% compared to MSST-RF, and 0.9% compared to ILR-DNN. The average F1 score of the proposed method was 99.51%, showing a 2.11% improvement over DCNN, a 1.53% improvement over Adaboost SVM, a 2.6% improvement over MSST-RF, and a 1.08% improvement over ILR-DNN. Finally, the average specificity of the proposed method was 99.69%, representing a 1.11%

increase compared to DCNN, a 0.86% increase compared to Adaboost SVM, a 1.36% increase compared to MSST-RF, and a 0.44% increase compared to ILR-DNN.

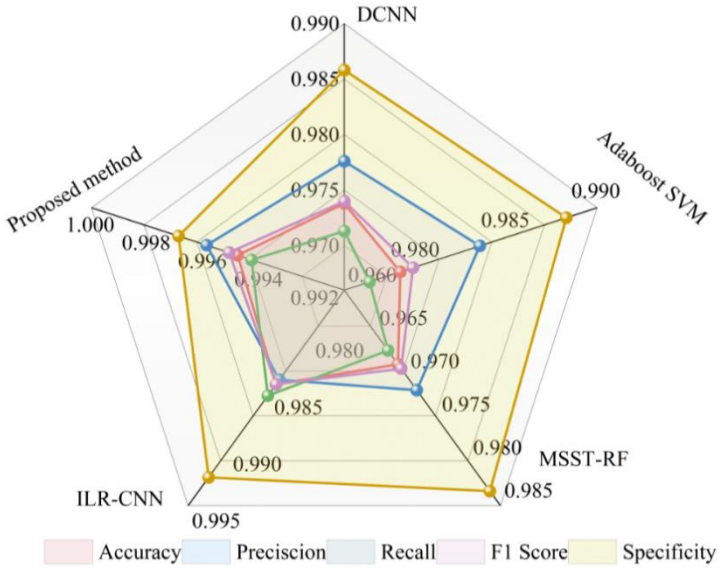


Fig. 11 Comparison of training set results

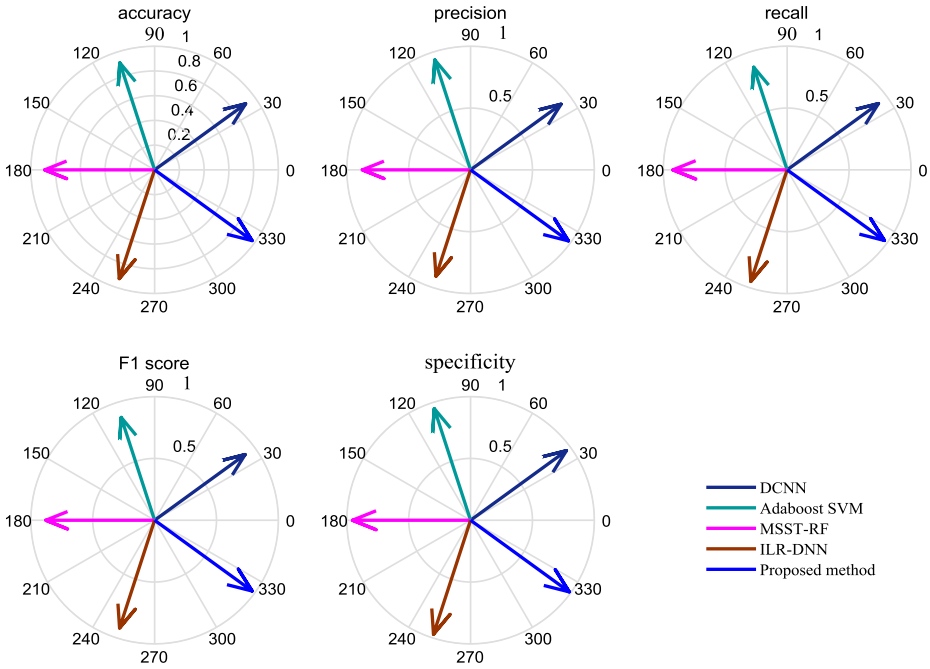


Fig. 12 Comparison of test set results

According to Fig. 12, the proposed method achieved an average accuracy of 97.38%, representing a significant improvement of 6.47% compared to DCNN, 6.29% compared to Adaboost SVM, 8.49% compared to MSST-RF, and 4.85% compared to ILR-DNN. Additionally, the average precision of the proposed method was 97.76%, an increase of 7.35% over DCNN, 4.21% over Adaboost SVM, 10.26% over MSST-RF, and 7.28% over ILR-DNN. Furthermore, the average recall rate of the proposed method was 97.13%, a notable enhancement of 5.94% compared to DCNN, 9.87% compared to Adaboost SVM, 4.82% compared to MSST-RF, and 2.26% compared to ILR-DNN. The average F1 score of the proposed method was 97.4%, showing a substantial improvement of 6.96% over DCNN, 9.93% over Adaboost SVM, 9.44% over MSST-RF, and 5.73% over ILR-DNN. Finally, the average specificity of the proposed method was 98.58%, representing a notable enhancement of 2.81% compared to DCNN, 3.58% compared to Adaboost SVM, 3.12% compared to MSST-RF, and 1.61% compared to ILR-DNN. Based on the comprehensive analysis of both the training and test sets, the method proposed in this paper achieved the best performance in terms of machining state classification.

7. CONCLUSIONS

This research first obtained the sensitive features of machining status based on SFT technology and time-frequency domain analysis methods. Then, an MISSA algorithm integrating improved circle chaos mapping, golden sine strategy, and enhanced Lévy flight strategy was proposed. Subsequently, an MSCNN-BiLSTM-ATM machining status monitoring model was constructed. Finally, the MISSA algorithm was used to optimize the hyperparameters in the monitoring model. The test and experimental results show that the proposed methods have achieved high recognition accuracy. The specific conclusions are as follows.

Using SFT technology to remove the spindle rotation frequency and its multiples from the vibration signal can effectively improve the significance of chatter features.

In the numerical experiments of CEC2005 complex functions, compared with other intelligent algorithms, the MISSA algorithm exhibited the best optimization stability and solution accuracy. The simulation running time was the lowest, only 5.32% of other intelligent algorithms.

Accuracy, precision, recall, F1 score, and specificity were used as performance indicators to evaluate the status classification model. The proposed chatter monitoring model achieved an average accuracy of 97.38%, an average precision of 97.76%, an average recall of 97.13%, an average F1 Score of 97.14%, and an average specificity of 98.58% on the test set.

Compared with DCNN, Adaboost SVM, MSST-RF, and ILR-DNN, the proposed method improved the average accuracy by up to 8.49%, average precision by up to 10.26%, average recall by up to 9.87%, average F1 Score by up to 9.93%, and average specificity by up to 3.58%.

Further validation of the proposed methodology across diverse machining conditions is necessary. In the current investigation, only vibration signals were employed for analysis, and the development of a multi-sensor feature fusion framework could potentially enhance detection precision. The conventional time-frequency feature extraction approach necessitates manual preprocessing procedures, which become labor-intensive when

handling large-scale datasets. A systematic feature importance evaluation framework can be established for automated selection of discriminative diagnostic features.

Acknowledgement: *This research was funded by the Henan Province Young Backbone Teachers Support Program in Higher Education (No.2023GGJS157), the Natural Science Foundation of Henan province (No. 222300420239), the Science and Technology Planning Project in Henan Province (No.232102241024) and Program for Innovative Research Team (in Science and Technology) in University of Henan Province (No.24IRTSTHN020).*

REFERENCES

1. Rahimi, M. H., Huynh, H. N., Altintas, Y., 2021, *On-line chatter detection in milling with hybrid machine learning and physics-based model*, CIRP Journal of Manufacturing Science and Technology, 35, pp. 25-40.
2. Jeong, K., Kim, W., Kim, N., Park, J., 2023, *Chatter detection in milling process with feature selection based on sub-band attention convolutional neural network*, The International Journal of Advanced Manufacturing Technology, 128(1-2), pp. 181-196.
3. Yesilli, M.C., Khasawneh, F.A., Otto, A., 2022, *Chatter detection in turning using machine learning and similarity measures of time series via dynamic time warping*, Journal of Manufacturing Processes, 77, pp. 190-206.
4. Matthew, D.E., Cao, H., Shi, J., 2024, *Advancing chatter detection: Harnessing the strength of wavelet synchrosqueezing transform and Hilbert-Huang transform techniques*, Journal of Manufacturing Processes, 127, pp. 613-630.
5. Li, K., He, S., Li, B., Liu, H., Mao, X., Shi, C., 2020, *A novel online chatter detection method in milling process based on multiscale entropy and gradient tree boosting*, Mechanical Systems and Signal Processing, 135, 106385.
6. Tran, M.Q., Liu, M.K., Tran, Q.V., 2020, *Milling chatter detection using scalogram and deep convolutional neural network*, The International Journal of Advanced Manufacturing Technology, 107(3), pp. 1505-1516.
7. Sun, Y., He, J., Ma, H., Yang, X., Xiong, Z., Zhu, X., Wang, Y., 2023, *Online chatter detection considering beat effect based on Inception and LSTM neural networks*, Mechanical Systems and Signal Processing, 184, 109723.
8. Aslan, D., Altintas, Y., 2018, *On-line chatter detection in milling using drive motor current commands extracted from CNC*, International Journal of Machine Tools and Manufacture, 32, pp. 64-80.
9. Zheng, X., Arrazola, P., Perez, R., Echebarria, D., Kiritsis, D., Aristimuño, P., Sáez-de-Buruaga, M., 2023, *Exploring the effectiveness of using internal CNC system signals for chatter detection in milling process*, Mechanical Systems and Signal Processing, 185, 109812.
10. Li, G., Bao, Y., Wang, H., Dong, Z., Guo, X., Kang, R., 2023, *An online monitoring methodology for grinding state identification based on real-time signal of CNC grinding machine*, Mechanical Systems and Signal Processing, 200, 110540.
11. Sestito, G.S., Venter, G.S., Ribeiro, K.S.B., Rodrigues, A. R., Silva, M.,2022. *In-process chatter detection in micro-milling using acoustic emission via machine learning classifiers*, The International Journal of Advanced Manufacturing Technology, 120(11), pp.7293-7303.
12. Bakhshandeh, P., Mohammadi, Y., Altintas, Y., Bleicger, F., 2024, *Digital twin assisted intelligent machining process monitoring and control*, CIRP Journal of Manufacturing Science and Technology, 49, pp. 180-190.
13. Gao, H., Shen, H., Yu, L., Yinling, W., Li, R., Nazir, B., 2021, *Milling chatter detection system based on multi-sensor signal fusion*, IEEE Sensors Journal, 21(22), pp. 25243-25251.
14. Tran, M. Q., Liu, M. K., Elsisi, M., 2022, *Effective multi-sensor data fusion for chatter detection in milling process*, ISA transactions, 125, pp. 514-527.
15. Zhou, G., Zhou, K., Zhang, J., Yuan, M., Wang, X., Feng, P., Zhang, M., Feng, F., 2024, *Digital modeling-driven chatter suppression for thin-walled part manufacturing*, Journal of Intelligent Manufacturing, 35(1), pp. 289-305.
16. Yin, C., Wang, Y., Ko, J. H., Lee, H. P., Sun, Y., 2024, *Attention-driven transfer learning framework for dynamic model guided time domain chatter detection.* Journal of Intelligent Manufacturing, 35(4), pp. 1867-1885.
17. Kuo, P.H., Luan, P.C., Tseng, Y.R., Yau, H.T., 2023, *Machine tool chattering monitoring by Chen-Lee chaotic system-based deep convolutional generative adversarial nets*, Structural Health Monitoring, 22(6), pp. 3891-3907.

18. Albertelli, P., Braghieri, L., Torta, M., Monno, M., 2019, *Development of a generalized chatter detection methodology for variable speed machining*, Mechanical Systems and Signal Processing, 123, pp. 26-42.
19. Yan, S.C., Sun, Y.W., 2022, *Early chatter detection in thin-walled workpiece milling process based on multi-synchrosqueezing transform and feature selection*, Mechanical Systems and Signal Processing, 169, 108622.
20. Wan, S., Liu, S., Li, X., Yan, K., Hong, J., 2023, *Milling chatter detection based on information entropy of interval frequency*, Measurement, 220, 113328.
21. Matthew, D.E., Shi, J., Hou, M., Cao, H., 2024, *Improved STFT analysis using time-frequency masking for chatter detection in the milling process*, Measurement, 225, 113899.
22. Li, D., Du, H., Yip, W.S., Tang, Y. M., To, S., 2024, *Online chatter detection for single-point diamond turning based on multidimensional cutting force fusion*, Mechanical Systems and Signal Processing, 206, 110850.
23. Wan, M., Wang, W. K., Zhang, W. H., Yang, Y., 2023, *Chatter detection for micro milling considering environment noises without the requirement of dominant frequency*, Mechanical Systems and Signal Processing, 199, 110451.
24. Thaler, T., Krese, B., Govekar, E., 2015, *Stability diagrams and chatter avoidance in horizontal band sawing*, CIRP annals, 64(1), pp. 81-84.
25. Wan, S., Li, X., Yin, Y., Hong, J., 2021, *Milling chatter detection by multi-feature fusion and Adaboost-SVM*, Mechanical Systems and Signal Processing, 156, 107671.
26. Sener, B., Gudelek, M.U., Ozbayoglu, A.M., Unver, H. O., 2021, *A novel chatter detection method for milling using deep convolution neural networks*, Measurement, 182, 109689.
27. Yesilli, M.C., Khasawneh, F.A., Otto, A., 2020, *On transfer learning for chatter detection in turning using wavelet packet transform and ensemble empirical mode decomposition*, CIRP Journal of Manufacturing Science and Technology, 28, pp. 118-135.
28. Yang, B., Guo, K., Zhou, Q., Sun, J., 2023, *Early chatter detection in robotic milling under variable robot postures and cutting parameters*, Mechanical Systems and Signal Processing, 186, 109860.
29. Lu, Y., Ma, H., Sun, Y., Liu, Z., Song, Q., 2022, *An early chatter detection method based on multivariate variational mode decomposition and chatter correlation factor*, IEEE/ASME Transactions on Mechatronics, 27(6), pp. 5724-5735.
30. Jauhari, K., Rahman, A. Z., Al, Huda, M., Azka, M., Widodo, A., Prahasto, T., 2024, *A feature extraction method for intelligent chatter detection in the milling process*, Journal of Intelligent Manufacturing, pp. 1-27.
31. Jauhari, K., Rahman, A. Z., Al, Huda, M., Azka, M., Widodo, A., Prahasto, T., 2023, *Building digital-twin virtual machining for milling chatter detection based on VMD, synchro-squeeze wavelet, and pre-trained network CNNs with vibration signals*, Journal of Intelligent Manufacturing, pp. 1-32.
32. Chen, K., Zhang, X., Zhao, W., 2023, *Automatic feature extraction for online chatter monitoring under variable milling conditions*, Measurement, 210, 112558.
33. Zhao, Y., Adjallah, K.H., Sava, A., Wang, Z., 2022, *Incipient chatter fast and reliable detection method in high-speed milling process based on cumulative strategy*, ISA transactions, 131, pp. 397-414.
34. Yesilli, M.C., Khasawneh, F.A., Mann, B.P., 2022, *Transfer learning for autonomous chatter detection in machining*, Journal of Manufacturing Processes, 80, pp. 1-27.
35. Kounta, C.A.K.A., Arnaud, L., Kamsu, F.B., Tangara, F., 2023, *Deep learning for the detection of machining vibration chatter*, Advances in Engineering Software, 180, 103445.
36. Unver, H.O., Sener, B., 2023, *A novel transfer learning framework for chatter detection using convolutional neural networks*, Journal of Intelligent Manufacturing, 34(3), pp. 1105-1124.
37. Xue, J., Shen, B., 2020, *A novel swarm intelligence optimization approach: Sparrow search algorithm*, Systems science & control engineering, 8(1), pp. 22-34.
38. Nadimi-Shahraki, M.H., Taghian, S., Mirjalili, S., 2021, *An improved grey wolf optimizer for solving engineering problems*, Expert Systems with Applications, 166, 113917.
39. Li, J., Chen, J., Shi, J., 2023, *Evaluation of new sparrow search algorithms with sequential fusion of improvement strategies*, Computers & Industrial Engineering, 182, 109425.
40. Hashim, F.A., Hussien, A.G., 2022, *Snake Optimizer: A novel meta-heuristic optimization algorithm*, Knowledge-Based Systems, 242, 108320.
41. Abdollahzadeh, B., Gharehchopogh, F. S., Mirjalili, S., 2021, *African vultures optimization algorithm: A new nature-inspired metaheuristic algorithm for global optimization problems*, Computers & Industrial Engineering, 158, 107408.