IEICE TRANS. INF. & SYST., VOL.E106–D, NO.5 MAY 2023

PAPER *Special Section on Deep Learning Technologies: Architecture, Optimization, Techniques, and Applications*

Intelligent Tool Condition Monitoring Based on Multi-Scale Convolutional Recurrent Neural Network

Xincheng CAO[†], **Bin YAO**[†a)], **Binqiang CHEN**[†], **Wangpeng HE**[††], **Suqin GUO**[†††], ***and* Kun CHEN**[†††], ***Nonmembers***

SUMMARY Tool condition monitoring is one of the core tasks of intelligent manufacturing in digital workshop. This paper presents an intelligent recognize method of tool condition based on deep learning. First, the industrial microphone is used to collect the acoustic signal during machining; then, a central fractal decomposition algorithm is proposed to extract sensitive information; finally, the multi-scale convolutional recurrent neural network is used for deep feature extraction and pattern recognition. The multi-process milling experiments proved that the proposed method is superior to the existing methods, and the recognition accuracy reached 88%.
key words: *digital manufacturing, tool condition monitoring, acoustic, deep learning, convolutional recurrent neural network*

1. Introduction

As soon as the machining process starts, the tool removes material by applying pressure and shearing forces to the workpiece, and the tool itself gradually wears and degrades [1]. Tool wear is designed to be a steady, cumulative process [2]. Nevertheless, at the industrial site, due to the changes of workpiece materials, cooling and lubrication, process parameters and other working conditions, the uncertainty of tool wear is aggravated, and the expected life is often not achieved [3]. For complex, high-value parts, a degraded tool that is not replaced in time can damage the workpiece and even the machine tool [4]. As a result, machine tool technicians have to stop frequently to check the tools and adopt a conservative strategy to retire the tools early [5]. It will lead to increased tool costs and downtime. Statistics show that machine downtime caused by tool abnormalities accounts for 10% to 40% of the total downtime [6]. Research by Kennametal shows that a sensitive and reliable tool condition monitoring system can save 30% of the machining cost of CNC machine tools [7], [8]. Therefore, tool condition monitoring (TCM) is widely recognized as an important technology to improve efficiency and reduce consumption in the metal cutting industry [9].

Tool condition monitoring can be roughly divided into physical-based methods and data-driven methods [10]. Physically-based methods try to establish the degradation model of the tool, and then determine the remaining life based on the accumulated cutting time. However, statistics prove that in the roughing or semi-finishing stage, the actual degradation process of the tool is affected by a series of accidental factors such as the quality of the tool, the change of the workpiece material, and the fluctuation of the cutting allowance. It is difficult to establish a universal mathematical model. In recent years, artificial intelligence has achieved rapid development. Data-driven artificial intelligence model has achieved a series of successful cases in structural health monitoring, mechanical fault diagnosis and other fields [11]–[13], and many studies have tried to apply artificial intelligence to tool condition monitoring system [14], [15]. When the data samples are sufficient, the artificial intelligence model fits the complex mapping relationship between the signal samples and the tool state through end-to-end learning. It can not only model the degradation process of tools, but also learn the personalized characteristics of different monitoring objects.

Based on the above research, a data-driven tool condition monitoring method based on convolutional neural network and recurrent neural network is proposed. The main contributions of this paper are summarized as follows:

(1) An indirect tool condition monitoring method based on industrial microphone is proposed, which is more suitable for on-line monitoring than dynamometer and acceleration sensor, and more sensitive than spindle current signal.

(2) A central fractal decomposition algorithm is proposed to preprocess the acoustics signal. The implicit wavelet packets are constructed to reduce the information loss caused by frequency band decomposition.

(3) A parallel multi-scale convolutional neural network is constructed, and the performance of extracting multi-scale impact features from the envelope demodulation spectrum is enhanced via the atrous convolution.

2. Related Work

Tool condition monitoring has always been a research hotspot in the field of advanced manufacturing, and it is an important task of workshop informatization and intellectualization. In recent years, scholars have published relevant research results, the core content of which includes sensing technology, data preprocess and recognition model.

Machine vision is a widely studied tool condition mon-

Manuscript received November 22, 2021.
Manuscript revised February 18, 2022.
Manuscript publicized June 16, 2022.
[†]The authors are with the School of Aerospace Engineering, Xiamen University, Xiamen 361005, China.
[††]The author is with the School of Aerospace Science And Technology, Xidian University, Xi'an 710071, China.
[†††]The authors are with Fujian Great Power Science and Technology Co., Ltd, Fuzhou 350000, China.
a) E-mail: yaobin@xmu.edu.cn
DOI: 10.1587/transinf.2022DLP0043

itoring technology [16], but it requires stopping machining and is difficult to overcome the interference of coolant and chips. Vibration signal is a useful indirect monitoring signal, which does not need to interrupt the cutting process. Zhou developed a tool holder system with integrated acceleration sensors [17]. Hassan extracted features from spindle vibration signals to identify tool wear conditions [18]. However, the sensor must be mounted close to the spindle, inevitably interfering with machine tool motion. In contrast, the acquisition of spindle current is more convenient. Yuan extracted features from the spindle current signal and then input them into the ensemble learning classifier for analysis [19]. Shen included machining parameters such as depth of cut as independent variables in the monitoring model [20]. However, under normal working condition, the spindle operates at a low workload, and the current response to tool wear is very weak, making it difficult to perceive tool degradation. To address this, the acoustic signal is chosen as the sensing means in this paper because it is more sensitive to tool degradation. Industrial microphones can be mounted relatively far from the spindle, protected from coolant and chips.

The advanced data processing algorithms to denoise and decompose the original data is very important for feature extraction. Zhou utilized wavelet transform modulus maxima estimation [21]. Jimenezde's experiment prove that the evolution of wavelet decomposition details is sensitive to tool wear progress [22]. Li extracted the wear time-domain characteristics of tool vibration signals through wavelet packet transform [23]. Chen adapted wavelet threshold algorithm to de-noise the drilling force and vibration acceleration [24].

The recognition model of tool condition has experienced the evolution from machine learning to deep learning. Ajayram constructed a statistical monitoring models for indexable inserts based on decision trees and random trees [25]. Ostasevicius adapted support vector machines to identify the wear of endmills [26]. Oo employed random forest and multiple linear regression [16]. However, experiments show that the expert design of manual features is the key to the success of machine learning algorithms, which leads to the limitation of generalization. With the development of deep learning, many researchers apply deep learning technology to tool condition monitoring. Ou used stacked automatic encoder to adaptively extract deep features from monitoring signals [27]. Furthermore, Harshavardhan assigned the two tasks of feature extraction and state recognition modeling to a single deep CNNs [28]. To observe the tool wear over a longer time span, Xu utilized recurrent neural networks (RNNs). The life cycle of the object is divided into several stages, and a corresponding gated recurrent unit network is trained for each stage [29]. Cai applies the long-short-term memory network (LSTM) to the long-distance fusion of multi-sensor information [30]. Deep learning model can handle larger size input and adaptively extract deep features. It reduces the dependence on feature engineering and improves the robustness of the model. For higher accuracy and reliability, this paper proposes a tool condition recognition scheme based on the combination of CNNs and RNNs, the former is responsible for feature extraction from the short-term acoustic signal, and the latter is responsible for modeling the tool degradation features in a wider time window.

3. The Proposed Method

3.1 Overview of the Proposed Method

The flowchart of the proposed method is shown in Fig. 1. First, an industrial microphone is used to collect the acoustic signal during the milling process. Then use the central fractal decomposition algorithm to preprocess the data and extract the feature vector. Since there are significant differences in the degradation of different tools, the acoustic record of the current moment are not enough to accurately identify the condition of the tool. Therefore, segments are cut from acoustic records at a large interval to construct acoustic feature vector sequences. Then input the acoustic feature vector sequence into the multi-scale convolutional recurrent network to extract deep features. Finally, the nonlinear classifier trained with MsCRNN is used to identify the stability state of the milling process.

3.2 Data Preprocessing via Central Fractal Decomposition

The dual-tree complex wavelet packet decomposition (DCWPD) uses parallel dual-tree decomposition coefficients to achieve information complementarity, thereby obtaining approximate translation invariance and reducing the loss of information. It has achieved many successful applications in the field of fault diagnosis. However, the frequency response of the dyadic wavelet will attenuate at the frequency band boundary. The information of the transition band cannot be reconstructed ideally.

To address this, implicit wavelet packets are constructed based on DCWPD. The dyadic wavelet packets

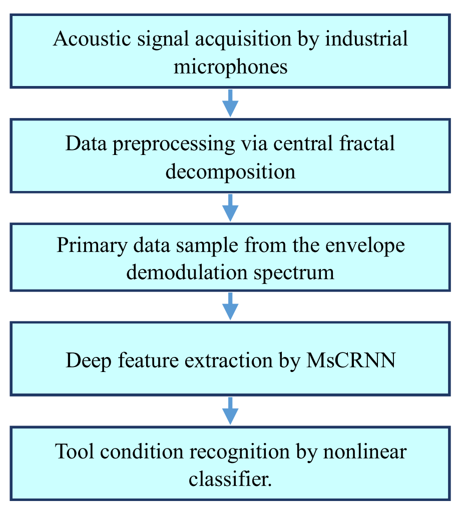


Fig. 1 Flow chart of the proposed method.

(wp_s) and the implicit wavelet packets (iwp_s) are combined to perform central fractal decomposition of the acoustic signal to improve the completeness of the reconstructed signal. Let the input acoustic signal be denoted as $x(n)$, the implicit wavelet packets are constructed as follows:

Step 1: Perform DCWPD on the input acoustic signal, such that $x(n)$ is decomposed into a set $D_k = D_k^j(n) \mid j = 1, 2, \ldots, 2^k$.

Step 2: Rearrange $D_k(n)$. Let the resulting set be $R_k(n)$. Let the binary index of D_k^j be

$$j = \sum_{m=0}^{k-1} 2^m n_m + 1, \tag{1}$$

Let the binary index of $R_k^{j'}$ be

$$j = \sum_{m=0}^{k-1} 2^m n'_m + 1, \tag{2}$$

The mapping between D_k^j and R_k^j is described as below:

$$n'_m = \begin{cases} n_m & m = k-1, \\ mod(n_m + n_{m+1}, 2) & m = 0, 1, \ldots, k-2, \end{cases} \tag{3}$$

Step 3: Generate the implicit wavelet packet using the following equation.

$$iwp_k^j(n) = R_k^{2j}(n) + R_k^{2j+1}(n), \ 1 \leq 2^{K-1} - 1, \tag{4}$$

The implicit wavelet packets form a series of central fractal wavelet packet groups, as shown in Fig. 2, which can continuously enlarge or refine the frequency domain view around any wavelet packet. In this paper, the wavelet basis function employed is Q-Shift 20 introduced by Chen in [31].

3.3 Construction of Primary Data Sample

The extraction process of the primary data sample at a single moment is shown in Fig. 3. Firstly, the original acoustic recording in 1 second is intercepted, and then the central fractal decomposition is performed. 15 sub-signals whose bandwidth is one-sixteenth of the sampling frequency are

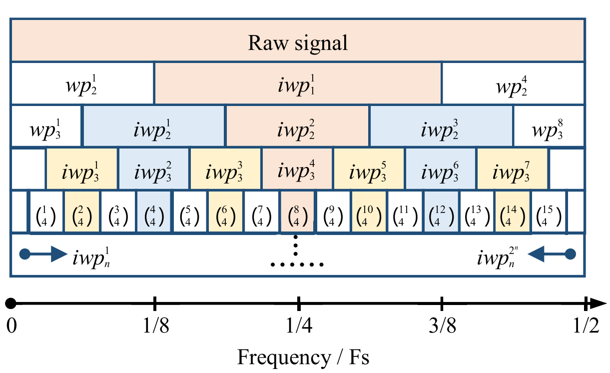


Fig. 2 Frequency-scale paving of implicit wavelet packets.

reconstructed. The Hilbert envelope demodulation spectrum near the characteristic frequency is used to form the short-term data sample. The characteristic frequency is set to the spindle rotation frequency, and the 2, 3, 6, 9 times of the spindle rotation frequency. Among them, the 6 times of the spindle rotation frequency is exactly the 2 times the cutting frequency. The 9 times frequency of the spindle rotation frequency is 3 times the cutting frequency.

Taking the current time as the reference and taking 12 seconds as the interval, extract the primary acoustic feature vectors of 5 historical moments to form the primary data sample.

3.4 Short-Term Feature Extraction Based on Parallel Multi-Scale Convolution

CNN has been proven to be an efficient tool for feature extraction. The classic CNN adopts a serial structure. In order to obtain a larger receptive field for the kernel, it is necessary to stack convolution layers and pooling layers in succession. This strategy achieves multi-scale feature extraction, but leads to large-scale computation. In this paper,

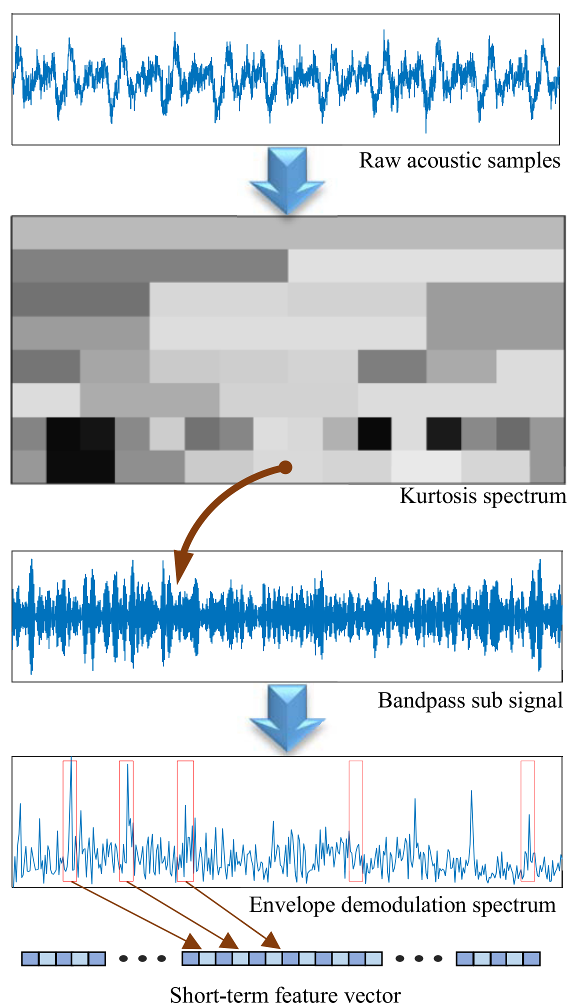


Fig. 3 Construction of short-term data sample.

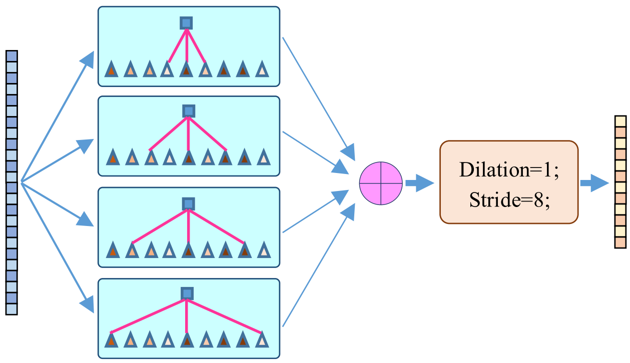


Fig. 4 The utilized parallel multi-scale convolution module.

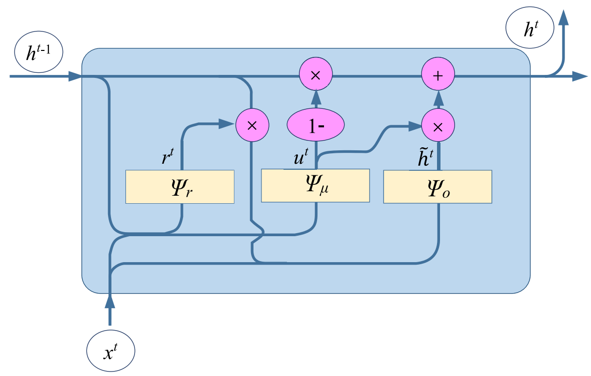


Fig. 5 The structure of multi-scale convolutional recurrent neural network.

Table 1 The parameters of the multi-scale convolutional subnetwork.

Layers	Channel number	Kernel size
1	8	32
2	16	16
3	32	8
4	8	5

the short-term feature extractor is used as part of the convolutional RNN, thereby, the computation is limited more strictly. Therefore, this paper uses diverse atrous convolution kernels to construct parallel multi-scale convolution module, as shown in Fig. 4. Atrous convolutions with different dilation rates are arranged side-by-side in the same layer to extract multi-scale features without increasing computation. Then, the feature vectors of different scales are connected into a convolutional layer with a stride of 8 for downsampling.

Atrous convolution, by adding holes in the convolution kernel, the receptive field of the convolution kernel can be expanded without the need for a pooling layer, while maintaining low calculations. Let $F: \mathbb{Z} \to \mathbb{R}$ be a discrete function. Let $\Omega_r = [-r, r] \bigcap \mathbb{Z}$ and let $k: \Omega_r \to \mathbb{R}$ be a discrete filter of size $2r + 1$. The discrete convolution operator $*$ can be defined as

$$(F * k)(p) = \sum_{s+t=p} F(s)k(t). \tag{5}$$

Let l be a dilation factor, then the atrous convolution $*_l$ can be defined as:

$$(F *_l k)(p) = \sum_{s+lt=p} F(s)k(t). \tag{6}$$

The multi-scale convolution module is used to form the convolution sub network, and the structural parameters are listed in Table 1. A total of four layers are used. As the data is transmitted forward, the channels are increased and the convolution kernel is reduced. Each channel is a multi-scale convolution module composed of four atrous convolution, and the dilation rate are fixed at 1, 2, 4 and 8. After the convolutional layers, the outputs of each channel are concatenated into a single feature vector. According to the data structure of the recurrent neural network, the output size of the short-term feature extractor should be consistent with the input. Therefore, two fully connected layers are set after the convolutional layer. The final output layer has as many neurons as the input, and the penultimate layer has twice as many neurons as the input layer.

3.5 Long-Term Modeling via Convolutional Recurrent Neural Network

The acoustic signal generated by the cutting process is affected by many accidental factors, and it is difficult to evaluate the reliability of the tool based on a short-term sample at a single moment. In order to solve this problem, MsCNN is embedded in a recurrent neural network to evaluate the stability of the processing process from a wider time window. The combination of feedforward connection and feedback connection empowers the recurrent neural network to process time series. The constructed multi-scale convolutional recurrent neural network (MsCRNN) is shown in Fig. 5. x^t is the feature vector input at the current moment, and h^{t-1} is the output of the previous feature vector processed by MsCRNN, which is also called the hidden cell state.

A total of 3 multi-scale convolutional neural networks (MsCNN) are used. From left to right, they are used to determine which historical information should be inherited into the processing of the feature vector at the current moment, which historical information should be updated, and extract features from the current input feature vector. The output of MsCRNN and the operation of each gate in it are defined by the following equations:

$$h^t = (1 - \mu^t) \otimes h^{t-1} + \mu^t \otimes \tilde{h}^t, \tag{7}$$

$$\tilde{h}^t = \tanh(\Psi_o([h^{t-1} \otimes r^t, x^t])), \tag{8}$$

$$\mu^t = \delta(\Psi_\mu([h^{t-1}, x^t])), \tag{9}$$

$$r^t = \delta(\Psi_r([h^{t-1}, x^t])), \tag{10}$$

Among them, $\otimes$ denotes pointwise multiplication, $[\cdot, \cdot]$ denotes vector connection, $\Psi(\cdot)$ denotes multi-scale convolutional feature mapping, $\delta(\cdot)$ denotes *sigmoid* activation function, $\tanh(\cdot)$ denotes *tanh* activation function. The *sigmoid* activation function maps the input to $[0, 1]$ to realize the function of the gate. The *tanh* activation function is used to adjust the range of alternative states to avoid gradient explosion.

4. Experiments Investigation

In order to verify the practicability of the proposed method, an end milling experiment was carried out.

4.1 Experiment Set-Up and Data Acquisition

To verify the effectiveness of the proposed method, a series of milling experiments were carried out. Experiments are conducted on DMTG VDL850A machining center, as shown in Fig. 6. Uncoated three flute endmills with a diameter of 10 mm are used, and the material is high speed steel. The workpiece is a rolled normalized JIS S45C steel block. It should be noted that this experiment simulates rough machining, so the oxide layer of the workpiece is retained. Industrial microphones are installed on the workbench and the sampling frequency is set to 12 KHz.

During the experiment, down milling and up milling is performed alternately. Five process schemes are implemented in turn, as listed in Table 2. Among them, a_e denotes radial cutting depth, a_p denotes axial cutting depth, a_f denotes feed rate and the spindle speed is set to 2500 r/min. For ease of description, these five milling cycles are defined as one test. After each test, use the tool microscope to collect the image of the end face of the endmill, check whether chipping occurs, and measure the wear width of the flank face.

4.2 Computer Hardware and Software

The software for preprocessing acoustic records and extracting primary feature vectors is based on the MATLAB platform and is implemented on a computer equipped with an Intel Core i7 central processing unit and 16 Gb memory.

The MsCRNN model is constructed by Python software. Keras application programming interface and Tensorflow framework are utilized. The experiments are conducted on an Intel Core i7 central processing unit with 16 Gb memory. MATLAB is used for image processing tasks. The cross entropy is utilized as the loss function, and the Adam optimizer is adapted. Cross entropy is an efficient objective function for combinatorial and continuous optimization and is now widely used to train neural networks for classification [32]. A series of advantages of the Adam optimization algorithm have been widely recognized, such as straightforward to implement, high computationally efficiency, low memory space requirements and invariant to diagonal rescaling of the gradients [33].

5. Results and Discussion

5.1 Instability Caused by Degradation of Endmill

In the experiment, 10 milling tests were performed for each endmill. In fact, most endmills have already chipped before the tenth test. To obtain complete experimental data, milling was continued for a period with the chipped endmills. Figure 7 illustrates the change of the effective value of the acoustic signal [1, 2000] frequency band. The purple line indicates the time of chipping, and the image on the right is the flank images of each endmill when it is chipped.

Figure 7 reveals the difference in the degradation of endmills. Endmill C chipped in the fifth test. Endmill E deteriorated in the form of wear and did not chip until the 9th test. Analyzing the milling parameters of the two, the axial cutting depth of the two is the same, and the feed speed of the endmill E is greater. This phenomenon prove that the actual capacity of cutting tool is affected by many accidental factors, such as the uniformity of the workpiece material

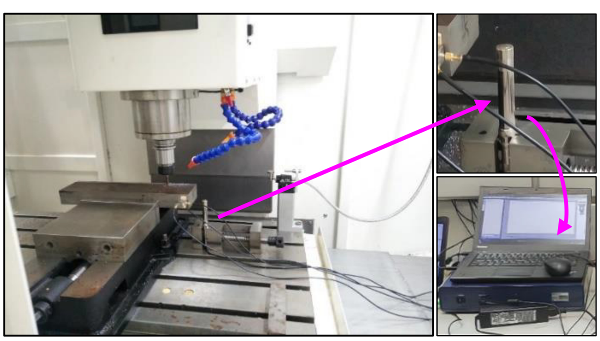


Fig. 6 The set-up of endmilling experiment.

Table 2 The process parameters.

Endmill	a_e (mm)	a_p (mm)	a_f (mm/rev)
A	10	0.5	0.053
B	10	0.5	0.067
C	10	0.6	0.053
D	10	0.6	0.060
E	10	0.6	0.067

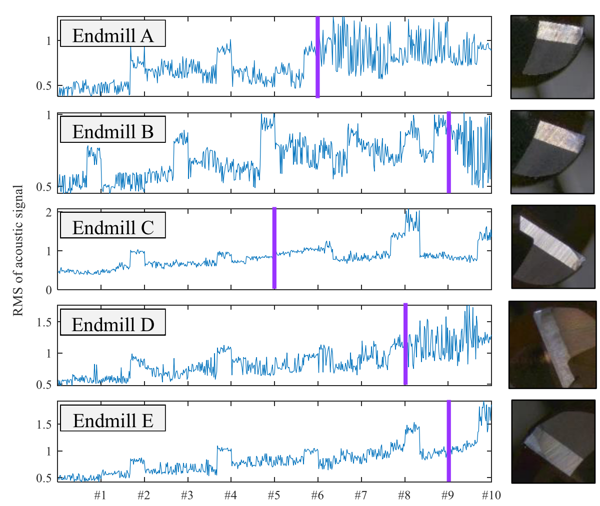


Fig. 7 Overview of the monitoring signals and photos of damaged endmill tips.

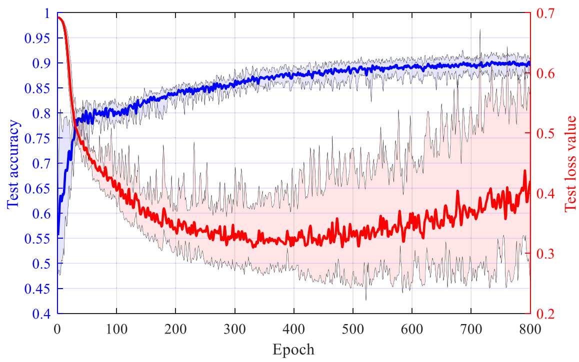


Fig. 8 The accuracy and loss value evolution curve when using the data of endmill D and E for testing.

and the accuracy of tool clamping.

For the identification experiments, label the data samples from milling test when chipping occurred as 'critical', label the samples of previous tests as 'reliable', and label the samples of the subsequent tests as 'chipped'.

5.2 The Learning Process of the Proposed MsCRNN

In order to verify the generalization and practicability of the proposed method, the samples of endmill A, B, and C are utilized as train set, and the samples of endmill D and E are used to test the model. The test accuracy and test loss of the five-fold crossover experiment are shown in Fig. 8. The test accuracy reached its highest in the 500th epochs. However, from the 400th epoch, the test loss gradually increases and fluctuates violently. This indicates that the model is overfitting. As analyzed earlier, there are significant differences in the individual property of endmills, and the size of the data sample is limited. This results in that the training dataset cannot fully cover all possibilities. As the training progresses, the model learns the detailed features of the training samples. Since there are difference between the test samples and the training samples, the model that over-matched to training samples is no longer suitable for the test samples. As a result, test losses gradually worsen and fluctuate violently.

Considering the accuracy and generalization ability comprehensively, the iterative optimization of MsCRNN is stopped at the 400th epoch. The test accuracy rate has reached 88%, which has certain reference value for tool management on the production site.

5.3 Comparison with CNN and NN-Based LSTM

In this section, the performance of CNN and NN-based GRU and the recommended MsCRNN are compared. Among them, CNN takes the primary feature vector at the current moment as input, and the input of GRU is the same as that of MsCRNN. The loss function and optimizer used in the training process of each model are also the same. Multiple tasks are defined, as listed in Table 3, the test set and training set are different in different tasks. Table 3 lists the average test accuracy achieved by different deep models when performing each task. First, analyze the performance differences when performing different tasks. All three models achieved the highest test accuracy in task $A/B/D/E \to C$, followed by task $A/B/C \to D/E$. The test accuracy achieved in tasks $A/B \to C/D/E$, $C/D/E \to A/B$, and $A/C \to B/E$ is lower, but the difference is not much. In the execution of task $B/E \to A/C$, the accuracy of various algorithms decreased significantly. Among the three models, CNN achieves the lowest test accuracy. An important reason is that CNN only performs diagnosis based on the acoustic features at the current moment. With the degradation of the tool, the acoustic characteristics show unpredictable fluctuations. It is necessary to synthesize the acoustic characteristics at multiple time points in a period of time to make an accurate diagnosis. The performance of NN-based GRU is between CNN and MsCRNN. In task $A/B/C \to D/E$ and task $A/B/D/E \to C$, the test accuracy of NN-based GRU and MsCRNN is not much different. However, in the last four tasks where the working conditions of the test set and the training set are much different, the test accuracy of the NN-based GRU is significantly lower than that of the MsCRNN.

Table 3 The results of different deep models.

Task	CNN	NN-based GRU	MsCRNN
$A/B/C \to D/E$	0.781	0.876	0.886
$A/B/D/E \to C$	0.790	0.889	0.895
$A/B \to C/D/E$	0.767	0.824	0.867
$C/D/E \to A/B$	0.779	0.829	0.871
$A/C \to B/E$	0.771	0.821	0.868
$B/E \to A/C$	0.733	0.804	0.843

Summarizing the results, the ability of MsCRNN to recognize the reliability of endmills is significantly better than that of CNN, and at the same time, it has stronger generalization ability than NN-based GRU.

5.4 Comparison with Other Methods

In this section, the proposed method is compared with the methods disclosed in recent literatures. One machine-learning-based algorithm and three deep-learning-based algorithms participated in the comparative experiment.

Zhou introduced a machine-learning-based method in [21]. This method first uses wavelet transform modulus maxima (WTMM) estimation to de-noise the raw signal, then extracts features such as means and standard deviations, and finally uses support vector machine (SVM) to identify tool wear status.

Duan proposed an WPD+CNN method of tool wear monitoring in [34], which performs wavelet package decomposition (WPD) of the original monitoring signal to construct a multi-band feature map, then employed CNNs to extract deep features.

In [35], Zeng proposed a method for converting temporal monitoring signals into two-dimensional images based on triangular matrix of angle summation (TMAS). Then extract the deep features and identify the tool condition via

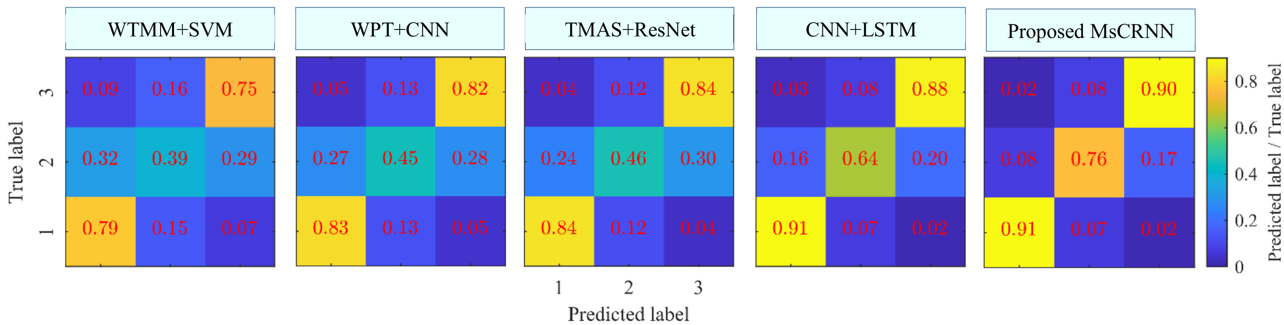


Fig. 9 Confusion matrix of test results achieved by various methods.

ResNet.

Zhang proposed a CNN+LSTM tool wear prediction method in [36]. This method uses a variety of sensors to collect monitoring signals, and then uses CNN to extract deep feature from the raw signal records of various sensors. The deep features of various monitoring signals are connected into a fusion feature vector, which is input into LSTM for further analysis. This method is also the combination of CNN and RNN, but it is significantly different from the method proposed in this paper. In the proposed method, CNN is embedded in RNN for learning. The input of MsCRNN is the sequence composed of instantaneous data, while in reference 38, the input of LSTM is the monitoring data at a single time.

Take the data samples of endmill B, C, and E as the training set, and take the data samples of endmill A and D as the test set. Implement five-fold crossover experiment to verify the applicability of each algorithm to the research case in this paper. Figure 9 illustrate the confusion matrix, where, label '1' denotes 'reliable', label '2' denotes 'critical', and label '3' denotes 'chipped'. Considering the significant difference in the number of samples under the three wear states, the confusion matrix display the proportion of each part of the sample to the total number of each category. For example, in the leftmost confusion matrix, the value in the cell with the true label '1' and the predicted label '2' is 0.15. The meaning is that, fifteen percent of the sample of reliable states were incorrectly predicted to be critical. The diagonal lines from the lower left corner to the upper right corner illustrate the recall of various category.

Among the five methods, the accuracy of the WTMM+SVM algorithm using manual features and machine learning is the lowest, and is lower than the accuracy reported in the literature. The possible reason is that the acoustic signal in this paper has severe fluctuations, and statistical features such as mean and kurtosis are contaminated by unknown noise. Using a deep learning model to adaptively extract deep features, the recognition accuracy of WT+CNN and TMAS+ResNet is more than 80% for both reliable and chipped samples, which has reference value. The algorithm proposed by Zhang uses CNN to extract deep features, and then uses LSTM for further analysis. The combination of the two deep models improves the recognition accuracy of the algorithm. However, this method is still based on the short-term signal samples at a single time, so the identification accuracy of the samples in the critical state is still not ideal.

The proposed method clipped short-term records from the raw monitoring signal at a longer time interval to form long-short-term data samples. The convolutional RNN is used to model the tool degradation process in a wider time window. The tool state is observed from a wider perspective, and the recognition recall of the critical state sample is improved by 12%.

Nevertheless, the recall of the proposed method is still less than 80% for identifying critical-state samples. According to the author's analysis, there are two reasons. On the one hand, the critical state is a transitional stage between stable wear and sharp chipping, which is similar to both states. On the other hand, the critical state has the fewest samples, and the data imbalance leads the model to pay more attention to the reliable and chipped types. It is accepted that only 8% of the critical-state samples were misidentified as reliable, and more were misidentified as chipped. Identifying a critical endmill as a tipping only slightly reduces tool utilization. However, a chipped endmill will damage the workpiece, and if it is not replaced in time, it will lead to more serious losses.

6. Conclusion

A deep-learning-based tool condition monitoring for digital manufacturing is proposed in this paper. The acoustic signal is firstly decomposed by the center fractal algorithm and then projected to the envelope demodulation space, which effectively removes the complex noise. Compared with ANN, MsCNN have fewer parameters and learn more features. MsCRNN integrates MsCNN and RNN and extracts deep features with higher adaptability and robustness. Fusing the current acoustic features and historical acoustic features, the model's ability to resist interference from unknown factors is significantly improved. The milling experiment reveals the uncertainty of endmill degradation and also verifies the advantage of the proposed method. Based on a single acoustic signal, the proposed method achieves a recognition accuracy of 88%, and the recall rate of chipped endmills reaches 90%.

The limitation of the proposed method is that the recognition accuracy of critical state samples is still not ideal, mainly because of the lack of ability to deal with the imbalance of data sets. In the next phase, research on data enhancement algorithms will be carried out, such as the use of generative adversarial models to generate critical-state samples. And research will be conducted on optimizing the adaptability of this method to adapt to other processing methods, such as forming milling, gear shaping, etc.

Acknowledgments

This work is supported by the National Key Research and Development Program of China (No.2020YFB1713503); the Fundamental Research Funds for the Central Universities (No.20720190009) and 2019 Industry-University-Research Cooperation Project of Aero Engine Corporation of China (No. HFZL2019CXY02).

References

[1] M. Kuntoğlu, A. Aslan, D.Y. Pimenov, Ü.A. Usca, E. Salur, M.K. Gupta, T. Mikolajczyk, K. Giasin, W. Kapłonek, and S. Sharma, "A review of indirect tool condition monitoring systems and decision-making methods in turning: Critical analysis and trends," Sensors, vol.21, no.1, 108, Jan. 2021.

[2] M. Kuntoğlu and H. Sağlam, "Investigation of progressive tool wear for determining of optimized machining parameters in turning," Measurement, vol.140, pp.427–436, 2019.

[3] Y. Zhou and W. Xue, "Review of tool condition monitoring methods in milling processes," Int. J. Adv. Manuf. Technol., vol.96, no.5-8, pp.2509–2523, 2018.

[4] Y. Li, C. Liu, J. Hua, J. Gao, and P. Maropoulos, "A novel method for accurately monitoring and predicting tool wear under varying cutting conditions based on meta-learning," CIRP Annals, vol.68, no.1, pp.487–490, 2019.

[5] D. Zhu, X. Zhang, and H. Ding, "Tool wear characteristics in machining of nickel-based superalloys," Int. J. Mach. Tools Manuf., vol.64, pp.60–77, 2013.

[6] C. Liu, Y. Li, J. Hua, N. Lu, and W. Mou, "Real-time cutting tool state recognition approach based on machining features in NC machining process of complex structural parts," Int. J. Adv. Manuf. Technol., vol.97, no.1-4, pp.229–241, 2018.

[7] K. Salonitis and A. Kolios, "Reliability assessment of cutting tool life based on surrogate approximation methods," Int. J. Adv. Manuf. Technol., vol.71, no.5-8, pp.1197–1208, 2014.

[8] A.G. Rehorn, J. Jiang, P.E. Orban, and E.V. Bordatchev, "State-of-the-art methods and results in tool condition monitoring: A review," Int. J. Adv. Manuf. Technol., vol.26, no.7-8, pp.693–710, 2005.

[9] S. Swain, I. Panigrahi, A.K. Sahoo, and A. Panda, "Adaptive tool condition monitoring system: A brief review," Proc. International Conference on Recent Advances in Materials, Manufacturing and Energy Systems (ICRAMMES), Vijayawada, India, pp.474–478, Jan. 2019.

[10] D. An, N.H. Kim, and J.-H. Choi, "Practical options for selecting data-driven or physics-based prognostics algorithms with reviews," Reliab. Eng. Syst. Saf., vol.133, pp.223–236, 2015.

[11] X. Cao, B. Chen, and N. Zeng, "A deep domain adaption model with multi-task networks for planetary gearbox fault diagnosis," Neurocomputing, vol.409, pp.173–190, Oct. 2020.

[12] J. Jiao, M. Zhao, J. Lin, and C. Ding, "Deep coupled dense convolutional network with complementary data for intelligent fault diagnosis," IEEE Trans. Ind. Electron., vol.66, no.12, pp.9858–9867, 2019.

[13] L. Wen, X. Li, L. Gao, and Y. Zhang, "A new convolutional neural network-based data-driven fault diagnosis method," IEEE Trans. Ind. Electron., vol.65, no.7, pp.5990–5998, 2018.

[14] X.-C. Cao, B.-Q. Chen, B. Yao, and W.-P. He, "Combining translation-invariant wavelet frames and convolutional neural network for intelligent tool wear state identification," Comput. Ind., vol.106, pp.71–84, 2019.

[15] F. Aghazadeh, A. Tahan, and M. Thomas, "Tool condition monitoring using spectral subtraction and convolutional neural networks in milling process," Int. J. Adv. Manuf. Technol., vol.98, no.9-12, pp.3217–3227, 2018.

[16] H. Oo, W. Wang, and Z. Liu, "Tool wear monitoring system in belt grinding based on image-processing techniques," Int. J. Adv. Manuf. Technol., vol.111, no.7-8, pp.2215–2229, Dec. 2020.

[17] C. Zhou, K. Guo, and J. Sun, "An integrated wireless vibration sensing tool holder for milling tool condition monitoring with singularity analysis," Measurement, vol.174, 109038, April 2021.

[18] M. Hassan, A. Sadek, and M.H. Attia, "Novel sensor-based tool wear monitoring approach for seamless implementation in high speed milling applications," CIRP Annals, vol.70, no.1, pp.87–90, 2021.

[19] J. Yuan, L. Liu, Z. Yang, and Y. Zhang, "Tool wear condition monitoring by combining variational mode decomposition and ensemble learning," Sensors, vol.20, no.21, 6113, Nov. 2020.

[20] Y. Shen, F. Yang, M.S. Habibullah, J. Ahmed, A.K. Das, Y. Zhou, and C.L. Ho, "Predicting tool wear size across multi-cutting conditions using advanced machine learning techniques," J. Intell. Manuf., vol.32, no.6, pp.1753–1766, Aug. 2021.

[21] A. Siddhpura and R. Paurobally, "A review of flank wear prediction methods for tool condition monitoring in a turning process," Int. J. Adv. Manuf. Technol., vol.65, no.1-4, pp.371–393, 2013.

[22] C. Zhou, K. Guo, and J. Sun, "Sound singularity analysis for milling tool condition monitoring towards sustainable manufacturing," Mech. Syst. Signal Proc., vol.157, 107738, 2021.

[23] A. Jiménez, M. Arizmendi, and J.M. Sánchez, "Extraction of tool wear indicators in peck-drilling of inconel 718," Int. J. Adv. Manuf. Technol., vol.114, no.9-10, pp.2711–2720, 2021.

[24] X.-R. Li, J.-M. Zhu, F.-Q. Tian, and H.-F. Pan, "Discrimination and prediction of tool wear state based on gray theory," J. Test. Eval., vol.48, no.6, pp.4262–4282, Nov. 2020.

[25] N. Chen, B. Hao, Y. Guo, L. Li, and N. He, "Research on tool wear monitoring in drilling process based on APSO-LS-SVM approach," Int. J. Adv. Manuf. Technol., vol.108, no.1-4, pp.2091–2101, 2020.

[26] K.A. Ajayram, R. Jegadeeshwaran, G. Sakthivel, R. Sivakumar, and A.D. Patange, "Condition monitoring of carbide and non-carbide coated tool insert using decision tree and random tree — A statistical learning," Proc. 28th International Conference on Processing and Fabrication of Advanced Materials (PFAM), Chennai, INDIA, pp.1201–1209, Dec. 2021.

[27] J. Ou, H. Li, G. Huang, and G. Yang, "Intelligent analysis of tool wear state using stacked denoising autoencoder with online sequential-extreme learning machine," Measurement, vol.167, 108153, July 2020.

[28] H. Mamledesai, M.A. Soriano, and R. Ahmad, "A qualitative tool condition monitoring framework using convolution neural network and transfer learning," Appl. Sci., vol.10, no.20, 7298, Oct. 2020.

[29] H. Xu, G.S. Hong, J.H. Zhou, J. Hong, and K.S. Woon, "Coarse-to-fine tool condition monitoring using multiple gated recurrent units," Proc. 45th Annual Conference of the IEEE Industrial Electronics Society (IECON), Lisbon, Portugal, pp.3737–3742, Oct. 2019.

[30] W. Cai, W. Zhang, X. Hu, and Y. Liu, "A hybrid information model based on long short-term memory network for tool condition monitoring," J. Intell. Manuf., vol.31, no.9-12, pp.1497–1510, 2020.

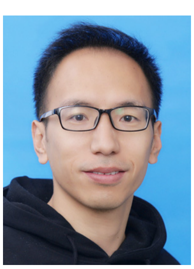
[31] B. Chen, Z. Zhang, Y. Zi, and Z. He, "Novel ensemble analytic discrete framelet expansion for machinery fault diagnosis," J. Mech. Eng., vol.50, no.17, pp.77–86, 2014.

[32] R. Rubinstein, "The cross-entropy method for combinatorial and continuous optimization," Methodol. Comput. Appl. Probab., vol.1, no.2, pp.127–190, 1999.
[33] D.P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.
[34] J. Duan, J. Duan, H. Zhou, X. Zhan, T. Li, and T. Shi, "Multi-frequency-band deep CNN model for tool wear prediction," Meas. Sci. Technol., vol.32, no.6, 065009, June 2021.
[35] Y.F. Zeng, R.L. Liu, and X.F. Liu, "A novel approach to tool condition monitoring based on multi-sensor data fusion imaging and an attention mechanism," Meas. Sci. Technol., vol.32, no.5, 055601, May 2021.
[36] X. Zhang, X. Lu, W. Li, and S. Wang, "Prediction of the remaining useful life of cutting tool using the Hurst exponent and CNN-LSTM," Int. J. Adv. Manuf. Technol., vol.112, pp.2277–2299, 2021.

Xincheng Cao received the master degrees in mechanical engineering from the School of Aerospace Engineering, Xiamen University. He is currently pursuing the Ph.D. degree in the School of Aerospace Engineering, Xiamen University. His main research interests include intelligent data analysis, deep learning, intelligent equipment and smart manufacturing.

Bin Yao received the Ph.D. degree from Xi'an Jiaotong University, in 2003. He is currently a Full Professor with the School of Aerospace Engineering, Xiamen University, China. His current research interests include shape of curved surface and NC technology, detection technology, and intelligent machining equipment.

Binqiang Chen received the B.S. degree in mechanical engineering from Sichuan University, in 2008, and the Ph.D. degree in mechanical engineering from Xi'an Jiaotong University, in 2013. His current research interests include smart manufacturing, structural health monitoring of equipment, and applied harmonic analysis.



Wangpeng He received the B.S. and Ph.D. degrees in mechanical engineering from Xi'an Jiaotong University, Xi'an, China, in 2007 and 2016, respectively. He then joined the School of Aerospace Science and Technology, Xidian University, where he is currently an Assistant Professor. His research interests include wavelet-based signal processing, sparsity techniques, and machinery fault diagnosis.

Suqin Guo received the B.S. degree from the College of Computer Application Software, Jilin University, Changchun, China, in 2014. Her research interest includes Information System, comprehensive energy services and artificial intelligence.

Kun Chen received the B.S. degree from the College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou, China, in 2009. His research interest includes Information System, comprehensive energy services and artificial intelligence.