

Research Article

Tool Wear Monitoring Algorithm Based on SWT-DCNN and SST-DCNN

Haining Gao ^{1,2}, Aimin Xie,³ Hongdan Shen,¹ Lei Yu,^{1,4} Yinling Wang,^{1,2} Yingjie Hu,¹ Yixiao Gao,¹ Jing Xu,¹ and Wanqing Wu¹

¹College of Energy Engineering, HuangHuai University, Zhumadian 463000, China

²Zhumadian International Joint Laboratory of Internet of Things for New Energy, Zhumadian 463000, China

³Zhengzhou Haozhong Electronic Technology Co., Ltd, Zhengzhou 450009, China

⁴Department of Computer Science, COMSATS University Islamabad, Abbottabad 22060, Pakistan

Correspondence should be addressed to Haining Gao; haininggao@huanghuai.edu.cn

Received 17 December 2021; Revised 2 January 2022; Accepted 4 January 2022; Published 18 January 2022

Academic Editor: Sheng Bin

Copyright © 2022 Haining Gao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The accurate monitoring of tool condition is of great significance to improve the machining quality and efficiency of parts, prolong the service life of tools and machine tools, and reduce the harm of manufacturing environment. In this dissertation, two methods based on synchronous compressed continuous wavelet transform and deep convolution neural network (SWT-DCNN) and synchronous compressed continuous wavelet transform and deep convolution neural network (SST-DCNN) are proposed to monitor tool wear. It is found that the recognition accuracy of SWT-DCNN method is 99.96%, and that of SST-DCNN method is 99.86%. The reason is that SWT method has good time-frequency energy aggregation. Compared with the SST-DCNN method, the recognition accuracy of the SWT-DCNN method is more stable. At the same time, it is found that the recognition rate of the SST-DCNN method in the process of normal tool monitoring is only 93.3%, which is easy to classify normal tools into the initial wear category. The experimental results show that the two methods proposed in this paper can effectively monitor the tool wear state.

1. Introduction

With the development of intelligent manufacturing technology, the process system is required to further improve the ability of active perception and independent decision-making. Therefore, tool condition monitoring has become a research hotspot in the machining field. The accurate monitoring of tool condition is of great significance to improve the machining quality and efficiency of parts, prolong the service life of tools and machine tools, and reduce the harm of manufacturing environment.

There are mainly two commonly used tool wear monitoring methods: one is the direct monitoring method based on machine vision [1] and the other is the indirect monitoring method based on sensor signals [2–6]. The visual monitoring method can obtain the tool wear more intuitively and quantitatively, but its premise is to stop cutting to

take a clear tool image, which leads to the reduction of machining efficiency [7]. The application of real-time wear monitoring methods based on sensor signals is more common. The tool wear state monitoring method based on sensor signals mainly collects the cutting force [2], vibration [3], acoustic emission [4], spindle current [5], and cutting temperature [6] in the machining process. Then, the machine learning method is used to determine the relationship between the signal and the wear amount through the analysis and processing of the signal. Finally, the wear state of the tool is judged according to the signal in the production process.

Many machine learning models are usually used in data-driven applications, such as extreme learning machine [3, 4], support vector machine [8, 9], and hidden semi-Markov model [10, 11]. Lei et al. [8] proposed a least squares support vector machine method for predicting tool life. Zhang and Zhang [9] established a tool wear monitoring model

combining trajectory similarity and support vector regression. Drouillet et al. [5] obtained the prediction model of tool remaining service life based on the artificial neural network by using spindle motor power. Lei et al. [3] used the hybrid GAPSO algorithm to optimize the initial weight and threshold of the extreme learning machine method. Then, they used the optimized extreme learning machine method to monitor the tool wear state. Zhou et al. [4] used the double-layer network structure to enhance the learning of time-frequency domain characteristics of acoustic signals. Then, the two angle kernel function is used to replace the preset hyperparametric problem of kernel function in the traditional extreme learning machine method, and finally, the tool wear monitoring in the milling process is realized.

A single sensor cannot capture all the characteristics of the tool wear [12]. To solve the above problems, researchers study multisensor fusion technology to detect chatter and tool wear. Aliustaoglu et al. [13] realized tool wear detection based on the statistical characteristics of cutting force, vibration and acoustic signal combined with a two-level fuzzy logic algorithm. Rizal et al. [14] realized tool wear detection based on the time-frequency domain characteristics of cutting force, vibration, torque, and temperature signals combined with Martin's system. Liu et al. [15] realized the detection of different machining states based on the time-frequency domain characteristics of force and vibration signals and support vector machine. Multisensor fusion technology can more accurately monitor the processing state, but there are great problems in industrial production.

With the development of computer technology, deep learning method shows outstanding performance in classification, especially in image classification. The advantage of deep learning is that it can automatically extract and learn representative features instead of performing manual feature extraction and selection. Li et al. [16] proposed a deep bidirectional long-term and short-term memory neural network model to predict the remaining service life of tools. Ma et al. [2] established a tool wear prediction model using convolution bidirectional long-term and short-term memory network and convolution bidirectional gated recursive unit. The minimum prediction errors are within 8%, which proves the effectiveness of the proposed method. An et al. [17] established a hybrid model combining the convolutional neural network and superimposed bidirectional and unidirectional LSTM networks to predict the remaining service life of tools, with an average prediction accuracy of 90%. Huang et al. [18] performed short-time Fourier transform on the vibration data in the cutting process to obtain the time-frequency domain diagrams under different tool states and combined with the convolutional neural network to monitor the tool state. Meanwhile, the convolutional neural network also shows excellent performance in pattern recognition when is applied to fault diagnosis [19] and cutting chatter [20, 21].

The tool condition monitoring signal obtained in the cutting process contains obvious unsteady characteristics, so time-frequency analysis has obvious advantages in processing machining signals. Firstly, the vibration signals

in the machining process are processed by synchronous compressed short-time Fourier transform (SST) and synchronous compressed continuous wavelet transform (SWT), and then the time-frequency domain diagrams with different tool states are obtained. Then, a deep convolution neural network (DCNN) model with an activation function of Leaky ReLU is constructed, which overcomes the problem of vanishing gradient. Finally, the deep neural network is trained with the time-frequency domain Atlas with different tool wear conditions, and two methods of automatically monitoring tool wear are obtained.

2. Theory

2.1. SST. Synchronous compression based on short-time Fourier transform is a time-frequency analysis method with high time-frequency resolution. In essence, the combination of time-frequency distribution rearrangement and short-time Fourier transform compresses the time-frequency distribution after short-time Fourier transform, so as to greatly improve the time-frequency resolution [22, 23].

The synchronous compression transformation of signal $x(t)$ is described as follows.

$$G(t, \omega) = \int_{-\infty}^{+\infty} x(u) \cdot g(u-t) \cdot e^{-i\omega u} du, \quad (1)$$

where $g(u-t)$ represents the time window function.

Based on Plancherel's theorem, equation (1) can be abbreviated as follows:

$$\begin{aligned} G(t, \omega) &= \int_{-\infty}^{+\infty} x(u) \cdot (g(u-t) \cdot e^{-i\omega u})^* du \\ &= \int_{-\infty}^{+\infty} x(u) \cdot (g_w(u))^* du \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{x}(\xi) \cdot (\hat{g}_w(\xi))^* d\xi, \end{aligned} \quad (2)$$

where $*$ represents conjugate operation, $\hat{g}_w(\xi)$ represents the Fourier transform of $g_w(u)$, and $\hat{x}(\xi)$ represents the Fourier transform of $x(u)$.

Set up $t' = u - t$, $\hat{g}_w(\xi)$ is described as follows:

$$\begin{aligned} \hat{g}_w(\xi) &= \int_{-\infty}^{+\infty} g(t') \cdot e^{i\omega(t+t')} \cdot e^{-i\xi(t+t')} dt' \\ &= e^{i\omega t - i\xi t} \cdot \int_{-\infty}^{+\infty} g(t') \cdot e^{i\omega t' - i\xi t'} dt' \\ &= e^{i\omega t - i\xi t} \cdot \hat{g}(\omega - \xi). \end{aligned} \quad (3)$$

Replace is described in equation (2) with equation (3) and introduce the phase shift operator $e^{i\omega t}$.

$$G(t, \omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{x}(\xi) \cdot (\hat{g}_w(\omega - \xi) \cdot e^{i\xi t}) du. \quad (4)$$

The Fourier transform of the signal $x(t)$ is described as follows:

$$\hat{x}(\xi) = 2\pi A \cdot \delta(\xi - 2\pi f_0). \quad (5)$$

Substitute equation (5) into equation (4), $G(t, \omega)$ is described as follows:

$$G(t, \omega) = A \cdot \hat{g}_\omega(\omega - 2\pi f_0) \cdot e^{i2\pi f_0 t}. \quad (6)$$

According to equation (6), the time-frequency energy of the signal will propagate within the frequency interval of $[2\pi f_0 - \Delta, 2\pi f_0 + \Delta]$, where Δ is the frequency range of the sliding time window function.

$$\begin{aligned} \partial_t G_e(t, \omega) &= \partial_t (A \hat{g}_\omega(\omega - 2\pi f_0) \cdot e^{i2\pi f_0 t}) \\ &= A \cdot (\omega - 2\pi f_0) \cdot e^{i2\pi f_0 t} \cdot j \cdot 2\pi f_0 \\ &= G_e(t, \omega) \cdot j \cdot 2\pi f_0, \end{aligned} \quad (7)$$

where $G_e(t, \omega) \neq 0$; then, the instantaneous frequency of f_0 is described as follows:

$$f_0(t, \omega) = -i \frac{\partial_t G_e(t, \omega)}{2\pi G_e(t, \omega)}, \quad (8)$$

$$SST(t, \omega) = \int_{-\infty}^{+\infty} G_e(t, \omega) \cdot \delta(\omega - f_0(t, \omega)) d\omega,$$

where $\int_{-\infty}^{+\infty} G_e(t, \omega) \cdot \delta(\omega - f_0(t, \omega)) d\omega$ is a synchronous compression operator and δ represents Dirac function.

2.2. SWT. Daubechies et al. [24] combined synchronous compression technology with continuous wavelet transform and proposed the SWT method. SWT increases the time-frequency resolution by compressing and rearranging the CWT transform coefficients in the frequency direction.

For any signal $s(t)$, its CWT is described in equation (10).

$$\begin{aligned} W_s(a, b) &= \langle s, \psi_{a,b} \rangle \\ &= \int s(t) a^{-(1/2)} \psi \left(\frac{t-b}{a} \right) dt, \end{aligned} \quad (9)$$

where $W_s(a, b)$ is the wavelet transform coefficient spectrum, a and b are scaling and displacement variables respectively, $*$ represents complex conjugate, and $\psi(t)$ is wavelet basis function.

According to Plancherel's theorem, transform $W_s(a, b)$ to the frequency domain; its expression is as follows.

$$W_s(a, b) = \frac{1}{2\pi} \int \hat{s}(\xi) a^{1/2} \hat{\psi} * (a\xi) e^{ib\xi} d\xi, \quad (10)$$

where ξ is the angular frequency and $\hat{s}(\xi)$ is the Fourier transform of signal s . When $W_s(a, b) \neq 0$, the instantaneous frequency $\omega_s(a, b)$ of a signal is described as follows:

$$\omega_s(a, b) = -i (W_s(a, b))^{-1} \frac{\partial}{\partial b} W_s(a, b). \quad (11)$$

Then, the synchronous compression transformation is carried out, the mapping from the starting point (b, a) to $(b, \omega_s(a, b))$ is established, and $W_s(a, b)$ is transformed from the time scale plane to the time-frequency scale plane to obtain a new time-frequency spectrum.

In order to suppress the ambiguity at the scale parameters, add the energy in the continuous interval $[\omega_l - (1/2)\Delta\omega, \omega_l + (1/2)\Delta\omega]$, and redistribute the distributed energy to the instantaneous frequency; the SWT can be obtained as follows.

$$SWT(\omega_l, b) = \frac{1}{\Delta\omega} \sum W_s(a_k, b) a_k^{-3/2} (\Delta a)_k, \quad (12)$$

where a_k is the k th discrete scale parameter, which satisfies $|\omega(a_k, b) - \omega_l| \leq \Delta\omega/2$, ω_l is the L -th center frequency, $\Delta\omega = \omega_l - \omega_{l-1}$, and $(\Delta a)_k = a_k - a_{k-1}$.

2.3. Convolutional Neural Network. Convolutional neural networks (CNN) are a kind of feedforward neural networks with depth structure including convolution calculation. It is a deep learning network as one of the representative algorithms of deep learning. It can extract high-level features from the input information and classify the input information according to its hierarchical structure with translation invariance. It has unique characteristics in image recognition and classification.

Convolutional neural networks mainly include an input layer, an output layer, and multiple hidden layers in structure. The hidden layer is usually composed of a convolutional layer, a pooling layer, and a fully connected layer. The output layer is usually composed of a fully connected layer and a classification layer.

In order to extract different features of the input feature map, the pixel values of the input features are convoluted by convolution check, and the convolution layer is calculated as follows [25].

$$x_{i,j}^{l+1} = f \left(\sum_{j=1}^L \sum_{i=1}^m (X_{i,j}^l \omega_{i,j}^l) + b \right), \quad (13)$$

where $X_{i,j}^l$ is the j th eigenvalue of the i th characteristic graph in the network layer l . L is the convolution kernel size. $\omega_{i,j}^l$ is the weight coefficient, b is the deviation value, and $f(\cdot)$ is the activation function.

The main function of the activation function is to present the nonlinear modeling ability of the network. Sigmoid, tanh, and ReLU activation functions are widely used. Sigmoid is the most widely used activation function, which has the shape of an exponential function. Its disadvantage is that it has very obvious saturation, and the derivatives on both sides of the function tend to 0. Tanh converges faster than sigmoid, reducing the number of iterations and causing the gradient to disappear. ReLU can maintain the gradient without attenuation at $x > 0$ to alleviate the gradient disappearance problem, but at $x < 0$, the gradient will also disappear, resulting in the corresponding weight cannot be updated, affecting the convergence of the network. Leaky-ReLU activation function can effectively solve the gradient disappearance problem, and its expression is as follows.

$$f(x) = \begin{cases} x, & \text{if } x > 0, \\ \alpha(e^x - 1), & \text{if } x < 0. \end{cases} \quad (14)$$

The function of the pooling layer is to sample the features obtained from the previous convolution layer to reduce the dimension, reduce the computational complexity, and avoid overfitting. The general form of pooling is described as follows:

$$x_{i,j}^l = f(\beta^l; \text{down}(x_{i,j}^{l-1}) + b), \quad (15)$$

where β is a multiplicative bias and $\text{down}(\cdot)$ is the down-sampling function.

Softmax function is a generalization of logistic classifier, which is mainly used for multiclassification problems. Suppose the set of training output samples is $X = \{x_1 \cdots x_i \cdots x_T\}$, where the category of input sample elements is $C = \{c_1 \cdots c_i \cdots c_K\}$ and the corresponding label is \bar{y} , so the probability of judging input sample x_i as a category in C set is $P = \{y_i = c_k | x_i\}$. The mathematical expression of Softmax function is described as follows:

$$\begin{aligned} P &= \{y_i = c_k | x_i\} \\ &= e^{x_i^{T c_k}} \frac{1}{\sum_{k=1}^K e^{x_k^{T c_k}}}, \end{aligned} \quad (16)$$

where $e^{x_i^{T c_k}}$ is the correlation coefficient between category c_k and the whole x_i classification category and $1/\sum_{k=1}^K e^{x_k^{T c_k}}$ is the normalization function.

3. Experiment and Parameter Design

3.1. Experimental Setup. VDL-1000E three-axis NC machine tool produced by the Dalian machine tool group was used for metal material cutting. The solid carbide ball end mill is SH300-B2-10015-H produced by the Xiamen Golden Heron Company with a diameter of 10 mm and the tooth number of 2. Cr12MoV harden steel with 45HRC hardness was selected as experimental. Cr12MoV hardened steel is mainly used for automobile panel dies. The chemical composition of the Cr12MoV steel parts is shown in Table 1. The workpiece is fixed on the fixture through M10 bolts, and the fixture is fixed with the machine tool workbench through vice. The included angle between the workpiece and the horizontal plane is 20.2° (as shown in Figure 1(a)). PCB acceleration sensor with a sensitivity of 10.42 mv/g and DH5922 acquisition system of Donghua were used to collect the acceleration signal of machining process (as shown in Figure 1(b)). VHX-100 ultradepth of field microscope is used to measure the tool flank wear (as shown in Figure 1(c)).

3.2. Parameter Setting. The processing parameters of this experiment are given, including the axial cutting depth is 0.25 mm, the spindle speed is 4000 rpm, the cutting width is 0.3 mm, and the feed rate per tooth is 0.12 mm/tooth. The tool wear status is defined as follows: the normal tool ($VB = 0mm$), the initial wear ($VB = 0.087mm$), the normal wear ($VB = 0.121mm$), and the sharp wear ($VB = 0.164mm$), as shown in Figure 2. The sampling frequency of this experiment is 5000 Hz, and the acceleration

sets under different tool wear states are collected, respectively. The collected data are divided into four status labels according to tool wear. The collected data is divided into training samples, test samples, and verification samples according to the ratio of 3:1:1. The number of sampling points in each section is set to 800. The number of training samples in each state is 1200, the number of test samples is 400, and the number of verification samples is 400.

The structure of the deep convolution neural network constructed in this study is shown in Figure 3. It includes 3 volume layers, 3 pool layers, 1 input layer, and full connection layer. The Adam adaptive optimizer is used to continuously update the network training parameters. The initial learning rate is 0.0001, and the attenuation rate is 0.9. 25 samples are used as a batch input convolutional neural network for training, and the number of iterations is 10. The cross entropy loss function is used to detect the training state of the convolutional neural network. The error function calculates the loss value of each iteration according to the error between the actual value and the expected value during the training period.

4. Result Analysis

4.1. Cutting Vibration with Different Tool States. The acceleration data in X direction with different tool wear states are shown in Figure 4. The maximum amplitude during the normal tool cutting is 8.6 m/s^2 . The maximum amplitude during the initial wear is 13.6 m/s^2 . The maximum amplitude during the normal wear is 15.1 m/s^2 . The maximum amplitude during the sharp wear is 18.3 m/s^2 . In conclusion, cutting vibration increases with the increase of tool wear.

4.2. Time-Frequency Domain Diagram of Different Tool States. Sample data under different tool states are converted into time spectrum by synchronous compressed short-time Fourier transform and synchronous compressed continuous wavelet transform, as shown in Figures 5 and 6.

It can be seen from Figures 5 and 6 that the trend of the tool wears characteristic frequency changes with time. When the total number of samples is large, the time-frequency graphs generated by the samples of the same tool state are still different in different degrees. The efficiency and accuracy of tool state diagnosis cannot be guaranteed by manual operation, so it is of great significance to us in the deep convolutional neural network to extract the features of time-frequency graphs of different tool wear.

In order to compare the time-frequency energy aggregation performance of this method, the time-frequency domain map in the state of sharp war was locally amplified, and the results are shown in Figure 7. Compared with the SST method, SWT method has better time-frequency energy aggregation, which is beneficial to improve time-frequency resolution.

4.3. Comparative Analyses of Recognition Models. In this paper, SST-DCNN method and SWT-DCNN method were used to monitor tool wear. According to the mean and

TABLE 1: Chemical composition table of Cr12MoV steel parts.

Element	C	Si	Mn	P	S	Cr	V	Mo	Ni	Cu
Content (%)	1.45-1.70	≤0.40	≤0.40	≤0.03	≤0.03	11.0-12.50	0.15-0.30	0.40-0.60	≤0.25	≤0.30

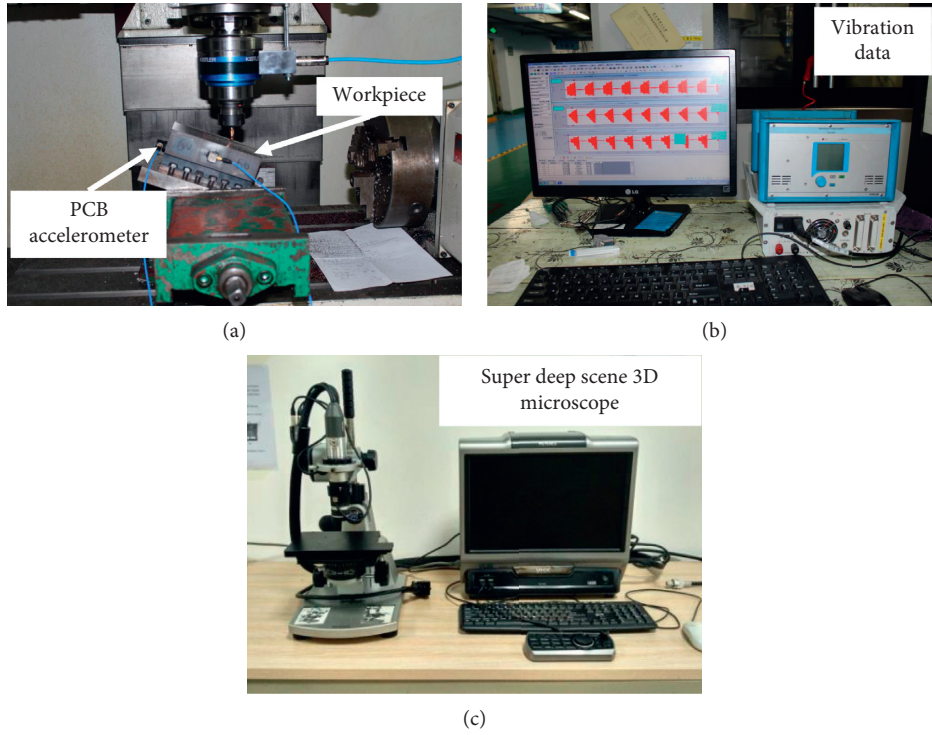


FIGURE 1: Processing and test site.

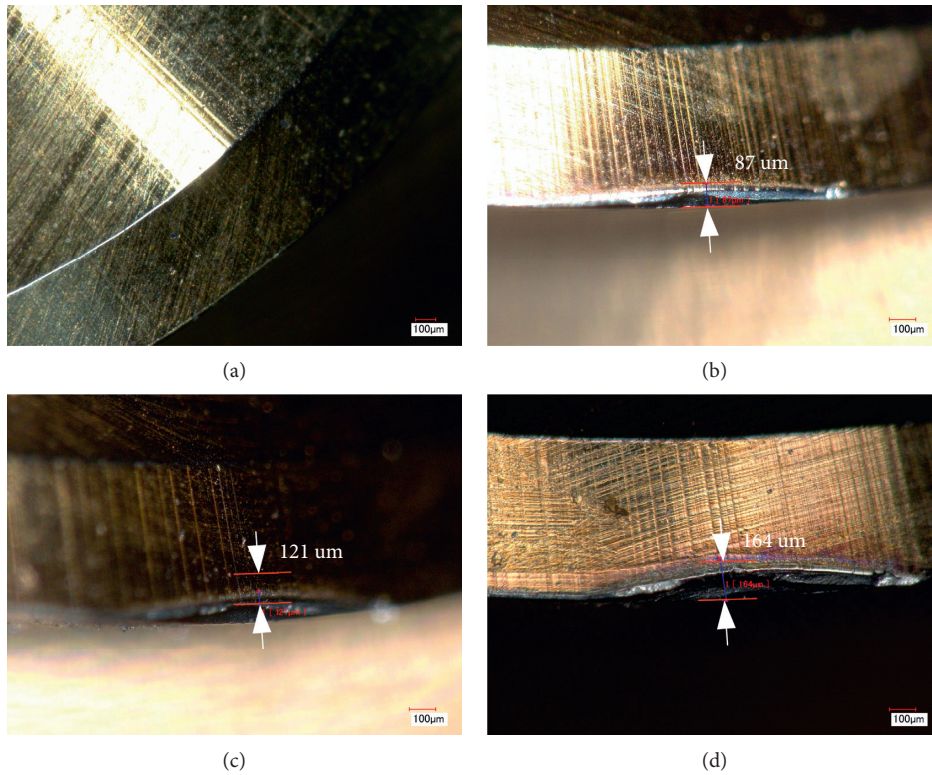


FIGURE 2: Different tool wear states.

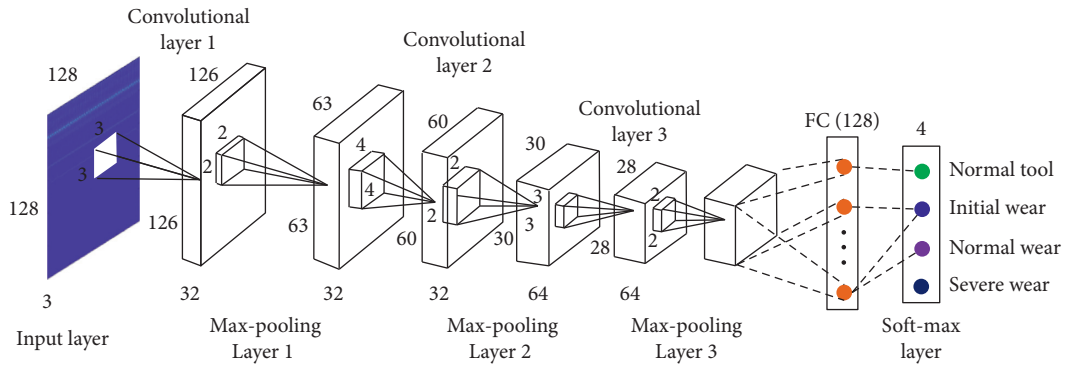


FIGURE 3: Deep convolution neural network structure.

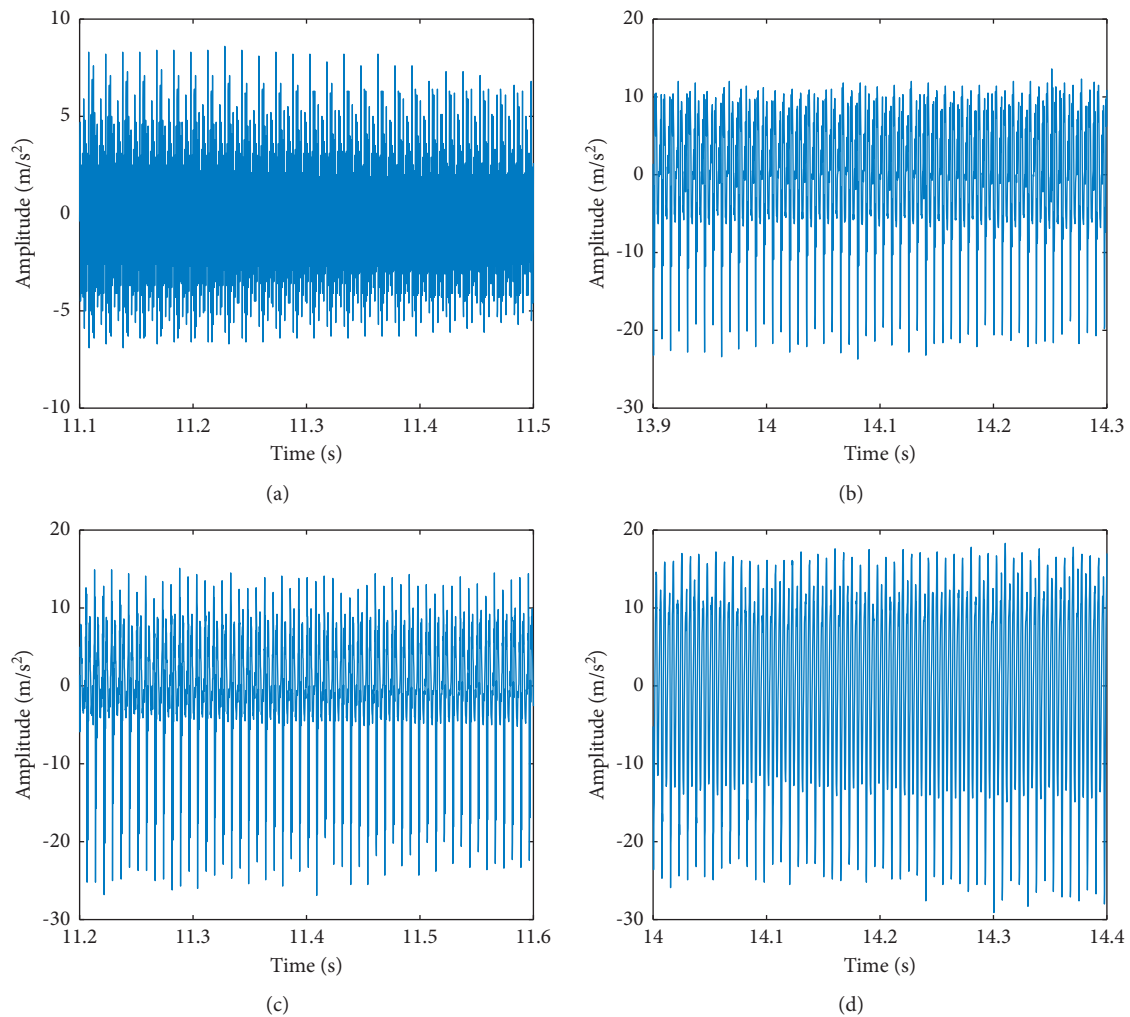


FIGURE 4: Cutting vibration. (a) The normal tool. (b) The initial wear. (c) The normal wear. (d) The sharp wear.

standard deviation of each iteration result of the data sets of the two methods, the loss value curve of the training set, the recognition accuracy of the verification set and the final recognition accuracy are obtained, so as to test the training speed, recognition accuracy, and generalization ability of the two methods. The loss function curve of the training set, recognition accuracy curve of verification set, and final recognition result is shown in Figure 8.

It can be seen from Figure 8(a) that the two methods proposed in this paper tend to converge at the fourth iteration. It can be seen from Figure 8(b) that, after the convergence of the convolution model with SWT-DNCC method, the oscillation of the recognition accuracy curve of the verification set is smaller and the recognition accuracy of the model is more stable. It can be seen from Figure 8(c) that the SWT-DNCC method is superior to the SST-DNCC

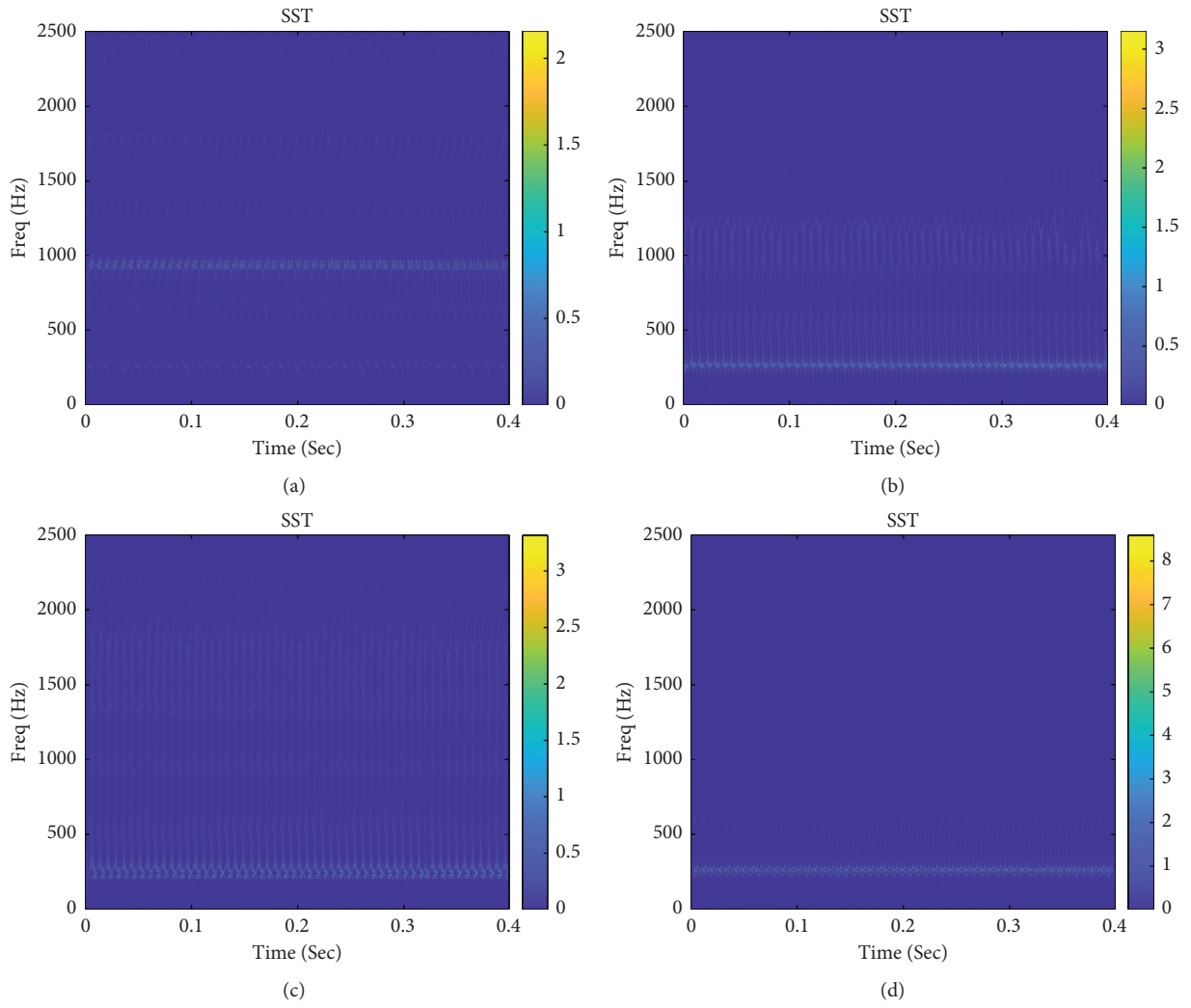


FIGURE 5: Time-frequency domain diagram with the SST method. (a) The normal tool. (b) The initial wear. (c) The normal wear. (d) The sharp wear.

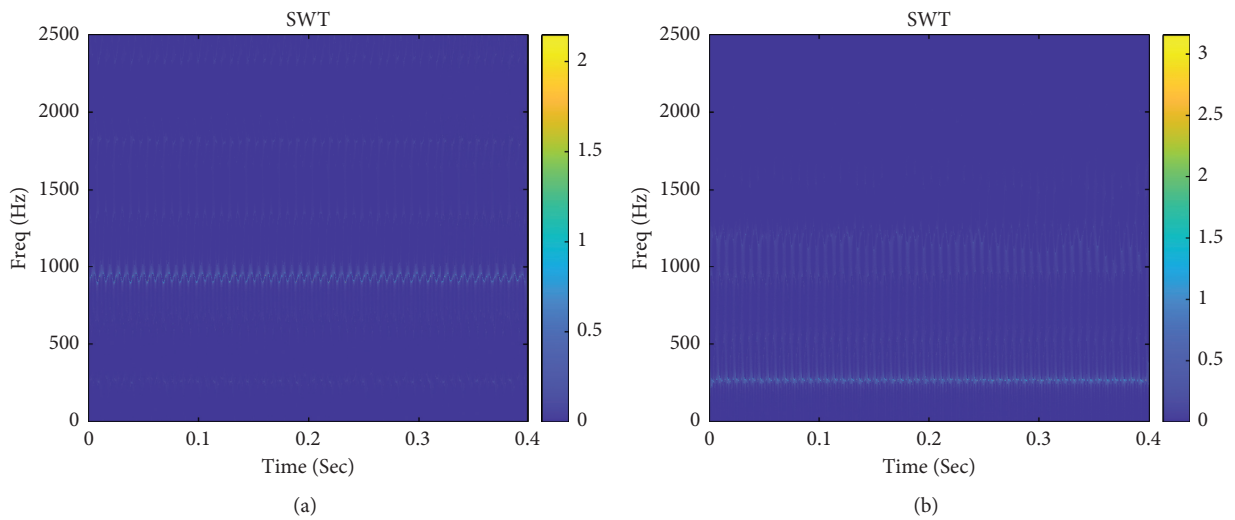


FIGURE 6: Continued.

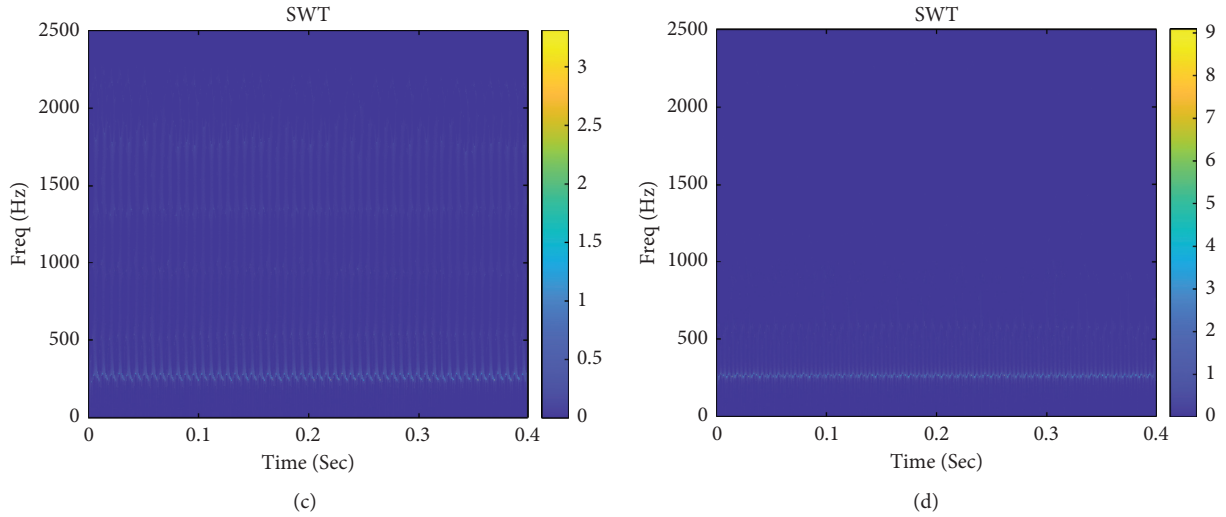


FIGURE 6: Time-frequency domain diagram with the SWT method. (a) The normal tool. (b) The initial wear. (c) The normal wear. (d) The sharp wear.

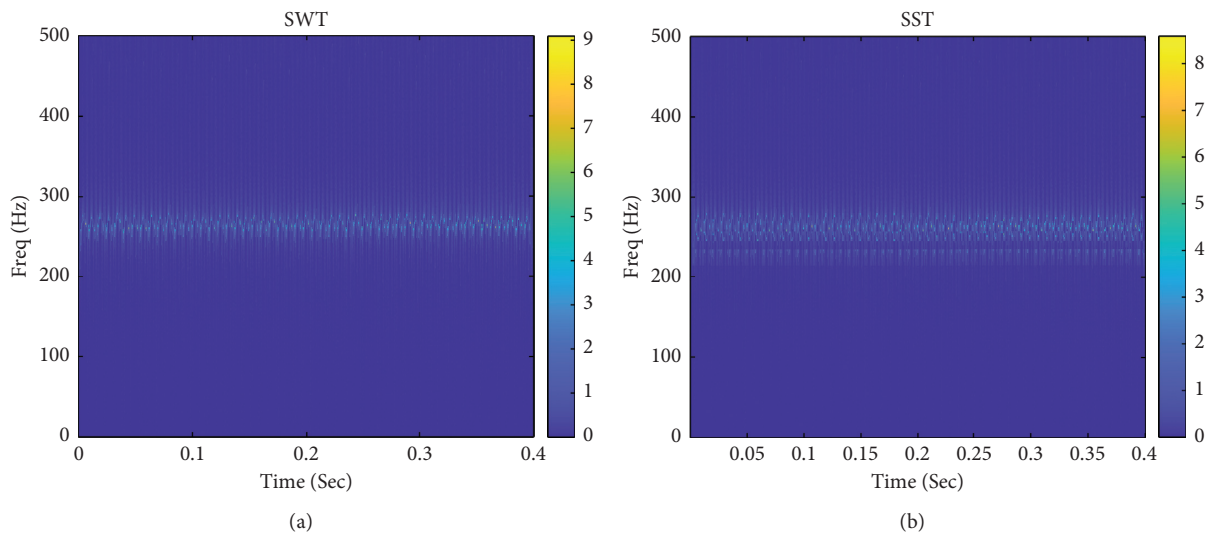


FIGURE 7: Local enlarged view of time-frequency domain with the sharp wear. (a) SWT method. (b) SST method.

method in the recognition accuracy of the training set, verification set, and test set. In conclusion, the characteristic information of the time spectrum obtained by the SWT-DNCC method is easier to be learned and extracted by the convolutional neural network.

In order to further investigate the misjudgment of the tool processing state by the method proposed in this paper, confusion matrix experiments were carried out on the test results, and the results are shown in Figure 9.

As can be seen from Figure 9(a), the diagnostic accuracy of the SST-DCNN method for normal tools is only 93.3%, and the probability of disqualifying normal tools as initial wear is 7.6%. The diagnostic accuracy of SWT-DCNN method is 98.1% and 99.6%, respectively. It can be seen from Figure 9(b) that the diagnostic accuracy of SWT-DCNN method is 97.5%, 99%, and 99.7% for normal tool, initial

wear tool, and normal wear tool, respectively. In conclusion, the diagnostic accuracy of SWT-DCNN method is higher than that of SST-DCNN method in various tool status.

In order to further verify the advantages of the method proposed in this paper, the experimental results of the method proposed in this paper are compared with those of the method of short-time Fourier transform combined with the deep convolution neural network (STFT-DCNN) and the method of continuous wavelet transform combined with deep convolution neural network (CWT-DCNN). The results are shown in Table 2.

It can be seen from Table 2 that the characteristic information of SWT and SST time-frequency diagrams is easier to be learned and extracted by the convolutional neural network. Therefore, the method proposed in this paper has higher recognition accuracy.

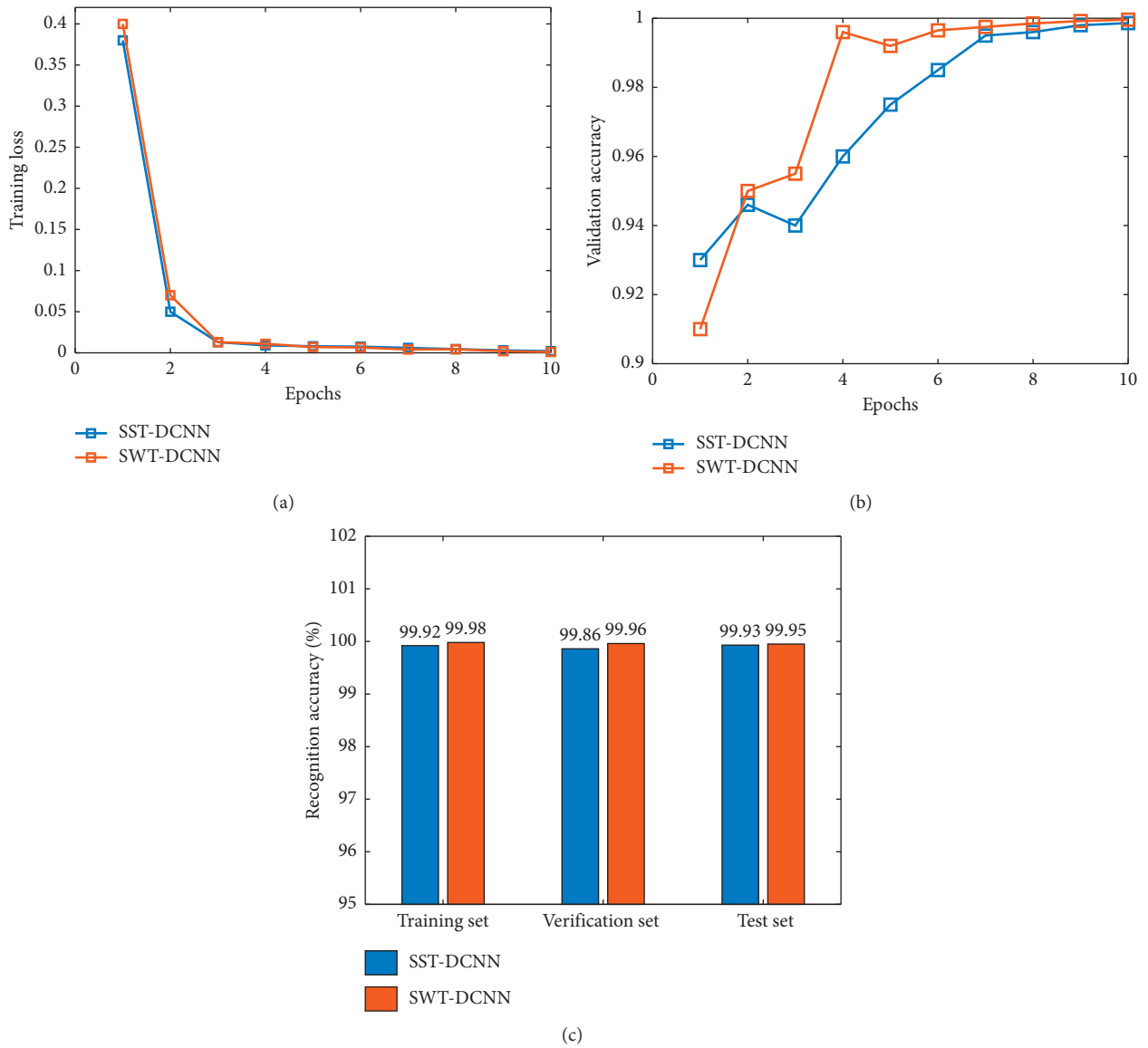


FIGURE 8: Loss function and recognition accuracy. (a) Value of the loss function (b) Identification accuracy of verification set. (c) Identification accuracy.

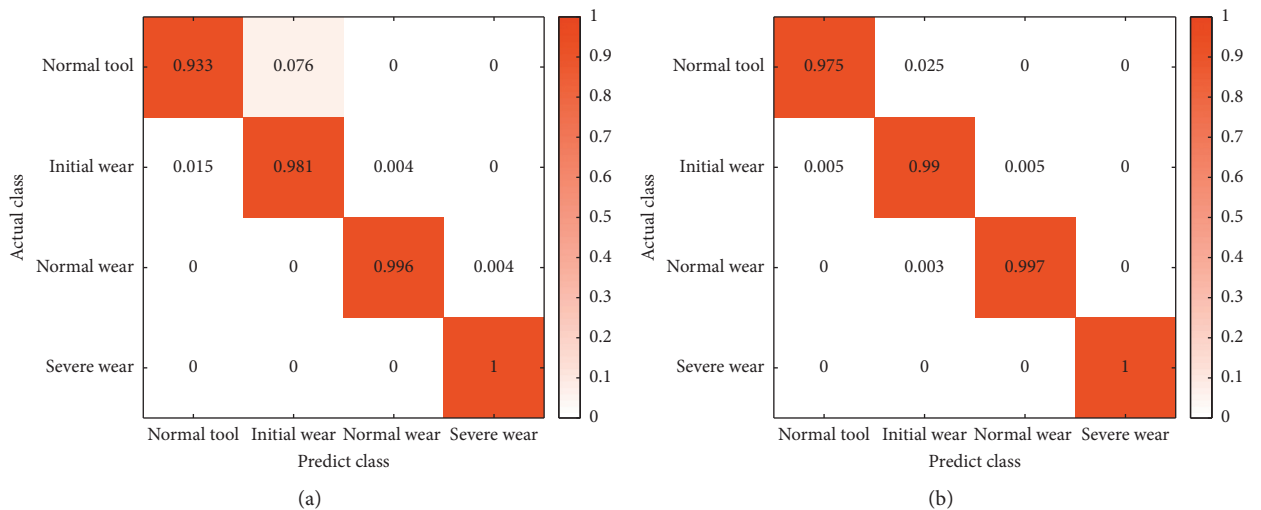


FIGURE 9: Different methods of confusion matrix. (a) SST-DCNN method. (b) SWT-DCNN method.

TABLE 2: Comparative analysis of recognition accuracy results.

Method	Recognition accuracy (%)		
	Training set	Verification set	Test set
SWT-DCNN	99.98	99.96	99.95
SST-DCNN	99.92	99.86	99.93
STFT-DCNN	99.82	99.78	99.89
CWT-DCNN	99.9	99.81	99.91

5. Conclusion

In this study, SWT-DCNN and SST-DCNN models were constructed to monitor the tool state during cutting. Experimental results show that the two models have high diagnostic accuracy. The specific conclusions are as follows:

- (1) Compared with the SST method, the characteristic information of the time spectrum obtained by the SWT method is easier to be learned and extracted by the convolutional neural network. The recognition accuracy of SWT-DCNN method is 99.96% and that of SST-DCNN method is 99.86%.
- (2) The diagnostic accuracy of the SST-DCNN method for normal tools is only 93.3%, which is easy to classify normal tools into the initial wear category.
- (3) Compared with SST method, the SWT method has better time-frequency energy aggregation, which is conducive to improving time-frequency resolution.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was funded by the Science and Technology Planning Project in Henan Province (Grant no. 212102210326), the Key Research Projects of Henan Higher Schools (Grant no. 21B460007), and Youth project of national scientific research project cultivation fund of Huan-guai University (Grant no. XKPY-202104).

References

- [1] R. Peng, J. Liu, X. Fu, L. Cuiya, and Z. Linfeng, "Application of machine vision method in tool wear monitoring," *International Journal of Advanced Manufacturing Technology*, vol. 116, no. 3, pp. 1357–1372, 2021.
- [2] J. Ma, D. Luo, X. Liao, Z. Zhang, Y. Huang, and J. Lu, "Tool wear mechanism and prediction in milling TC18 titanium alloy using deep learning," *Measurement*, vol. 173, Article ID 108554, 2021.
- [3] Z. Lei, Q. Zhu, Y. Zhou, B. Sun, W. Sun, and X. Pan, "A GAPSO-enhanced extreme learning machine method for tool wear estimation in milling processes based on vibration signals," *International Journal of Precision Engineering and Manufacturing-Green Technology*, vol. 8, no. 3, pp. 745–759, 2021.
- [4] Y. Zhou, B. Sun, W. Sun, and Z. Lei, "Tool wear condition monitoring based on a two-layer angle kernel extreme learning machine using sound sensor for milling process," *Journal of Intelligent Manufacturing*, vol. 33, pp. 1–12, 2020.
- [5] C. Drouillet, J. Karandikar, C. Nath, A.-C. Journeaux, M. El Mansori, and T. Kurfess, "Tool life predictions in milling using spindle power with the neural network technique," *Journal of Manufacturing Processes*, vol. 22, pp. 161–168, 2016.
- [6] M. Plogmeyer, G. González, V. Schulze, and G. Bräuer, "Development of thin-film based sensors for temperature and tool wear monitoring during machining," *Tm - Technisches Messen*, vol. 87, no. 12, pp. 768–776, 2020.
- [7] X. Cui, K. Yan, and J. Guo, "Bio-inspired fabrication, mechanical characterization and cutting performance evaluation of Al₂O₃/TiC micro-nano-composite ceramic with varying microscopic surfaces," *Ceramics International*, vol. 45, no. 7, pp. 8286–8299, 2019.
- [8] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, and J. Lin, "Machinery health prognostics: a systematic review from data acquisition to RUL prediction," *Mechanical Systems and Signal Processing*, vol. 104, pp. 799–834, 2018.
- [9] C. Zhang and H. Zhang, "Modelling and prediction of tool wear using LS-SVM in milling operation," *International Journal of Computer Integrated Manufacturing*, vol. 29, no. 1, pp. 76–91, 2016.
- [10] K. Zhu and T. Liu, "Online tool wear monitoring via hidden semi-Markov model with dependent durations," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 1, pp. 69–78, 2017.
- [11] T. Liu and K. Zhu, "A switching hidden semi-Markov model for degradation process and its application to time-varying tool wear monitoring," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2621–2631, 2020.
- [12] H. Gao, H. Shen, L. Yu, W. Yinling, R. Li, and B. Nazir, "Milling chatter detection system based on multi-sensor signal fusion," *IEEE Sensors Journal*, vol. 21, no. 22, Article ID 25243, 2021.
- [13] C. Aliustaoglu, H. M. Ertunc, and H. Ocak, "Tool wear condition monitoring using a sensor fusion model based on fuzzy inference system," *Mechanical Systems and Signal Processing*, vol. 23, no. 2, pp. 539–546, 2009.
- [14] M. Rizal, J. A. Ghani, M. Z. Nuawi, and C. H. C. Haron, "Cutting tool wear classification and detection using multi-sensor signals and Mahalanobis-Taguchi System," *Wear*, vol. 376–377, pp. 1759–1765, 2017.
- [15] C. Liu, Y. Li, G. Zhou, and W. Shen, "A sensor fusion and support vector machine based approach for recognition of complex machining conditions," *Journal of Intelligent Manufacturing*, vol. 29, no. 8, pp. 1739–1752, 2018.
- [16] H. Li, W. Wang, Z. Li, L. Dong, and Q. Li, "A novel approach for predicting tool remaining useful life using limited data," *Mechanical Systems and Signal Processing*, vol. 143, Article ID 106832, 2020.
- [17] Q. An, Z. Tao, X. Xu, M. El Mansori, and M. Chen, "A data-driven model for milling tool remaining useful life prediction with convolutional and stacked LSTM network," *Measurement*, vol. 154, Article ID 107461, 2020.
- [18] Z. Huang, J. Zhu, J. Lei, L. Xiaoru, and T. Fengqing, "Tool Wear Monitoring with Vibration Signals Based on Short-Time Fourier Transform and Deep Convolutional Neural

- Network in milling,” *Mathematical Problems in Engineering*, vol. 2021, Article ID 9976939, 14 pages, 2021.
- [19] Z. Chen, A. Mauricio, W. Li, and K. Gryllias, “A deep learning method for bearing fault diagnosis based on Cyclic Spectral Coherence and Convolutional Neural Networks,” *Mechanical Systems and Signal Processing*, vol. 140, Article ID 106683, 2020.
- [20] M. Q. Tran, M. K. Liu, and Q. V. Tran, “Milling chatter detection using scalogram and deep convolutional neural network,” *International Journal of Advanced Manufacturing Technology*, vol. 107, no. 3, pp. 1505–1516, 2020.
- [21] H. N. Gao, D. H. Shen, L. Yu, and W. Zang, “Identification OF cutting chatter through deep learning and classification,” *International Journal of Simulation Modelling*, vol. 19, no. 4, 2020.
- [22] G. Thakur and H.-T. Wu, “Synchrosqueezing-based recovery of instantaneous frequency from nonuniform samples,” *SIAM Journal on Mathematical Analysis*, vol. 43, no. 5, pp. 2078–2095, 2011.
- [23] S. Wang, X. Chen, G. Cai, X. Li, and Z. He, “Matching demodulation transform and synchrosqueezing in time-frequency analysis,” *IEEE Transactions on Signal Processing*, vol. 62, no. 1, pp. 69–84, 2013.
- [24] I. Daubechies, J. Lu, and H.-T. Wu, “Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool,” *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 243–261, 2011.
- [25] W. Zhao, S. Li, A. Li, B. Zhang, and Y. Li, “Hyperspectral images classification with convolutional neural network and textural feature using limited training samples,” *Remote sensing letters*, vol. 10, no. 5, pp. 449–458, 2019.