Study on Self-Learning System for Milling Chatter Feature Based on Hybrid Preprocessing Model and Improved Convolutional Clustering

Pengcheng Guo

School of Mechanical and Automotive Engineering, Xiamen University of Technology, Xiamen, 361021, China.
Fujian Province Key Laboratory of Green Intelligent Cleaning Technology and Equipment, Xiamen, 361021, China.

Sijie Cai

School of Mechanical and Automotive Engineering, Xiamen University of Technology, Xiamen, 361021, China.
Fujian Province Key Laboratory of Green Intelligent Cleaning Technology and Equipment, Xiamen, 361021, China.
Xiamen Key Laboratory of Robotics Systems and Digitization, Xiamen, 361021, China. E-mail: xmucsj@163.com

Shuixuan Chen, Song Gao and Zhihua Zhang

School of Mechanical and Automotive Engineering, Xiamen University of Technology, Xiamen, 361021, China.
Fujian Province Key Laboratory of Green Intelligent Cleaning Technology and Equipment, Xiamen, 361021, China.

Jianyu Chen

XIAMEN LOTA INTERNATIONAL CO.,LTD, Xiamen, 361021, China.

(Received 25 March 2024; accepted 25 July 2024)

Milling chatter is a self-excited vibration phenomenon during the cutting process, typically occurring in machining operations with lower stiffness. It significantly impacts the machining precision and surface morphology of the workpiece. To summarize the vibration signal features under operating conditions and enhance the precision of the intelligent recognition model for milling chatter features, this paper proposes an adaptive learning model for chatter feature recognition based on the improved convolutional clustering algorithm. This study uses wavelet analysis to extract chatter features and improved convolutional clustering to identify chatter and stable cutting signals. This method selected appropriate mapping techniques based on differences in chatter signals under various conditions and uses convolution for feature extraction. Improve convolutional clustering evaluates sample similarity through energy transfer, less influenced by signal space structure compared to Euclidean or Mahalanobis distances. This method provides a unified criterion for evaluating different signal feature representation spaces, thereby improving the accuracy and efficiency of chatter sample recognition. The recognition accuracy of chatter phenomena at different spindle speed ranges reaches 95 % and 96.3 % respectively. The results show that this method can effectively distinguish between chatter signals and stable cutting signals.

1. INTRODUCTION

Milling chatter is a phenomenon of self-excited vibration that occurs during the interaction of physical fields and the machining system in the material removal process. It often manifests in the machining of complex surfaces, especially in the manufacturing of components with low stiffness such as impeller blades. This phenomenon has a significant impact on the precision machining quality and efficiency of such components. By studying the principles of chatter, establishing Stability Lobe Diagram (SLD), and planning process parameters reasonably, it is possible to avoid the occurrence of chatter during the machining process. However, due to the influence of material hardness variations and sudden changes in the external environment, irregular chatter may occur during the machining process that does not follow established patterns. Therefore, it is significant to identify chatter caused by sudden changes. Among them, establishing an accurate and efficient chatter signal feature recognition model is a core step in constructing a chatter feature monitoring system. Researching the adaptive learning functionality of chatter signal feature recognition systems is a prerequisite for further enhancing the accuracy of the chatter feature recognition system model during operational conditions.

Chatter is classified into regenerative chatter, frictional chatter, and coupled chatter, with regenerative chatter being the most common type encountered in practical machining processes. Previous researchers, through the study of regenerative chatter, established the SLD, which is used to express the influence of spindle speed and axial depth of cut on machining stability. Some researchers have proposed a method to obtain a SLD through experiments. They design an inclined workpiece to change the axial depth of cut and deter-

mine the boundary points of the entire SLD by changing the spindle speed multiple times.[1,2] Some researchers have also studied Reliability Lobe Diagram (RLD)[3] to predict the region where chatter is likely to occur. Summarizing the signal features during milling chatter, predicting the occurrence of chatter signals through the recognition of signal feature models, and adjusting process parameters to avoid chatter phenomena are essential methods to improve milling machining accuracy and efficiency. Signals that effectively represent chatter features during machining include, acoustic pressure signals,[4] spindle acceleration vibration signals,[5,6] acoustic emission signals,[7,8] current signals,[9,10] power signals,[11] and so on. Threshold methods and pattern recognition methods are common approaches for chatter recognition.[12] In the pattern recognition method, techniques such as wavelet transform (WT),[13,14] Ensemble Empirical Mode Decomposition (EEMD)15, or Variational Mode Decomposition (VMD)16 are employed for signal decomposition to extract time-frequency domain features. These features are then combined with algorithms such as Deep Neural Networks (DNN), Convolutional Neural Net-works (CNN), Support Vector Machines (SVM), Artificial Neural Networks (ANN), etc., to construct chatter detection systems,[4,17–19] thereby improving the accuracy of chatter recognition. In the threshold method, kurtosis,[20] variance21, and Root Mean Square (RMS)[22] can be used as time-domain signal differentiation indicators for distinguishing between machining chatter and stable cutting. In the frequency domain, methods such as Teager-Huang,[23] Vold-Kalman,[24] standard deviation,[25] and others are commonly employed for chatter signal feature extraction.

Due to the lack of adaptive learning functionality in most intelligent chatter feature recognition systems, they are unable to induce data patterns during operation to correct the model. Additionally, the manual summarization of patterns and the inputting of data into the dataset involve a significant workload. To address the issues and enhance the predictive accuracy of unsupervised learning models, this paper proposes an improved convolutional clustering based unsupervised learning method applied in the chatter recognition system. This method enables labeling of collected signals through data features during the model's working process. It facilitates real-time learning of data types, updating data labels, expanding the dataset, and adjusting the boundaries of the chatter feature recognition model. Using wavelet and wavelet packet transforms to establish a hybrid preprocessing model lays the foundation for subsequent chatter identification. By adjusting the parameters of convolutional layers and pooling layers to reduce the dimensionality of the mapping space obtained from the hybrid preprocessing model, improving the accuracy of chatter feature recognition. In addition, utilizing Wasserstein distance improved the clustering method, providing a unified criterion for evaluating different signal feature representation spaces, thereby enhancing the accuracy and efficiency of chatter sample recognition.

2. SELF-LEARNING ALGORITHM MODEL FRAMEWORK FOR CHATTER FEATURE INTELLIGENT RECOGNITION SYSTEMS

The self-learning algorithm framework for the intelligent recognition system of chatter features is illustrated in Fig. 1,

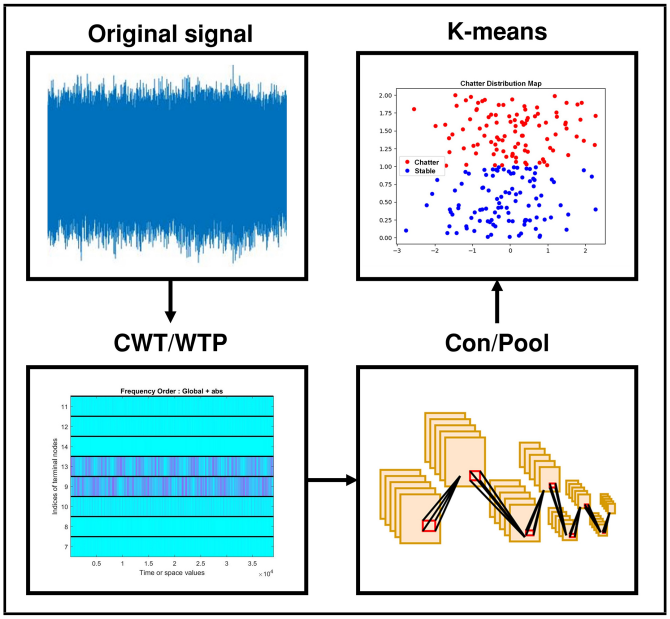

Figure 1. Intelligent recognition system flowchart.

the steps to implement the self-learning algorithm are as follows: (1). Build a vibration signal acquisition system for milling processes using an accelerometer. Establish a vibration signal dataset that includes both stable cutting and cutting chatter processes. (2). Combining wavelet/wavelet packet analysis methods to transform signal features into time-frequency domain representations. Plot the energy spectrum of the cutting vibration signal to map the characteristic signal components of chatter phenomena. (3). Further extract chatter signal features through convolutional layers and compress signal dimensions using pooling layers. (4). Utilize clustering analysis to perform unsupervised learning on chatter signal features and stable cutting signal features. Establish a unified chatter signal feature recognition model through improved convolutional clustering for different spindle speed conditions, addressing the issue of non-uniform hard threshold values for chatter features under different operating conditions. By applying this method to label the cutting experimental samples, it achieves adaptive learning and optimization of chatter signal recognition.

3. MILLING STABILITY ANALYSIS

Regenerative chatter is the critical stable state in the milling process. To study the phenomenon of regenerative chatter, it is essential to first investigate the dynamic model of the milling process. The dynamic model of the milling process is illustrated in Fig. 2, in the figure, Φ_j is the instantaneous cutting angle of the j tooth (rad), Ω is the spindle speed (r/min), ae represents the radial depth of cut (mm), v_t represents the feed direction of the tool, F_{tj} and F_{nj} are respectively the tangential cutting force (N) and radial cutting force (N) experienced by the j tooth.

Due to the presence of two degrees of freedom vibration systems in the X and Y directions on the milling system, the dynamic cutting force differential equation of the milling system is given in Eq. (1):

$$M\begin{bmatrix} x''(t) \\ y''(t) \end{bmatrix} + C\begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} + K\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} F_x \\ F_y \end{bmatrix}; \quad (1)$$

In the equation, Fx and Fy are the cutting forces on the

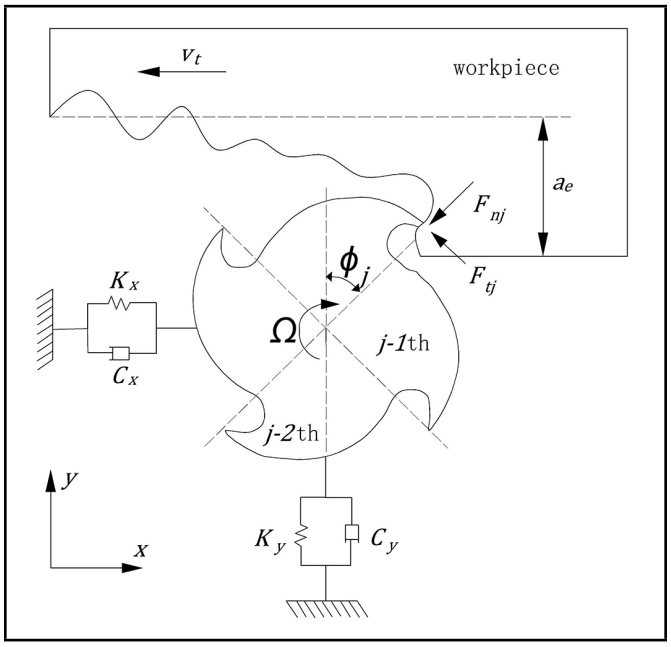

Figure 2. Milling dynamics model.

milling cutter in the X and Y directions, respectively. x, x', x" and y, y', y" represent the displacement, velocity, and acceleration on the tool, respectively. M, C, and K represent the mass, damping, and stiffness matrices, respectively.

$$M = \begin{bmatrix} m_{xx} & m_{xy} \\ m_{yx} & m_{yy} \end{bmatrix}, \quad C = \begin{bmatrix} c_{xx} & c_{xy} \\ c_{yx} & c_{yy} \end{bmatrix}, \quad K = \begin{bmatrix} k_{xx} & k_{xy} \\ k_{yx} & k_{yy} \end{bmatrix}. \quad (2)$$

Combining Eq. (1), the transfer function matrix can be either solved using the harmonic response analysis method or measured directly through modal experiments, as shown in Eq. (3).

$$G(i\omega) = \begin{bmatrix} G_{xx}(i\omega) & G_{xy}(i\omega) \\ G_{yx}(i\omega) & G_{yy}(i\omega) \end{bmatrix}; \quad (3)$$

where $G_{xx}(i\omega)$ and $G_{yy}(i\omega)$ are the direct transfer functions in x and y directions, and $G_{xy}(i\omega)$ and $G_{yx}(i\omega)$ are the cross transfer functions.

Let r be the tool tip vibration at the current time t, and r_0 be the vibration at the same phase angle from the previous tool tip pass. As shown in Eq. (4)

$$\begin{cases} r = [x(t), y(t)]^T \\ r_0 = [x(t-T), y(t-T)]^T \end{cases}. \quad (4)$$

The projections of the vibration on the x and y axes of the coordinate system are defined as $x(t)$ and $y(t)$, respectively.

Regenerative chatter studies focus on the system's unstable operating process, which can be explained as the divergent state of the system's free vibration. This vibration frequency, known as the chatter frequency, is related to the system's poles and is independent of the forced vibration (spindle speed) frequency. Let the chatter frequency be the first-order modal frequency ω_c of the machining system. The cutting force caused by chatter can be expressed as a harmonic form with frequency ω_c, be known as $F_e^{-i\omega_c t}$. Currently, the phase difference between the vibration of the previous tool tip at the same phase angle and the vibration at the current moment is $\omega_c t$. Therefore, combining the frequency response function, we obtain Eq. (5).

$$\begin{cases} r(i\omega_c) = G(i\omega 0) F_e^{i\omega_c t} \\ r_0(i\omega_c) = e^{-i\omega_c t} r(i\omega_c). \end{cases} \quad (5)$$

At this time, the projection of the uncut chip thickness caused by vibration in the coordinate system can be represented by $\Delta = [(x - x_0), (y - y_0)]^T$, as shown in Eq. (6).

$$\Delta(i\omega_c) = r(i\omega_c) - r_0(i\omega) = (1 - e^{-i\omega_c t}) e^{i\omega_c t} G(i\omega_c) F; \quad (6)$$

by substituting Eq. (6) into Eq. (7).

$$F(t) = \frac{1}{2} a_p K_t A_0 \Delta(t). \quad (7)$$

The critical condition for regenerative chatter can be obtained. At this point, the amplitude of the cutting force induced by vibration is magnified by the dynamic cutting force, indicating the onset of chatter. When both sides of the equation are equal, it represents the critical state at which regenerative chatter occurs. Using the critical chatter expression, the relationship between the critical depth of cut and spindle speed can be determined.

$$F e^{i\omega_c t} = -\frac{1}{2} a_p K_t (1 - e^{i\omega_c t}) A_0 G(i\omega_c) F e^{i\omega_c t}; \quad (8)$$

in the equation, a_p is the axial depth of cut, K_t is the tangential force cutting coefficient, A_0 is the directed dynamic milling force coefficient matrix, and $\Delta(t)$ is the dynamic displacement matrix in the X and Y directions. $F(t)$ is the dynamic milling force.

The relationship between the critical depth of cut a_p and spindle speed, also known as the stability lobe diagram, can be determined using Eq. (8).

4. STUDY ON HYBRID PREPROCESSING MODEL BASED ON SIGNAL FEATURES UNDER DIFFERENT OPERATING CONDITIONS

Based on the above method, a SLD can be established. Through the SLD, rational planning of process parameters can effectively avoid the occurrence of chatter. However, during the cutting process, sudden changes in material hardness and external environmental factors may occur, leading to situations that do not conform to the SLD. At this point, utilizing vibration signals for chatter recognition becomes crucial.

Figure 3 shows the time domain diagram of vibration signals during the milling process, representing stable cutting, weak chatter, and strong chatter conditions, respectively. As shown in the figure, the amplitude is smaller and the signal is orderly in a stable cutting state. When chatter occurs, the amplitude increases and the signal becomes chaotic. As the energy of the chatter increases, it severely affects the lifespan of the tool and the machine. To better analyze the chatter signals, this paper employs continuous wavelet transform/wavelet packet transform methods to map time-domain signals into the time-frequency domain space for expression, establishing a hybrid preprocessing model and laying the foundation for subsequent chatter feature identification.

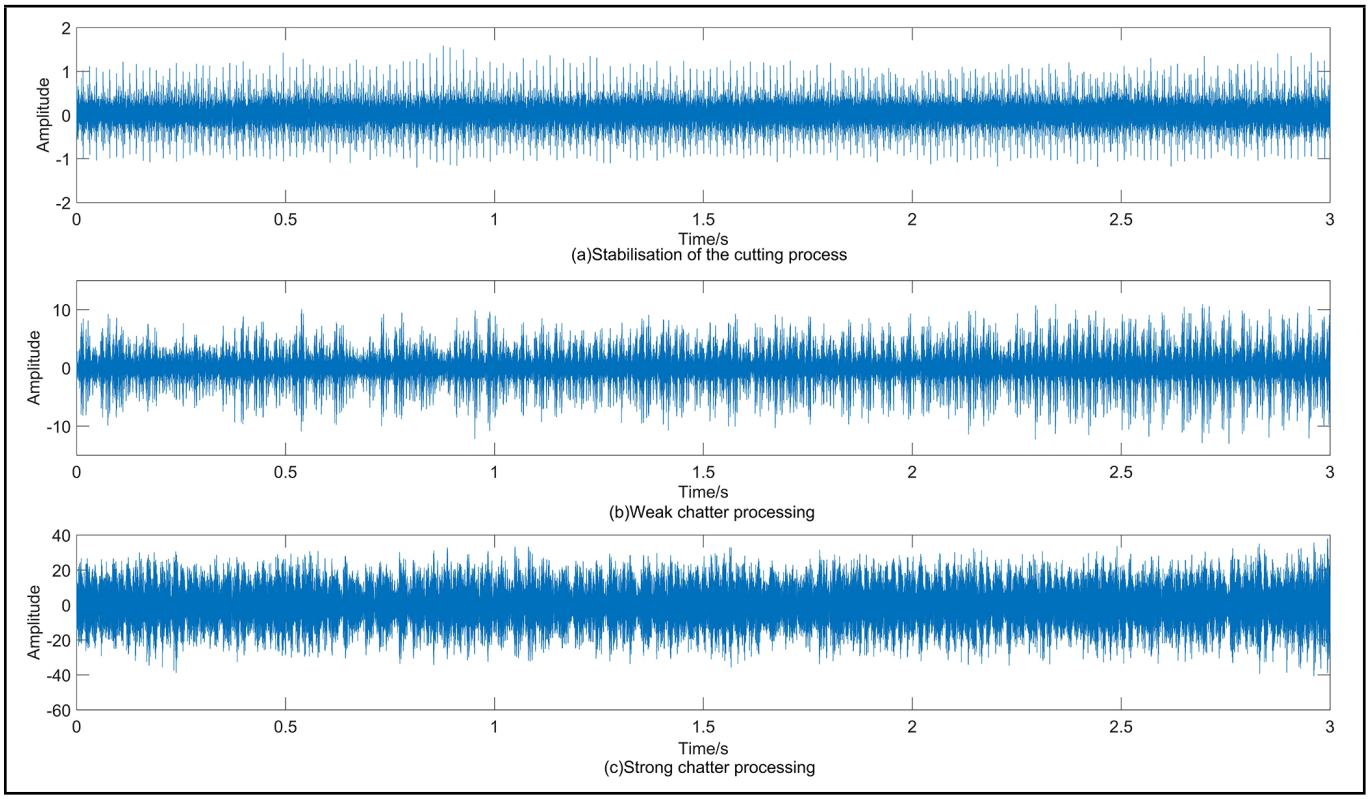

Figure 3. Time-domain diagrams of milling signals in different states.

4.1. Continuous Wavelet Transform (CWT)

The mathematical model for WT is given by Eq. (9):

$$G = W \times x; \tag{9}$$

Here, W is the wavelet basis vector, as shown in Eq. (10); x is the input signal of length N, and G is the wavelet coefficient matrix.

$$W = [\psi(a_1, b_1)\psi(a_1, b_2)L\psi(a_1, b_N)\psi(a_2, b_1)L \\ \psi(a_2, b_N)L\psi(a_M, b_N)]^T; \tag{10}$$

In the equation, a is the scale factor, b is the translation factor, M is the scale number, and N is the signal length.

The CWT is defined as shown in Eq. (11):

$$WT[f(a,b)] = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t)\psi(\frac{t-b}{a})dt; \tag{11}$$

where $f(t)$ is the original signal, and $\psi(t)$ is the wavelet basis function.

The Complex Morlet function graph is depicted in Fig. 5 (a), showing its balance in both the frequency and time domains, making it highly effective for analyzing signals with distinct frequency characteristics, particularly suitable for chatter signals. Its mathematical expression is given in Eq. (12).

$$\psi(\omega) = \frac{\sigma}{\sqrt{\pi}} e^{-\sigma^2\omega^2} e^{i2\pi f_c \omega}; \tag{12}$$

Here, ω represents time, f_c represents the central frequency, and σ represents the wavelet window width.

4.2. Wavelet Packet Transform (WPT)

Because CWT only decomposes the low-frequency band signal during signal decomposition, it has lower resolution in decomposing high-frequency band signals. WPT, on the other hand, is an optimization of CWT. In its decomposition of signals, it utilizes low-pass filter $h(x)$ and high-pass filter $g(x)$ to convolve with the signal. This process decomposes the signal into low-frequency and high-frequency regions, thereby enhancing the resolution of CWT in high-frequency segments. The formula for WPT is given by Eq. (13).

$$\begin{cases} d_l^{j,2n} = \sum_k h_{k-2l} d_k^{j-1,n} \\ d_l^{j,2n+1} = \sum_k g_{k-2l} d_k^{j-1,n} \end{cases} \tag{13}$$

Where $d_l^{j,2n}$ and $d_l^{j,2n+1}$ represent the WPT decomposition coefficients, and h_{k-2l} and g_{k-2l} are the low-pass and high-pass filter coefficients for the WPT.

In Figure 4 (a), the wavelet packet tree diagram represents the three-level wavelet packet decomposition of the acceleration signal collected during the machining process. The third-level wavelet packet decomposition divides the signal in the $0 \sim 6500$ Hz range into eight segments with equal bandwidths. Figure 4 (b) represents the time-frequency diagram of the wavelet packet, used to describe the time-frequency domain features of the signal. The vertical axis is divided into regions based on the third-level wavelet packet decomposition, with numbers from small to large corresponding to the frequency bands of the third-level wavelet packet tree from left to right nodes.

In this paper, the Meyer wavelet is chosen as the basis function for the WPT. Geometrically, the compact support in the time domain of the Meyer wavelet ensures its time-domain characteristics. Moreover, the Meyer wavelet exhibits mul-

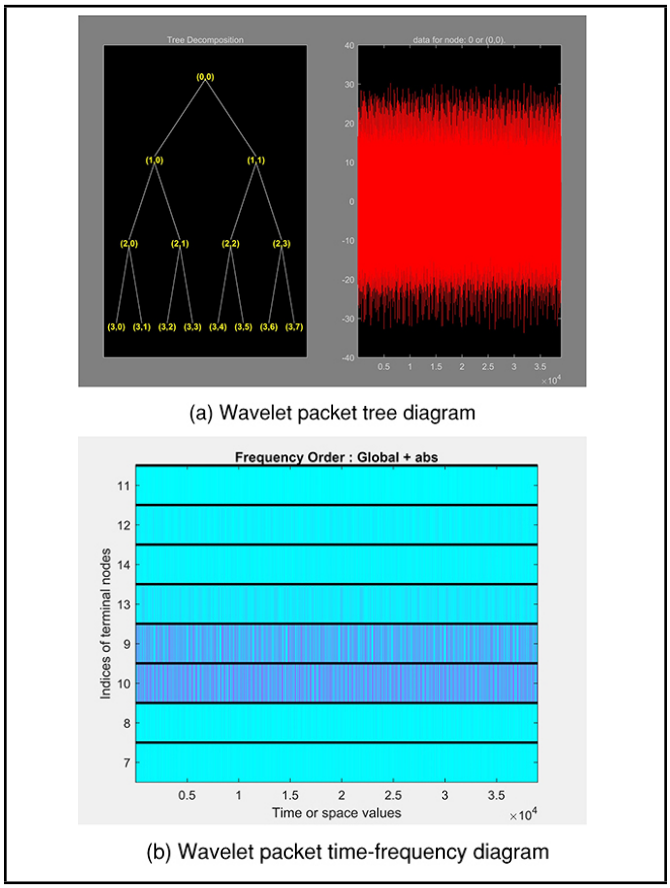

Figure 4. The wavelet packet tree and time-frequency diagram of the vibration signal.

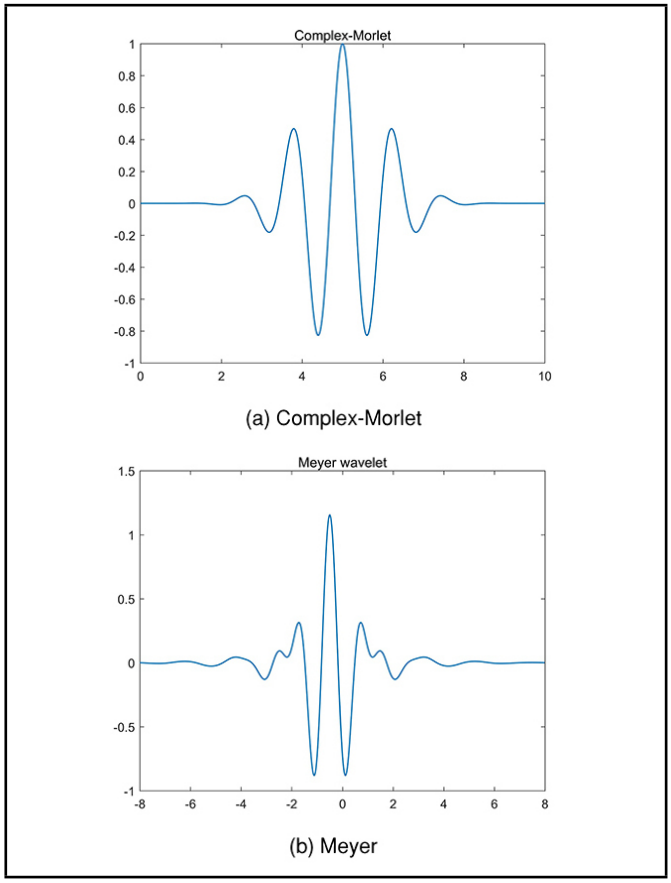

Figure 5. Wavelet basis function.

tiscale resolution characteristics, making it suitable for non-stationary signals. It allows for the decomposition and reconstruction of signals at different time and frequency scales, enabling adaptation to local variations in signal features. Thus, in the extraction of chatter features, the Meyer wavelet can effectively partition various sub-bands. The function graph is depicted in Fig. 5 (b). Meyer wavelet's Fourier transform is shown in Eq. (14).

$$\hat{\psi}(\omega) = \begin{cases} \frac{1}{\sqrt{\pi}} e^{\frac{i\omega}{e}} \sin\left[\frac{\pi}{2} v(x)\left(\frac{3}{2\pi}|\omega|-1\right)\right] & \frac{2\pi}{3} \le |\omega| \le \frac{4\pi}{3} \\ \frac{1}{\sqrt{\pi}} e^{\frac{i\omega}{e}} \cos\left[\frac{\pi}{2} v(x)\left(\frac{3}{4\pi}|\omega|-1\right)\right] & \frac{4\pi}{3} \le |\omega| \le \frac{8\pi}{3} \\ 0 & |\omega| \notin \left[\frac{2\pi}{3}, \frac{8\pi}{3}\right]; \end{cases} \tag{14}$$

where $v(x)$ satisfies:

$$v(x) = \begin{cases} 1 & 1 \le x \\ v(x) + v(1-x) = 1 & 0 \le x \le 1 \\ 0 & x \le 0 \end{cases}. \tag{15}$$

As shown in Fig. 6 (a) represents the vibration signal at 1250 r/min spindle speed, 200 mm/min feed rate, and 0.5 mm depth of cut, while in Fig. 6 (b) represents the vibration signal at 2400 r/min spindle speed, 200 mm/min feed rate, and 2.0 mm depth of cut. As shown in the figure, at spindle speeds between 1000 r/min and 3000 r/min, the difference between chatter signals and stable cutting signals is primarily evident in the low-frequency region. Since the CWT has higher resolution in the low-frequency range compared to the high-frequency range, it can better handle the distinction between chatter signals and stable cutting signals at spindle speeds between 1000 r/min and 3000 r/min. In Figure 7 (c) represents the vibration signal at a spindle speed of 4100 r/min, a feed rate of 200 mm/min, and a depth of cut of 0.5 mm. Figure 7 (d) represents the vibration signal at a spindle speed of 4100 r/min, a feed rate of 200 mm/min, and a depth of cut of 3.5 mm. At spindle speeds of 3000 r/min to 5000 r/min, the energy of natural frequency and self-excited vibration increases. There is more interference in the low-frequency region, and the difference between chatter signals and stable cutting signals in the high-frequency range is significant. CWT cannot provide the same high resolution for extracting information in both high and low-frequency bands. WPT can decompose high-frequency signals and low-frequency signals at the same level. This approach reflects the energy levels in each frequency band, distinctly distinguishing between chatter and interference information.

5. MODELING OF INTELLIGENT RECOGNITION SYSTEM SELF-LEARNING ALGORITHM BASED ON CONVOLUTIONAL CLUSTERING

5.1. Research On Feature Extraction Method Based On Standard Deviation

By employing the method of CWT/WPT, the vibration signal was transformed into a time-frequency diagram. To enable the model to distinguish between chatter and stable cutting signals in the time-frequency diagram, feature extraction was applied to the time-frequency diagram.

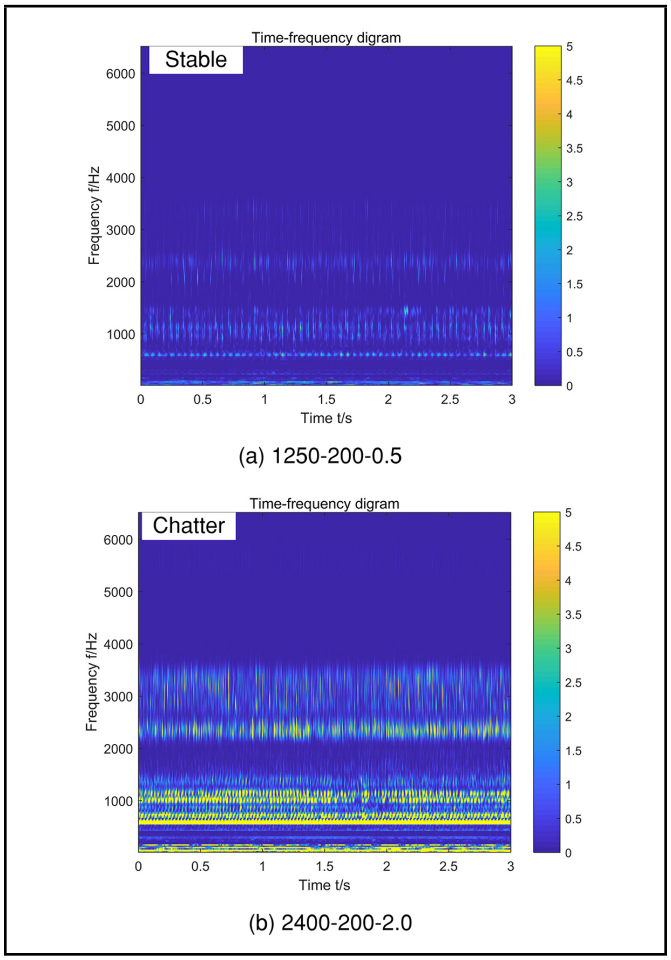

Figure 6. CWT time-frequency plot of stabilised/fibrillating signals.

Feature extraction based on the standard deviation of energy in the time-frequency diagram can help distinguish between the two signals to some extent. Mathematical formula for standard deviation is as follows.

$$S_N = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}; \tag{16}$$

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \cdots x_N}{N}; \tag{17}$$

where N was the number of pixels in the image, x_i was the energy value at the i point in the current channel. $\bar{x}$ was the average energy value of the current channel.

From the formula, it can be observed that the standard deviation was calculated based on the mean. A larger standard deviation indicates greater overall energy variation in the current channel, reflecting differences in energy across the image space. This can help distinguish between chatter signals and stable cutting signals to some extent.

In Figure 7 (a) represents the vibration signal at 1250 r/min spindle speed, 200 mm/min feed rate, and 0.5 mm axial depth of cut. The sound is steady during processing, and no chatter marks are left on the surface, indicating stable machining signals. In Figure 7 (b) shows the vibration signal at 2400 r/min spindle speed, 200 mm/min feed rate, and 2.0 mm axial depth of cut. The sound was sharp during processing, and chatter marks were visible on the surface, indicating chatter signals. Through experimental comparison, it was found that the features differences between chatter signals and stable cutting signals in the wavelet packet time-frequency diagrams were not significant at speeds between 1000 r/min and 3000 r/min. After extracting features using the standard deviation and performing clustering, the clustering accuracy was 81 %.

Figure 7 (c) represents the vibration signal at a spindle speed of 4100 r/min, a feed rate of 200 mm/min, and an axial cutting depth of 0.5 mm. The sound during processing was stable, and the surface shows no chatter marks, indicating a stable cutting signal; Fig. 7 (d) represents the vibration signal with a 3.5 mm axial depth of cut under the same conditions. The sound during processing was sharp, and the surface exhibited chatter marks, indicating chatter. Through experimental comparison, it was found that in the speed range of 3000 r/min to 5000 r/min, the differences in the wavelet packet time-frequency diagram are apparent. However, utilizing standard deviation for feature extraction and subsequent clustering was limited. Since standard deviation can only assess global differences to determine features and cannot directly discern the spatial distribution of energy, the clustering accuracy is only 74.15 %.

To better extract chatter signal features and improve the accuracy of clustering algorithms, a method capable of recognizing the distribution of energy in the time-frequency domain space is needed. In this paper, we introduce a feature extraction method based on convolutional layers and pooling layers.

5.2. Feature Extraction Method Based On Convolutional And Pooling Layers

To compensate for the insufficient feature extraction capability of the standard deviation method, this paper proposes a data preprocessing method based on convolutional and pooling layers. The method of using convolutional and pooling layers transformed the time-frequency diagram into a low-frequency space that encapsulated the chatter features. In this space, certain basis vectors can represent the features of chatter signals. At this time, chatter features were minimally affected by other interference information. The boundary between chatter samples and stable cutting samples became clearer in the space, thereby extracting effective chatter features. The convolutional layer traverses the image with different-sized convolutional kernels (filters) to obtain feature information from the image. The convolution operation is shown in Eq. (18).

$$f(z) = \sum_{j=1}^{L}\sum_{i=1}^{C} Z_{i,j} \times w_{i,j} + b; \tag{18}$$

In the equation: L and C represented the length and width of convolutional kernels, respectively; $Z_{i,j}$ represented the eigenvalue of the input matrix at position (i, j); $w_{i,j}$ and b represented the weight and bias of the convolutional kernel, respectively; $f(z)$ represented the output obtained from the convolution.

The convolutional kernel traversed the input image with the set stride to obtain a new feature map. The convolution operation involves multiplying and adding the corresponding elements of the input image and the convolutional kernel. Figure 8 (a) it can be observed that the input image undergoes convolutional operation with a 3×3 kernel to generate the resulting feature map. To control the size of the output image, padding with zero values can be applied to the outer layer of

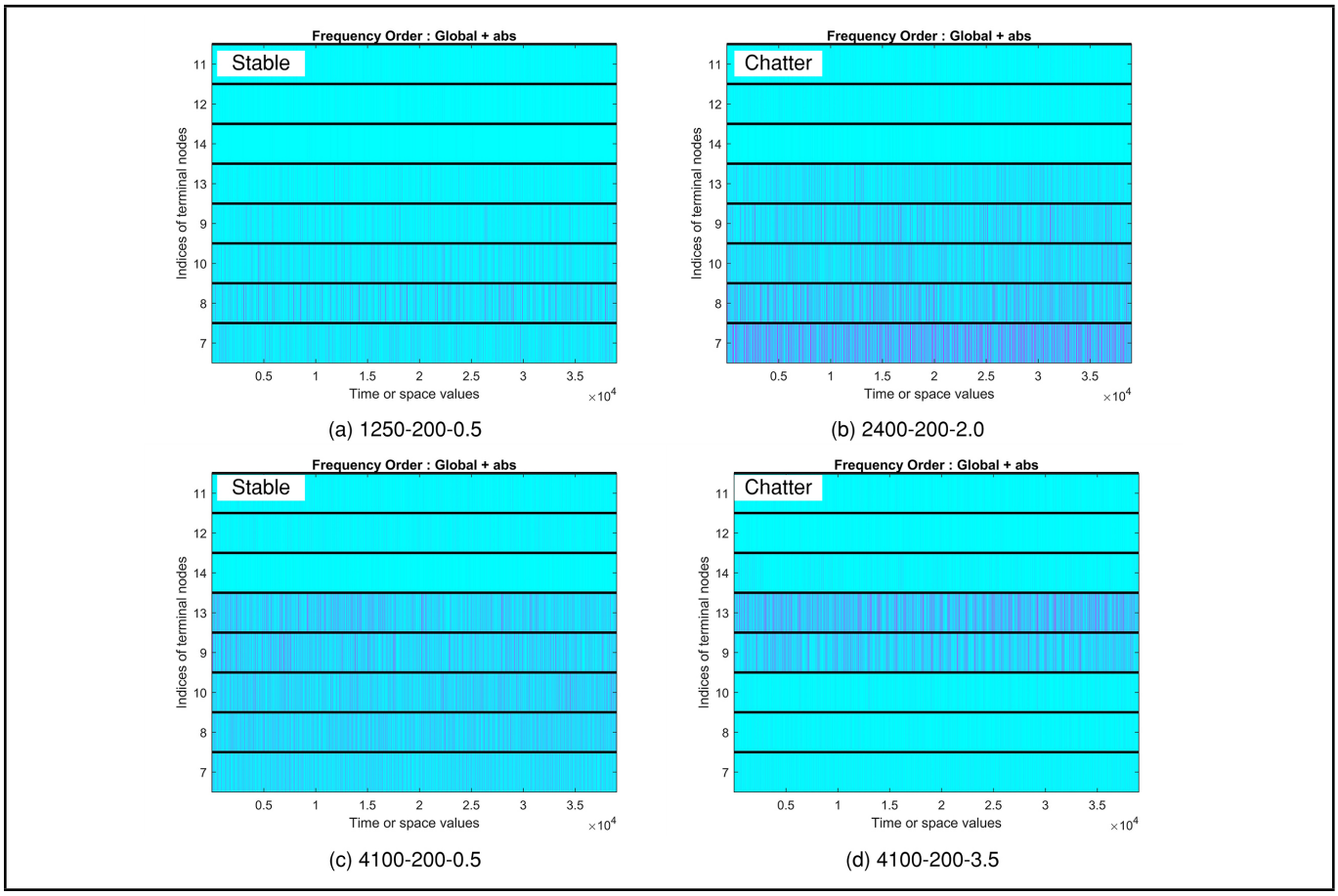

Figure 7. WPT Time-Frequency diagram under variable speed/cutting depth.

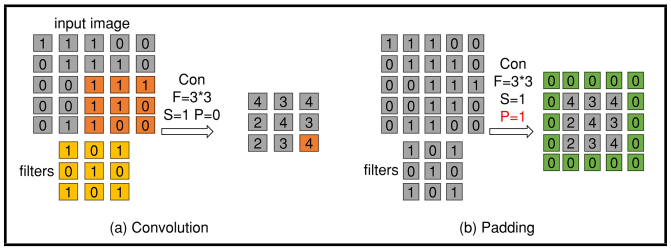

Figure 8. Convolution and edge padding.

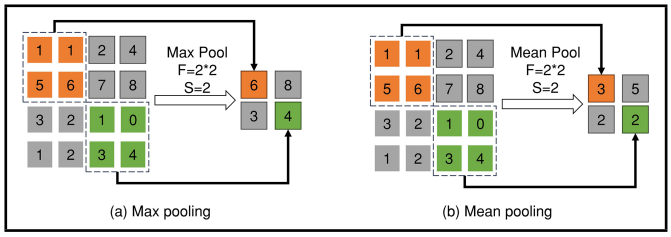

Figure 9. Pooling operation.

the output image. The size of the output image after convolution can be calculated using Eq. (19) and Eq. (20):

$$H = \frac{H_1 + 2P - F}{S} + 1; \tag{19}$$

$$W = \frac{W_1 + 2P - F}{S} + 1. \tag{20}$$

The formula relates the size of the output image H and W after convolution with the size of the input image H_1 and W_1, where P represented the padding amount, determining the number of zeros added to the output image, S is the stride, and F was the width of the convolution kernel.

As shown in Fig. 8 (b), edge padding resolves the issue of reduced image size after convolution. When padding was set to 1, the convolved image was surrounded by an additional layer of zeros.

The main purpose of the pooling layer is to reduce the dimensions of the feature maps, thereby decreasing the size of the image while retaining its important features. Like the convolutional layer, the pooling layer also employs a filter to perform down sampling on the feature maps. However, the filter in the pooling layer does not involve data in the calculation, which helps prevent overfitting to some extent. The common pooling methods include max pooling and mean pooling. Max pooling is illustrated in Fig. 9 (a), where a filter traverses the feature map, extracting the maximum feature value and placing it in the corresponding position of the new feature map. The specific operation is shown in Eq. (21).

$$f(x) = \max\{x_n | n = 1, 2, 3 \cdots i\}. \tag{21}$$

Mean pooling, as shown in Fig. 9 (b), computes the mean value of the feature map to generate the output. The calculation is expressed by Eq. (22).

$$f(x) = \text{mean}\{x_n | n = 1, 2, 3 \cdots i\}; \tag{22}$$

In the equation: x_n was the activation value of the neuron; i was the number of elements in the current pooling kernel; $f(x)$ represents the output of the pooling layer.

From Fig. 10, it can be observed that feature maps can undergo feature extraction and compression through two pathways. When clustering the signals from the speed range of

1000 r/min to 3000 r/min after continuous wavelet decomposition, using the second pathway, features are compressed to a dimension of $1 * 32$ through convolution and pooling. In the remaining spindle speed ranges and data processing methods, the first pathway is used to compress the features to a dimension of $1 * 64$.

5.3. Improved K-Means Algorithm Based Self-Learning Method For Chatter Features

This paper employs the K-means algorithm for unsupervised learning on one-dimensional feature vectors processed by convolutional and pooling layers. The algorithm steps are as follows. (1) Randomly select two samples from the dataset as the initial centers of two clusters; (2) According to the Wasserstein distance formula, calculate the distance between each sample point in the dataset and the centers of the two clusters. Assign each sample point to the cluster center closest to it, dividing the dataset into two clusters; (3) Using the sample points in each cluster, calculate the centroid of each cluster based on the mean, where the obtained centroids are not required to be within the dataset; (4) Repeat steps 2 and 3 until the sample points in the clusters no longer change, then stop the iteration and output the classification results.

Traditional K-means algorithm uses Euclidean distance to calculate the distance between each sample point and the cluster center. As shown in Eq. (23).

$$dis(x_i, c_j) = \sqrt{\sum_{t=1}^{m}(x_{it} - c_{jt})^2}; \tag{23}$$

In the formula, x_i represents the i-th object, c_j represents the j-th cluster center, $1 \leq j \leq k$; x_{it} represents the t-th attribute of the i-th object, $1 \leq t \leq m$, c_{jt} represents the t-th attribute of the j-th cluster center.

In high-dimensional space, sample distribution becomes sparse, and the basis vectors in each dimension are not necessarily orthogonal to each other. This phenomenon leads to the uncertainty of the Euclidean distance criterion, making the Euclidean distance unreliable in high-dimensional space. Compared to Euclidean distance, the Wasserstein distance quantifies the minimum cost of transforming one sample into another, identifying the similarity between samples through this minimum cost. Thus, it can comprehensively assess the distance between two samples by considering the differences across multiple dimensions, making it suitable for determining the similarity of samples in various dimensional spaces, particularly in high-dimensional spaces. Because the convolutional results of images are typically high-dimensional samples, this paper adopts the Wasserstein distance to enhance the K-means algorithm.

The transformation principle of the Wasserstein distance is illustrated in Fig. 11. The original vector is denoted as P, and the target vector is denoted as Q. The original vector P is transformed into a joint vector α with the target vector Q through the inverse matrix F_Y^{-1}. Subsequently, using the matrix F_x, it is transformed back into the target vector Q. The goal was to minimize the distance between F_Y^{-1} and F_x, making the transformation cost minimal. When this condition is met, the two vectors are considered most similar. The distance between P and Q at this point is given by:

$$W(P, Q) = ||F_x^{-1} - F_Y^{-1}||_p. \tag{24}$$

The specific process for calculating the distance is as follows: $P = \{(l_i, w_i)|i = 1, 2, \cdots, n\}$ was a one-dimensional vector with n data samples, where li and wi were the coordinates and weights of the i-th element of P. $Q = \{(l_j, w_j)|j = 1, 2, \cdots, m\}$ was a one-dimensional vector with m data samples, where l_j and w_j were the coordinates and weights of the j-th element of Q. To minimize the overall migration cost, it is necessary to find a flow function $F = (f_{ij})_{n*m}$, where f_{ij} represents the amount to be moved between l_i and l_j. The minimum cost is:

$$WD(P, Q, F) = \sum_{i=1}^{n}\sum_{j=1}^{m} d_{ij} f_{ij}. \tag{25}$$

Subject to the following constraints:

$$f_{ij} \geq 0, i = 1, 2, \cdots, n, j = 1, 2 \cdots m; \tag{26}$$

$$\sum_{j=1}^{m} f_{ij} \leq w_i, i = 1, 2, \cdots n; \tag{27}$$

$$\sum_{i=1}^{n} f_{ij} \leq w_j, j = 1, 2, \cdots m; \tag{28}$$

$$\sum_{i=1}^{n}\sum_{j=1}^{m} f_{ij} = \min(\sum_{i=1}^{n} w_i, \sum_{j=1}^{m} w_j). \tag{29}$$

Here, d_{ij} represents the distance between l_i and l_j.

The optimal $F* = (f*_{ij})_{n*m}$ is the sought-after Wasserstein distance and is found as shown in Eq. (30):

$$WD(P, Q) = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} d_{ij} f_{ij}^*}{\sum_{i=1}^{n}\sum_{j=1}^{m} f_{ij}^*}. \tag{30}$$

Sequentially compare the distance from each object to each cluster center, assign the object to the cluster of the nearest center, and obtain k clusters $\{s_1, s_2, \cdots, s_k\}$. The subsequent cluster determines its new centroid by averaging all its internal points, with the mean calculated as in Eq. (31):

$$c_1 = \frac{\sum_{i=1}^{|S_l|}}{|s_l|} x_1; \tag{31}$$

In the formula, c_l represents the centroid of the l-th cluster, where $1 \leq l \leq k$, and $|s_l|$ represents the number of objects in the l-th cluster.

6. MILLING CHATTER EXPERIMENT

Figure (12) illustrates the surface topography, time-domain signals, and time-frequency domain signals during the milling process at different spindle speeds and axial depths of cut. From the figure, it can be observed that stability is maintained with small axial depths of cut, while the tendency for chatter increases with larger axial depths of cut. Based on the experimental results in the figure, the approximate critical chatter

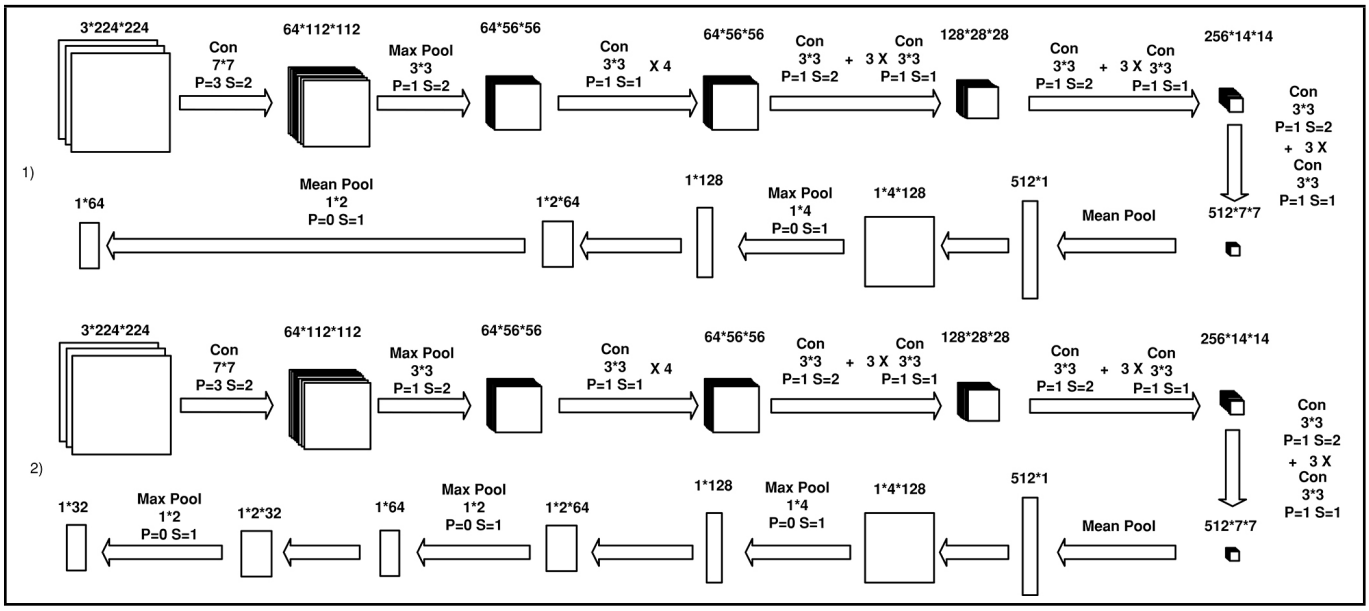

Figure 10. Improved convolutional clustering network structure.

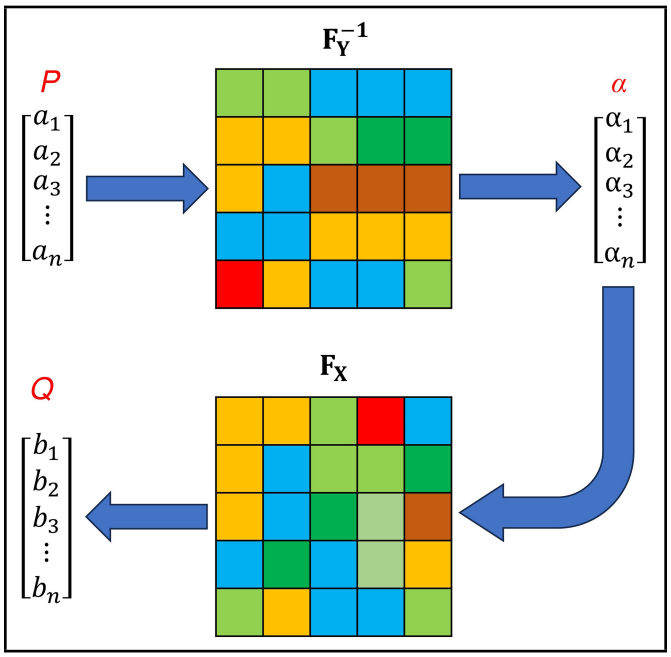

Figure 11. Wasserstein distance principle.

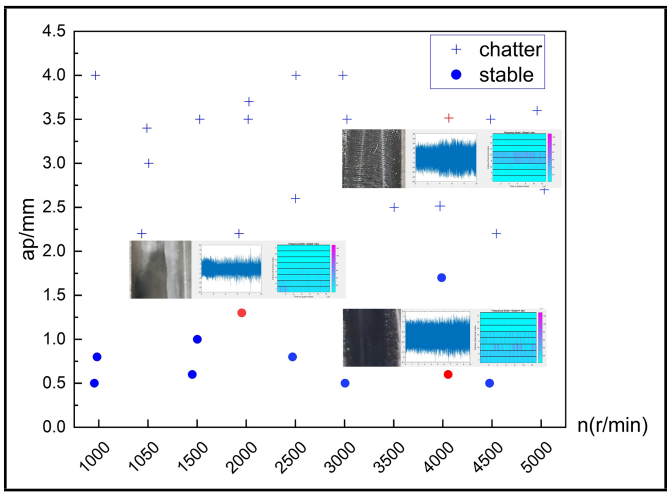

Figure 12. Distribution diagram of chatter and stable cutting samples.

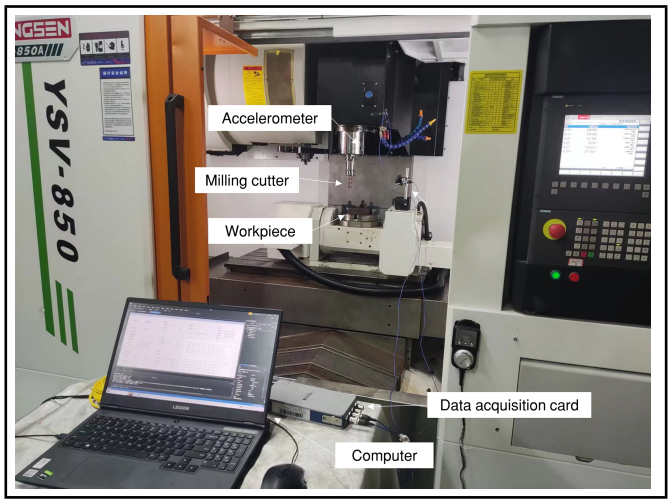

Figure 13. Milling experimental platform.

positions can be determined, guiding the planning of experimental parameters.

The milling experiment, as shown in Fig. 13, consists of a CNC machine tool, milling cutter, workpiece, acceleration sensor, data acquisition card, and computer. The tool was a tungsten steel vertical milling cutter with a hardness of 60 and four teeth. The tool was suitable for a speed range of 1000 r/min to 5000 r/min. The acceleration sensor used was the PCB 352C03 type sensor, with a sensitivity of 9.77 mV/g; The data acquisition card used is the NI-9234 type acquisition card; The sampling frequency was set to 13000 Hz. The workpiece has dimensions of 100 mm ∗130 mm ∗25 mm and is made of 45# modulation steel. Based on the applicable speed range of the tool and the performance of chatter at different speeds, the experimental groups are divided. The experimental groups are categorized into speed ranges of 1000 r/min to 3000 r/min and 3000 r/min to 5000 r/min.

By planning experimental parameters, milling experiments were conducted to validate the accuracy of the algorithm using experimental data. 300 sets of experimental samples were configured in the range of 1000 r/min to 3000 r/min, with 150 sets as chatter samples and 150 sets as stable cutting samples. In

the range of 3000 r/min to 5000 r/min, 325 sets of experimental samples were configured, with 220 sets as chatter samples and 105 sets as stable cutting samples.

Figure 14 illustrates the variation in the number of iterations under the influence of different initial cluster centers. Whereas a and b form one pair; c and d constitute another pair. The Figures 14 (a) and 14 (c) represent the sum of distances from the final cluster members to the center at each iteration. As shown in Fig. 14 (a) when there is a significant difference between the initial cluster centers and the actual cluster centers, it leads to an increase in the number of iterations, thus seeking the real cluster centers. When the initial cluster centers are chosen more reasonably, as shown in Fig. 14 (c) the number of iterations significantly decreases. The Figures 14 (b) and 14 (d) represent the distance between adjacent cluster centers. When the initial cluster centers are poorly chosen, with changes in cluster membership, the cluster centers may experience significant movement. The distance between adjacent cluster centers can fluctuate greatly but tends to stabilize with iterations until the distance between two points becomes 0, as shown in Fig. 14 (b). After 13 difference calculations, the distance between the two cluster centers in this cluster becomes 0. When the initial cluster centers are chosen more reasonably, there is less variation in cluster membership, and the movement of cluster centers is smaller. The movement between adjacent cluster centers is relatively smooth, as shown in Fig. 14 (d). After 5 difference calculations, the distance between the two cluster centers becomes 0.

Perform feature extraction on the experimental group data using standard deviation and convolutional layers with pooling layers, respectively. Compare the accuracy of the algorithm after feature extraction using these two methods. According to Tab. 1, it can be observed that in the feature extraction methods based on CWT and standard deviation, the accuracy reaches 95 % at speeds ranging from 1000 r/min to 3000 r/min. This indicates that at this speed range, there is a clear boundary in the overall energy distribution between chatter signals and stable cutting signals. However, at speeds ranging from 3000 r/min to 5000 r/min, the accuracy of standard deviation is 89.84 %; In the WPT, the accuracy of the standard deviation feature extraction method at spindle speeds ranging from 1000 r/min to 3000 r/min is 81 %; At spindle speeds ranging from 3000 r/min to 5000 r/min, the accuracy is 74.15 %, reflecting that the overall energy distribution in the time-frequency diagram at this speed range does not effectively differentiate between chatter signals and stable cutting signals. In the feature extraction method based on CWT and improve convolutional clustering, the accuracy is 95 % at spindle speeds ranging from 1000 r/min to 3000 r/min, and 82.46 % at spindle speeds ranging from 3000 r/min to 5000 r/min. However, in the feature extraction method combining WPT and improve convolutional clustering, the accuracy is 68.66 % at spindle speeds ranging from 1000 r/min to 3000 r/min. This reflects that the difference between chatter signals and stable cutting signals in the high-frequency range is not significant at this speed range. At spindle speeds ranging from 3000 r/min to 5000 r/min, the accuracy is 96.3 %, indicating that at this speed range, the difference between chatter signals and stable cutting signals is mainly distributed in the high-frequency region. From the results, it can be observed that CWT resolves chatter signal features based on overall energy magnitude. The signal analysis using CWT at spindle speeds ranging from 1000 r/min to 3000 r/min is more effective than WPT. On the other hand, WPT resolves chatter features by concentrating on high-energy clusters. At spindle speeds ranging from 3000r/min to 5000r/min, WPT is superior to CWT due to the clear concentration of chatter energy.

Table 1. Clustering accuracy of different feature extraction methods.

Feature extraction methods	Spindle speed n(r/min)	Accuracy
Standard deviation (CWT)	1000∼3000	95 %
Standard deviation (CWT)	3000∼5000	89.84 %
Standard deviation (WPT)	1000∼3000	81 %
Standard deviation (WPT)	3000∼5000	74.15 %
Improve convolutional clustering (CWT)	1000∼3000	95 %
Improve convolutional clustering (CWT)	3000∼5000	82.46 %
Improve convolutional clustering (WPT)	1000∼3000	68.66 %
Improve convolutional clustering (WPT)	3000∼5000	96.3 %

Table 2. Model accuracy with different distance formulas.

Formulas	Spindle speed n(r/min)	Accuracy
Euclidean distance	3000∼5000	90.04 %
Wasserstein distance	3000∼5000	96.3 %

Based on the above conclusions, this study adopts the CWT method to process samples in the spindle speed range of 1000 r/min to 3000 r/min and utilizes the WPT method to process samples in the spindle speed range of 3000 r/min to 5000 r/min. The improve convolutional clustering method is employed for signal feature judgment, establishing a chatter feature recognition system.

Due to the error in distance measurement in high-dimensional space, caused by Euclidean distance, the clustering results are inaccurate. Therefore, Wasserstein distance is used as a replacement for Euclidean distance in the clustering algorithm's distance measurement formula, improving the accuracy of the algorithm, as shown in Tab. 2.

When stable cutting samples are mistakenly classified as chatter samples and chatter control is applied, the impact on the machining quality is relatively small. However, when chatter samples are mistakenly classified as stable cutting samples and no chatter control is applied, the impact on the machining quality is significant. In Tab. 3, recall rate for chatter samples was calculated using Eq. (32).

$$R = \frac{TP}{TP + FN}; \tag{32}$$

R denotes the chatter samples recall; TP represents the number of correctly recognized chatter samples, and FN represents the number of chatter samples not recognized.

From the results in the Tab. 3, it can be observed that the recall of chatter signal feature judgment using the improve convolutional clustering method is generally superior to the standard deviation method. The improve convolutional clustering method is more reliable.

Based on the above research results, this study ultimately selects the improve convolutional clustering method to estab-

Table 3. Chatter sample recall rate under different feature extraction methods.

Feature extraction methods	Spindle speed n(r/min)	Recall
Standard deviation (CWT)	1000∼3000	90 %
Standard deviation (CWT)	3000∼5000	88.63 %
Improve convolutional clustering (CWT)	1000∼3000	92.66 %
Improve convolutional clustering (WPT)	3000∼5000	94.45 %

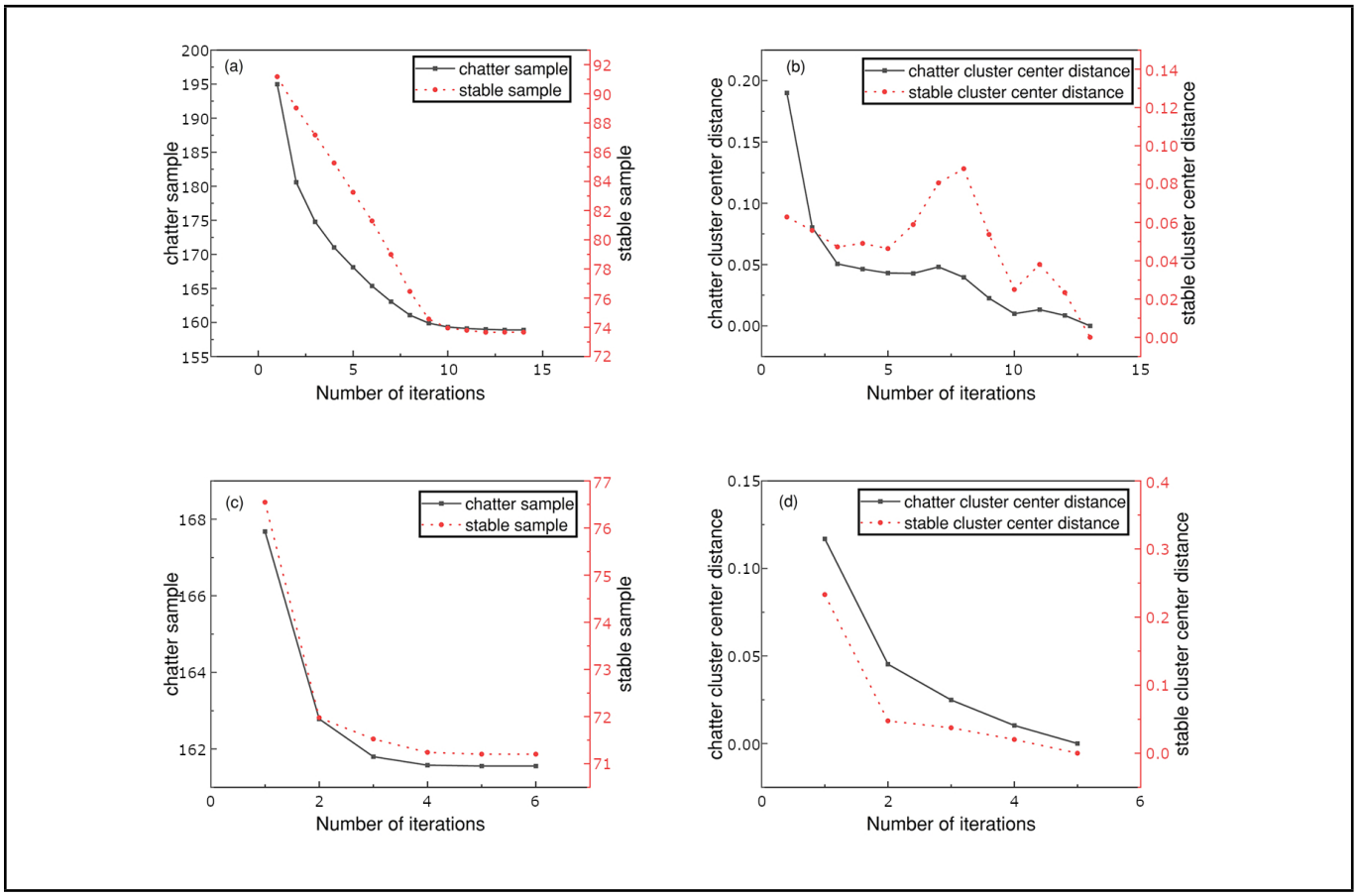

Figure 14. Iterative process diagram of the improved convolutional clustering algorithm.

Table 4. Chatter sample recall rate under different feature extraction methods.

Feature extraction methods	Spindle speed n(r/min)	Accuracy	Recall
Improve convolutional clustering (CWT)	1000∼3000	95 %	92.66 %
Improve convolutional clustering (WPT)	3000∼5000	96.3 %	94.45 %

lish an unsupervised learning model. The accuracy and recall of the model are shown in Tab. 4.

7. CONCLUSIONS

This paper proposes a improve convolutional clustering method for unsupervised learning of chatter signal features in milling processes. The recognition accuracy of chatter at different spindle speeds reached 95 % and 96.3 % respectively. The results show that, at speeds between 1000 r/min-3000 r/min, chatter features are more pronounced in the low-frequency region. The use of CWT can better distinguish between chatter signals and stable cutting signals. At speeds between 3000 r/min-5000 r/min, the increase in self-excited vibration energy introduces more interference in the low-frequency region. However, the difference between chatter signals and stable cutting signals is significant in the high-frequency range, and the use of WPT can better distinguish between them. In summary, this paper draws the following conclusions:

1. Establishing a hybrid preprocessing model for different operating conditions can improve the ability to identify chatter features.
2. Based on different spindle speed ranges, the optimal mapping space structure for chatter features was analyzed, resulting in an improved classification accuracy of the chatter recognition system.
3. Utilizing Wasserstein distance as a discriminant for sample distances in the clustering model effectively addresses the issue of decreased accuracy in high-dimensional spaces for the clustering model.

ACKNOWLEDGEMENT

This work supported by Xiamen University of Technology high level talent projects (4010522016), China.

REFERENCES

[1] Quintana G, Ciurana J, and Teixidor D. A new experimental methodology for identification of stability lobes diagram in milling operations, *International Journal of Machine Tools and Manufacture*, **48**(15), 637-1645, (2008). https://doi.org/10.1016/j.ijmachtools.2008.07.006

[2] Quintana G, et al. Sound mapping for identification of stability lobe diagrams in milling processes, *International Journal of Machine Tools and Manufacture*, **49**(3-4), 203-211, (2009). https://doi.org/10.1016/j.ijmachtools.2008.11.008

[3] Jiang GJ, Wu DW, Zhang N, et al. Chatter reliability prediction of side milling aero-engine blisk, *J Mech Sci Technol*,

34, 4005–4013, (2020). https://doi.org/10.1007/s12206-020-2211-z

[4] Mishra R and Singh B. Prediction of milling chatter using SBLMD-ANN, *Journal of Mechanical Science and Technology*, **36**(2), 877-882, (2022). https://doi.org/10.1007/s12206-022-0135-5

[5] Kim SK, and Lee SY. Chatter prediction of end milling in a vertical machining center, *Journal of Sound & Vibration*, **241**(4), 567-586, (2001). https://doi.org/10.1006/jsvi.2000.3144

[6] Gu D, et al. Three degrees of freedom chatter stability prediction in the milling process. *Journal of Mechanical Science and Technology*, **34**(9), 3489-3496, (2020). https://doi.org/10.1007/s12206-020-0801-4

[7] Sestito GS, Venter GS, Ribeiro KSB, et al. In-process chatter detection in micro-milling using acoustic emission via machine learning classifiers, *Int J Adv Manuf Technol*, **120**, 7293–7303, (2022). https://doi.org/10.1007/s00170-022-09209-w

[8] Li M, Jiang X, Huang D. Surface defect and chatter monitoring in robotic drilling CFRP composites using acoustic emission technique, *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, **238**(4), 605-615, (2024). https://doi.org/10.1177/09544054231166781

[9] Vashisht RK, and Peng Q. Online Chatter Detection for Milling Operations Using LSTM Neural Networks Assisted by Motor Current Signals of Ball Screw Drives, *Journal of Manufacturing Science and Engineering*, **143**(1), 1-29, (2020). https://doi.org/10.1115/1.4048001

[10] Peng D, et al. Prediction of milling force based on spindle current signal by neural networks, *Measurement*, **205**, 112153, (2022). https://doi.org/10.1016/j.measurement.2022.112153

[11] Li K, et al. A novel online chatter detection method in milling process based on multiscale entropy and gradient tree boosting, *Mechanical Systems and Signal Processing*, **135**, 106385, (2020). https://doi.org/10.1016/j.ymssp.2019.106385

[12] Lida Z, and Liu C. Recent progress of chatter prediction, detection and suppression in milling, *Mechanical Systems and Signal Processing*, **143**, 106840, (2020). https://doi.org/10.1016/j.ymssp.2020.106840

[13] Yuan Y, Jing X, Li H, et al. Chatter detection based on wavelet coherence functions in micro-end-milling processes, *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, **233**(9), 1934-1945, (2019). https://doi.org/10.1177/0954405418808214

[14] Yungchen Y, et al. Real-time chatter detection and automatic suppression for intelligent spindles based on wavelet packet energy entropy and local outlier factor algorithm, *International Journal of Advanced Manufacturing Technology*, **103**(1-4), 297-309, (2019). https://doi.org/10.1007/s00170-019-03551-2

[15] Yang F. Timely online chatter detection in end milling process, *Mechanical Systems and Signal Processing*, **75**, 668-688, (2016). https://doi.org/10.1016/j.ymssp.2016.01.003

[16] Wang R, Niu J, Sun Y. Chatter identification in thin-wall milling using an adaptive variational mode decomposition method combined with the decision tree model, *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, **236**(1-2), 51-63, (2022). https://doi.org/10.1177/0954405420933705

[17] Postel M, Bugdayci B and Wegener K. Ensemble transfer learning for refining stability predictions in milling using experimental stability states, *Int J Adv Manuf Technol*, **107**, 4123–4139, (2020). https://doi.org/10.1007/s00170-020-05322-w

[18] Tran MQ, Liu MK. and Tran QV. Milling chatter detection using scalogram and deep convolutional neural network, *Int J Adv Manuf Technol*, **107**, 1505–1516, (2020). https://doi.org/10.1007/s00170-019-04807-7

[19] Wang Y, Bo Q, Liu H, et al. Mirror milling chatter identification using Q-factor and SVM, *Int J Adv Manuf Technol*, **98**, 1163–1177, (2018). https://doi.org/10.1007/s00170-018-2318-x

[20] Liming W, et al. Two New Kurtosis-based Similarity Evaluation Indicators for Grinding Chatter Diagnosis under Non-stationary Working Conditions, *Measurement*, **176**(2), 109215, (2021). https://doi.org/10.1016/j.measurement.2021.109215

[21] Yun C, et al. Feature extraction using dominant frequency bands and time-frequency image analysis for chatter detection in milling, *Precision Engineering*, **56**, 235-245 (2018). https://doi.org/10.1016/j.precisioneng.2018.12.004

[22] Maojun L, Dingxiao H and Xujing Y. Chatter stability prediction and detection during high-speed robotic milling process based on acoustic emission technique, *Int J Adv Manuf Technol*, **117**, 1589–1599, (2021). https://doi.org/10.1007/s00170-021-07844-3

[23] Rong Mao L, et al. Acoustic Signal Analysis by Teager-Huang Transform for Milling Chatter Recognition, *Sensors and materials: An International Journal on Sensor Technology*, **3**(1), 32, (2020). https://doi.org/10.18494/SAM.2020.2671

[24] Dong, X., Tu, G., Wang, X. et al. Real-time chatter detection via iterative Vold-Kalman filter and energy entropy, *Int J Adv Manuf Technol*, **116**, 2003–2019, (2021). https://doi.org/10.1007/s00170-021-07509-1

[25] Yanqing Z, et al. MaxEnt feature-based reliability model method for real-time detection of early chatter in high-speed milling, *ISA Transactions*, **113**, 39-51, 2021. https://doi.org/10.1016/j.isatra.2020.07.022