



Combining translation-invariant wavelet frames and convolutional neural network for intelligent tool wear state identification

Xin-Cheng Cao^a, Bin-Qiang Chen^{a,*}, Bin Yao^a, Wang-Peng He^b

^a School of Aerospace Engineering, Xiamen University, Xiamen 361005, PR China

^b School of Aerospace Science and Technology, Xidian University, Xi'an 700071, PR China



ARTICLE INFO

Article history:

Received 29 August 2018

Received in revised form 21 December 2018

Accepted 30 December 2018

Available online 9 January 2019

Keywords:

Tool wear state monitoring
Spindle vibration
End milling
Wavelet frame
Convolutional neural network (CNN)

ABSTRACT

On-machine monitoring of tool wear in machining processes has found its importance to reduce equipment downtime and reduce tooling costs. As the tool wears out gradually, the contact state of the cutting edge and the workpiece changes, which has a significant influence on the vibration state of the spindle. The performance of traditional intelligent fault diagnosis methods depend on feature extraction of dynamic signals, which requires expert knowledge and human labor. Recently, deep learning algorithms have been applied widely in machine health monitoring. In this paper, we present a novel intelligent technique for tool wear state recognition using machine spindle vibration signals. The proposed technique combines derived wavelet frames (DWFs) and convolutional neural network (CNN). Constructed based on dual tree wavelets, DWF are equipped with merits of centralized multiresolution and nearly translation-invariance. In this method, DWFs are employed to decompose the original signal into frequency bands of different bandwidths and different center frequencies, which are more pronounced as the tool wears. Further, the reconstructed sub-signals are stacked into a 2-D signal matrix to match the structure of 2-D CNN while retaining more dynamic information. The 2-D convolutional neural network is utilized to automatically recognize features from the multiscale 2-D signal matrix. End-milling experiments were performed on a S45C steel workpiece at different machining parameters. The experiment results of the recognition for tool wear state show the feasibility and effectiveness of the proposed method.

© 2019 Elsevier B.V. All rights reserved.

1. Introduction

Cutting tool plays a very important role in modern manufacturing industry. Tool wear, as a result of the interaction between tool and workpiece, is one of the major limitation to machining processes. As early as 1990, some surveys showed that, on average, 6.8% of the downtime was caused by tool failure [1]. As the automation of the manufacturing environment increased, this proportion quickly rose to 20% [2]. It is predicted that an accurate and reliable tool condition monitoring system will help maximize the effective working life of the tool and reduce downtime by pre-arranging tool change time nodes [3–5]. Therefore, tool condition monitoring and tool wear prediction have received more and more attention.

A typical intelligent monitoring system is usually consists of four parts: signal acquisition, data processing, feature engineering,

and monitoring model. Signal acquisition aims at collecting original signal by sensors such as dynamometer, accelerometer, acoustic emission (AE) sensors, current/power sensors and so on. The dynamometer is the most widely explored in research [6,7]. Sanchez monitored the cutting force during dry machining of titanium alloys, and the correlation between cutting force and tool wear is very obvious [8]. AZMI uses cutting force to monitor tool health during milling of GFRP composites [9]. However, the severe intrusiveness of the dynamometer to production makes it unsuitable for manufacturing environment. NAKI presented an intelligent system composed of four neural networks to process AE signals, the RBF neural network has the highest recognition accuracy and reaches more than 95% for all three wear states [10]. However, the extremely high sampling frequencies required by AE measurement is not conducive to online monitoring. The obvious advantage of the power sensor is that it does not interfere with the production process, and many modern machine tools are already equipped with power sensors. NIAKI presented an extended Kalman filter based on spindle power for tool wear, which effectively improves the prediction accuracy [11]. The same method is used to monitor the wear of the milling cutter during

* Corresponding author at: 422 South Si'ming Road, Siming District, Xiamen 361005, Fujian, PR China.

E-mail address: cbq@xmu.edu.cn (B.-Q. Chen).

milling the gamma-prime strengthened alloys, and the prediction error is not more than 18% [12]. Vibration monitoring has also been applied in the diagnosis of tool wear, as the amplitude and effective value of the vibration signal increase significantly as the tool wears. Yan R et al. introduced a closed loop calibration system for vibration sensors to improve calibration accuracy and efficiency [13]. Hsieh extracted five features from the [0,30] Hz spectrum of the spindle vibration signal to train the neural network, and the recognition accuracy of the micro-milling wear state is reported to be 100% [14]. Juho Ratava et al. measured the vibration signal during the machining process of impeller cast, and established a model to monitor the damage of the tool. After a series of variable parameter experiments, the average recognition accuracy of the model reached 80% [15].

An appropriate signal processing must be performed before feature extraction, which basically includes data segmentation and digital filtering. Yan R et al. proposed an adaptive partition window to accumulate all local feature information and align them [16]. The original signal may be also converted into frequency domain using Fourier transform, or time-frequency domain by wavelet transform or wavelet packet transform [17]. Wavelet transform utilizes a decaying wave atom, whose translation and scaling form multi-resolution analysis (MRA) [18]. Cui et al. used db3 wavelet to effectively extract the fault impact waveform of the rolling element bearing from the vibration signal with low SNR [19]. The dual tree complex wavelet transform (DTCWT) is a recent enhancement to conventional discrete wavelet transforms with the attractive properties of approximate shift invariance and inhibited frequency aliasing [20]. Chen et al. proposed an iterated method for constructing a dual tree complex wavelet base with the enhanced frequency aliasing property [21]. Wang et al. demonstrated the robustness of DTCWT in extracting vibration features interfered by strong noises [22]. Empirical mode decomposition (EMD) and multi-scale analysis have also achieved high degree of empirical success.

Feature engineering aims at transforming the processed signal data into several key feature parameters. Statistical method is the most commonly used technique to extracting features from processed signals in both the time domains (mean, root mean square (RMS), standard deviation, skew, kurtosis, ratio of different directions, etc.) and the frequency domain (single harmonic power, harmonic ratio, power spectral amplitudes, power spectral density, etc.) [23,24]. Then, we can select a subset from the original feature set using methods such as correlation index ranking [25,26] or genetic algorithms [27]. It is also possible to extract a feature vector with a lower dimension from the original feature value set by means of principal component analysis (PCA) [28] or locally linear embedding (LLE) and so on. Finally, the tool wear condition can be evaluated with feature vector parameters based on one certain monitoring model, such as artificial neural network (ANN) [7,29–31], Fuzzy inference systems [32,33], hidden Markov model (HMM) [34,35], and support vector machine (SVM) [36,37].

Although it has achieved some success in some problems, there is a clear flaw in the typical intelligent monitoring system. It greatly depends on whether the features have exactly caught the critical property of the tool wear conditions [17]. Due to the limitation of modeling capability, the traditional modeling models cannot process inputs in very high dimension. This means that uncertain actual conditions require different feature vectors, and manual adjustments are often required, which puts high demands on the individual expertise and extensive experience of the operator and reduces the applicability and reliability of the entire tool condition monitoring system.

With the development of artificial intelligence technologies, the feature engineering that requires a lot of manual intervention may be not essential. The new network structure, new training strategies, and improved computer computing capabilities enable the deep

learning model to directly adopt high-dimensional input training models with great performance within acceptable time cost. Convolutional neural networks (CNNs), have shown their success in various computer vision applications [38,39], where input data are usually 2-D data. CNN models can also process 1-D signals [40,41].

Recently, the application of CNN network model in mechanical fault diagnosis has also been successful. Sun W et al. proposed an intelligent gear fault diagnosis method using a complex wavelet enhanced CNN, and the test accuracy was reported as 99.79% [42]. Wang et al. presented an adaptive CNN for rolling element bearing fault diagnosis, and confirmed that the CNN-based method is more effective than other intelligent methods [43]. Janssens applied CNN to the fault diagnosis of rotating machinery, and achieved great classification results without extensive knowledge [44]. Jing proposed an adaptive multi-sensor data fusion method based on CNN to detect faults in a planetary gearbox [45]. Wang proposed a CNN-based hidden Markov model, and the accuracy of bearing fault classification is reported as 98% [46].

Based on the above analysis, this paper proposes an intelligent tool wear condition monitoring method based on redundant wavelet expansion and deep learning techniques. 2-D vibration images in matrix form, transformed from the original 1-D spindle vibration signal via derived wavelet frames (DWFs), are used as the input of the model. Different signal representation forms and different modeling methods will be compared to demonstrate the performance of this proposed method.

2. Translation-invariant feature extraction using derived wavelet frames

Transient signatures are important in indicating dynamic process relevant to tool wear and breakage. Due to lack of shift-invariance, conventional wavelet transforms are comparatively weak in identifying repetitive transient vibration [47,48].

In this section, we introduce derived wavelet frames (DWFs), which realizes nearly translation-invariant multiresolution analysis. DWFs incorporate dyadic dual-tree complex wavelet packets and additional nondyadic implicit wavelet packets. The latter ones are beneficial for extracting transition band features.

2.1. Dual-tree complex wavelet packet decomposition

Dual-tree complex wavelet packet decomposition (DCWPD) is constructed based on dual-tree complex wavelet basis, which consists of two scaling functions and two wavelet functions. In orthonormal cases, for the complex valued wavelet function

$$\psi^{\mathbb{C}}(t) = \psi^{\Re}(t) + j \cdot \psi^{\Im}(t) \quad (1)$$

there is a restriction of approximate Hilbert transform shown as

$$\psi^{\Im}(t) = H\{\psi^{\Re}(t)\} \quad (2)$$

where $\psi^{\Re}(t), \psi^{\Im}(t)$ are real part and imaginary part of $\psi^{\mathbb{C}}(t)$ respectively, and $j = \sqrt{-1}$ is the imaginary unit. Equivalently, a half sample delay equation exists for impulse response functions of wavelets ($h_1^{\Re}(t), h_1^{\Im}(t)$).

$$h_1^{\Im}(t)(n) \approx h_1^{\Re}(n - 0.5) \quad (3)$$

Let the \mathbb{Z} transform of a discrete series $\{x(n)\}$ be represent as

$$X(z) = Z\{x(n)\} = \sum_{n=-\infty}^{+\infty} x(n)z^{-n} \quad (4)$$

The more similar is the filter-bank function to the original signal to be analyzed, the better the defect-related features will be extracted [49]. In this research, the filter-bank structure of dual

tree complex wavelet packet decomposition (DCWPD) is shown in Fig. 1. The notation ' \cdot ' represents \Re or \Im . That is, the two filter branches share the same filtering structure but without data-flow in between. More details about the two-scale relationship can be found in articles written by Kingsbury and Selesnick who initially proposed dual tree complex wavelet transform [50].

As shown in Fig. 1, DCWPD can be regarded as two separate tree structured iterative decompositions using special hybrid wavelet bases. The related time-frequency atoms are shown in Fig. 2. The employed filters can be classified into three categories:

i) Wavelet basis at the initial decomposition stage. These filters are $\{h_{10}^{\Re}(n), h_{11}^{\Re}(n), h_{10}^{\Im}(n), h_{11}^{\Im}(n)\}$ (Fig. 2(a, b)), and they satisfy the following equation.

$$\begin{cases} h_{10}^{\Im}(n) = h_{10}^{\Re}(n-1) \\ h_{11}^{\Im}(n) = h_{11}^{\Re}(n-1) \end{cases} \quad (5)$$

ii) Conventional dual-tree complex wavelet basis to obtain wavelet series $\{d_{k,j}^{\mathbb{C}}(n) | k \geq 2, j = 1, 2\}$. The associated filters are $\{h_0^{\Re}(n), h_1^{\Re}(n), h_0^{\Im}(n), h_1^{\Im}(n)\}$ (Fig. 2(c, d)).

iii) Additional basis to generate extended wavelet packet series $\{d_{k,j}^{\mathbb{C}}(n) | k \geq 2, j \geq 2\}$. The associated filters are

$$\begin{cases} f_0(n) = h_{10}^{\Im}(n) \\ f_1(n) = h_{11}^{\Re}(n-1) \end{cases} \quad (6)$$

2.2. Strategy for constructing implicit wavelet packets

The frequency response of dyadic wavelet is not of ideal brick wall shape. Therefore, performances of conventional wavelet packets are not perfect in extracting incipient vibration signatures located in transition bands. In order to address this problem, implicit wavelet packets are constructed based on DCWPD. Let the input vibration signal be denoted as $\{x(n)\}$, the construction strategy for implicit wavelet packets (IWP) is proposed as follow.

Step 1) Perform multi-scale decomposition based on dual tree wavelet packet decomposition on the input signal, such that $\{x(n)\}$ is transformed into a set

$$D_k = \{D_k^j(n) | j = 1, 2, \dots, 2^k\}$$

Step 2) Rearrange the elements in the set D_k . Let the resulting set be $R_k = \{R_k^j(n) | j = 1, 2, \dots, 2^k\}$. Elements in the new set are arranged according to order of central frequency of wavelet packet.

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

For R_k^j , let the binary coding of the index j be

$$j = \sum_{m=0}^{k-1} 2^m n_m + 1 \quad (7)$$

A new integer j' is expressed as

$$j' = \sum_{m=0}^{k-1} 2^m n'_m + 1 \quad (8)$$

where the parameters n'_m is defined as

$$n'_m = \begin{cases} n_m, m = k-1 \\ \text{mod}(n_m + n_{m+1}, 2), m = 0, 1, \dots, k-2 \end{cases} \quad (9)$$

Step 3) Generate the implicit wavelet packet using the following equation

$$iwp_k^j(n) = R_k^{2j}(n) + R_k^{2j+1}(n), 1 \leq 2^{k-1} - 1 \quad (10)$$

2.3. Frequency-scale paving of derived wavelet frames

Different from DCWPD. Implicit wavelet packets realize multiresolution analysis around fixed central frequencies. As shown in Fig. 3, for instance, elements in the following set

$$IWPS_{1,1} = \{iwp_k^j | k \in \mathbb{Z}, k \geq 1, j = 2^{k-1}\} \quad (11)$$

share a common central frequency $f_s/4$, where f_s denotes sampling frequency of the signal. While at the next decomposition stage, two

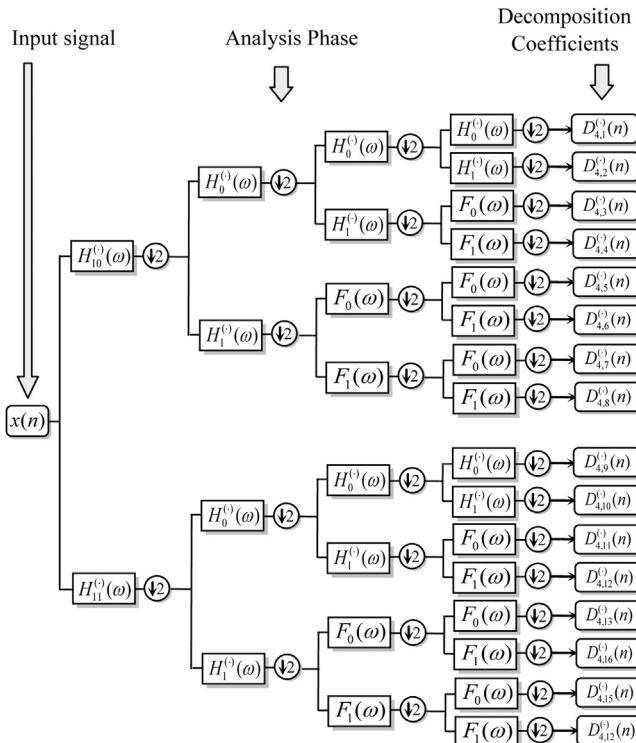


Fig. 1. Filter-bank structure of dual tree wavelet packet decomposition.

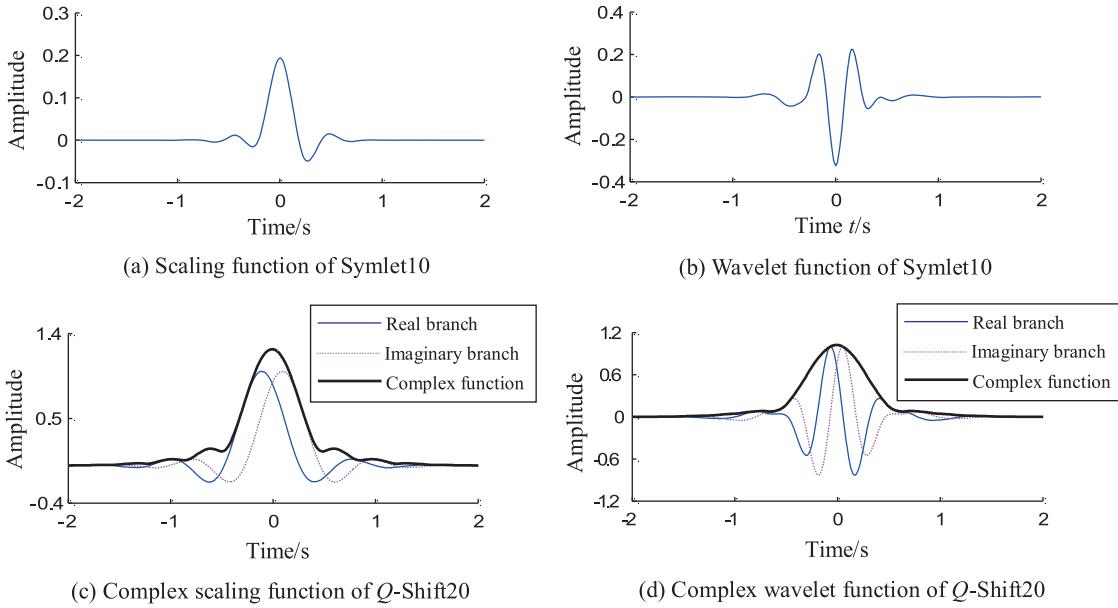


Fig. 2. Hybrid wavelet bases of dual tree complex wavelet packet decomposition.

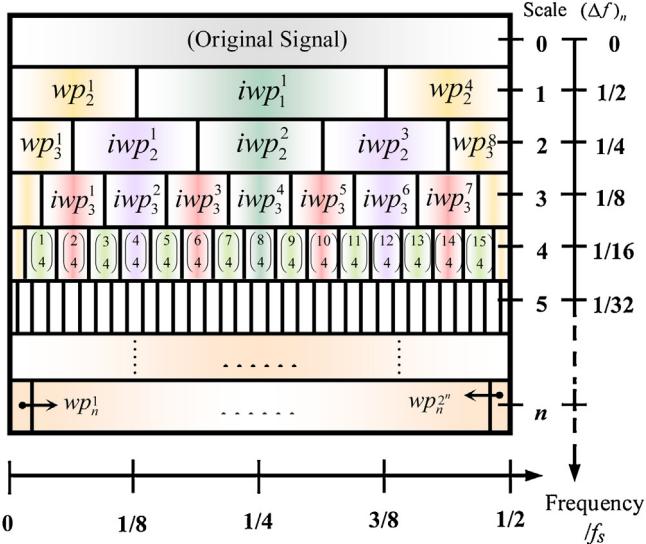


Fig. 3. Frequency-scale paving of implicit wavelet packets.

additional such set, $IWPS_{1,1}$ and $IWPS_{1,2}$, are generated. Mathematical definitions of such sets can be expressed as

$$IWPS_{kj} = \{iwp_{k+k'-1}^{(2j-1)\times 2^{k'-1}} \mid k' \in \mathbb{Z}, k' \geq 1\} \quad (12)$$

$$(k, j \in \mathbb{Z}, k \geq 1, j \geq 1, j \leq 2^{k-1})$$

The initial element in this set is $iwp_k^{1+(j-1)2^{k-1}}$. The passing-band width of this initial element is $f_s/2^{k+1}$. The temporal responses in an IWP set are also approximate similar with each other. In Fig. 4 we plot the first four wavelet functions of $IWPS_{1,1}$. Slight asymmetry appears in the side bands of the implicit wavelets due to asymmetry orthonormal basis at the first decomposition stage.

3. Fundamentals of convolution neuro network

The convolution neural network (CNN) employs multiple processing layers to process the input data, benefiting from the

greedy learning process to obtain essential and robust features, while reducing the training parameters by local field sensing, weight replication, and sub-sampling. This section will introduce the basic architecture and training process of CNN neural network.

3.1. Basic structure of CNN

As shown in Fig. 5, there exist three main layers in the architecture, the input layer, the hidden layer and the output layer. The hidden layer consists of a series of sub-perception units (indicated by the dashed outline on the wireframe) connected by a specific topology. Typically, a sub-perception unit consists of a convolution layer and a pooling layer. Through the convolution layer, we get a series of feature maps, whose dimensions are much smaller than the original input. Subsequently, the feature maps are subject to the activation function, and finally run through the pooling layer. The latter is primarily responsible for reducing the dimension of the output feature map by subsampling. In the output layer, all of the neurons are fully connected to a classifier, such as softmax regression, and finally the classifier gives the final classification.

3.2. Convolution layer

The convolution layers employ a series of learnable filter (or named kernel) to convolve with the input then generates a series of distinctive feature maps by the activation unit. From the perspective of the transfer function, the filter is a linear invariant system. For a clear description, some common notations are listed below: $\mathbf{x}_i^{S_{l-1}}$ denotes the i -th feature map generated from $(l-1)$ -th pooling layer; $\mathbf{w}_j^{C_l}$ denotes the weight matrix of the j -th filter in the l -th convolution layer; $\mathbf{b}_j^{C_l}$ denotes the j -th element of the bias of the l -th layer; $f(\bullet)$ denotes the activation function, and $*$ denotes the 2-D convolution operation. The j -th feature map generated from l -th convolution layer can be represented as the following:

$$\mathbf{x}_j^{C_l} = f \left(\sum_{i \in M_j} \mathbf{x}_i^{S_{l-1}} * \mathbf{w}_j^{C_l} + \mathbf{b}_j^{C_l} \right) \quad (13)$$

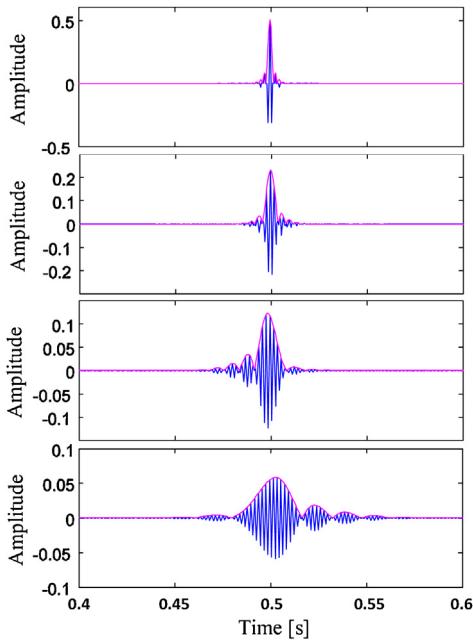


Fig. 4. Temporal responses of the first four implicit wavelet functions.

where M_j denotes a selection of feature maps generated from $(l-1)$ -th pooling layer. Thus, as information is propagated forward in the hidden layer, different detail features can be combined to generate more advanced features of a larger scale.

The activation function is the core of machine learning, which turns the model into a nonlinear model to enhance the expressive power of the model. *ReLU* is a simple and effective activation function, which is defined as follows:

$$\text{ReLU}(x) = \begin{cases} x & x > 0 \\ 0 & x \leq 0 \end{cases} \quad (14)$$

As shown in Fig. 6(a), when the input is greater than 0, the gradient is always equal to 1, so that it can converge quickly; when the input is less than 0, the output is always equal to 0, the so-called one-sided suppression, which endows the model with good sparsity. Compared to the saturating nonlinearities such as sigmoid and tangent functions, the *ReLU* function can effectively solve the gradient diffusion problem and show better generalization capacity with respect to large scale training datasets [51]. However, the compulsory sparseness also reduces the effective capacity of

the model, so there are gradually many improvements, such as Parametric Rectifier, as shown as follows:

$$f(x) = \begin{cases} x & x > 0 \\ \alpha x & x \leq 0 \end{cases} \quad (15)$$

If α is fixed at -0.01, it is called *LReLU* (Fig. 6(b)). If it is a Gaussian random number, it is called Random Rectifier (*RReLU*, Fig. 6(b)).

In order to achieve higher computational efficiency of the model, the *ReLU* is employed in this paper, and the weights are initialized based on Gaussian distribution, with zero mean and small standard deviation.

3.3. Pooling layer

The function of the pooling layer immediately following the convolutional layer is to reduce the dimensions of the input feature maps while improving the robustness of the CNN model [52]. Pooling is proposed to remove unnecessary irrelevant details but retain the task-related information, which providing a form of translation invariance [53] as well as reduce the computational complexity and sensitivity to ambient influence [54].

Average-pooling (Fig. 7 (a)) and Max-pooling (Fig. 7 (b)) are two of the most common pooling methods across various tasks. Especially, it is reported that max-pooling is particularly well suited to the separation of features that are very sparse [55]. In this research, max-pooling is chosen for the pooling layers, as defined as follows

$$\mathbf{x}_j^{S_l} = f(\beta_j^{S_l} \text{down}(\mathbf{x}_j^{C_l}) + b_j^{S_l}) \quad (16)$$

where $\mathbf{x}_j^{S_l}$ denotes the j -th feature map in the l -th pooling layer; $\beta_j^{S_l}$ denotes the j -th scaling factor of the l -th pooling layer, $b_j^{S_l}$ denotes the j -th bias of the l -th pooling layer, and $\text{down}(\bullet)$ represents the subsampling function.

3.4. Back propagation based fine-tuning

The greedy layer-wise training process consists of forward pass and back-propagation pass. The forward pass takes the 2-D matrix as the inputs, go through the whole network, and calculate the errors between the model prediction and the objective. The back-propagation pass fine-tune the network using the error feedback. These two process alternate until the network converges. The back propagation process is of various arithmetic, in this paper, the algorithms of back-propagation pass for the three components can be separately summarized as follows.

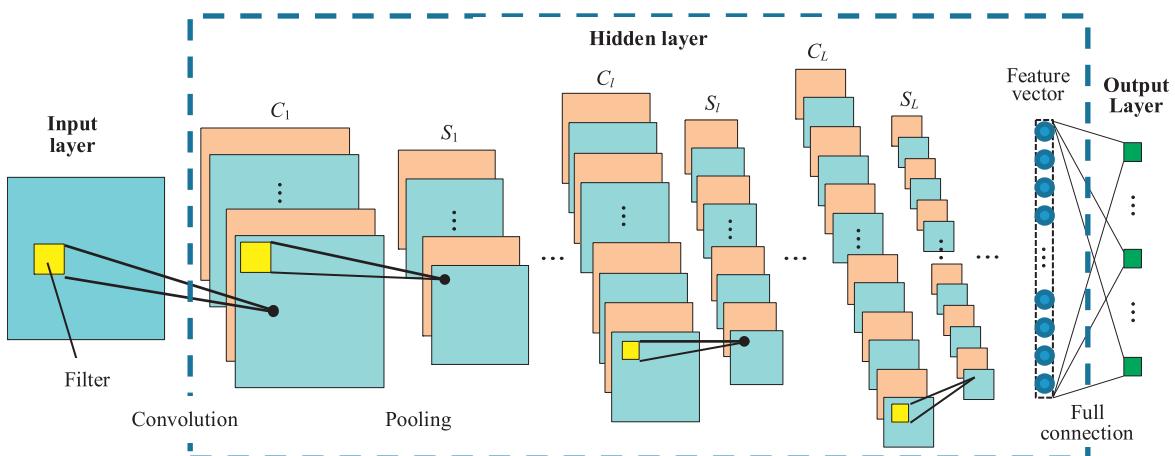


Fig. 5. The basic structure of the CNN model.

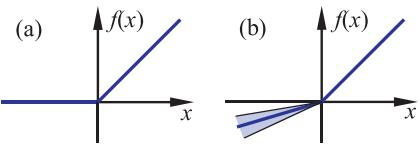


Fig. 6. The comparison of several activation functions.

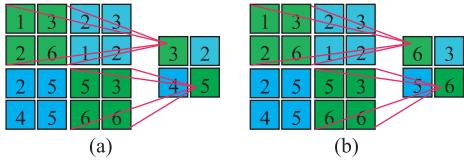


Fig. 7. Two common pooling methods. (a) average-pooling (b) max-pooling.

- Full-connection layer to pooling layer

$$\frac{\partial L}{\partial \mathbf{w}^l} = \mathbf{x}^{l-1} (\delta^l)^T \quad (17)$$

$$\frac{\partial L}{\partial \mathbf{b}^l} = \delta^l \quad (18)$$

L is the loss function, and the categorical_crossentropy loss function is employed in this research to evaluate the network and provide the error feedback. It is defined as follows for the last layer, where n denotes the number of samples used in training, and m denotes the number of classes. \hat{y}_{im} denotes the value of the m -th element of the m -dimensional predicted vector for the i -th sample, and y_{im} denotes the value of the m -th element of the m -dimensional objective vector for the i -th sample.

$$L = - \sum_{i=1}^n \hat{y}_{i1} \log y_{i1} + \hat{y}_{i2} \log y_{i2} + \dots + \hat{y}_{im} \log y_{im} \quad (19)$$

\mathbf{w}^l denotes the connection weight vector between the $(l-1)$ -th layer and the l -th layer; \mathbf{b}^l denotes the bias of the l -th layer; \mathbf{x}^{l-1} denotes the feature vector generated for the $(l-1)$ -th layer; δ^l denotes the error sensitivities of the l -th layer, and it can be calculated as follows:

$$\delta^l = (\mathbf{w}^{l+1})^T \delta^{l+1} \cdot f(\mathbf{x}^{l-1}) \quad (20)$$

For the last full-connection layer, δ^L can be calculated as $\delta^L = f(\mathbf{x}^{L-1}) \cdot (\hat{\mathbf{y}} - \mathbf{y})$, where $\hat{\mathbf{y}}$ denotes the predicted vector and \mathbf{y} denotes the objective vector.

- Pooling layer to convolutional layer

$$\frac{\partial L}{\partial \mathbf{w}_{j,i}^{C_l}} = \sum_{u,v} (\delta_j^{C_l})_{u,v} (p_j^{S_{l-1}})_{u,v} \quad (21)$$

$$\frac{\partial L}{\partial \mathbf{b}_j^{C_l}} = \sum_{u,v} (\delta_j^{C_l})_{u,v} \quad (22)$$

where $w_{j,i}^{C_l}$ denotes the i -th element of the j -th filter in the l -th convolution layer; $\mathbf{b}_j^{C_l}$ denotes j -th element of the bias of the l -th convolution layer; $p_j^{S_{l-1}}$ is the patch in the j -th input matrix of the l -th pooling layer in order to compute the element at (u, v) in the output of the l -th pooling layer; the sensitivities for l -th

convolution layer are calculated as follows

$$\delta_j^{C_l} = \beta_j^{S_l} (f'(\mathbf{x}_j^{C_l}) \cdot up(\delta_j^{S_l})) \quad (23)$$

where $up(\cdot)$ denotes an up-sampling process which conducts an inverse transformation to restore the dimension of the feature maps in the sub-sampling layer to the dimension before sub-sampling, $f'(\bullet)$ denotes the first derivative of the activation function $f(\bullet)$.

- Convolution layer to pooling layer

$$\frac{\partial L}{\partial \mathbf{b}_j^{S_l}} = \sum_{u,v} (\delta_j^{S_l})_{u,v} \quad (24)$$

$$\frac{\partial L}{\partial \mathbf{w}_j^{S_l}} = \sum_{u,v} (\delta_j^{S_l})_{u,v} \quad (25)$$

$\beta_j^{S_l}$, $down(\bullet)$ and $b_j^{S_l}$ denote the same things as in Eq. (16). The sensitivities for l -th pooling layer are calculated as follows:

$$\delta_j^{S_l} = \sum_{i \in M_j} \delta_i^{C_{l+1}} * \mathbf{w}_j^{C_l} \quad (26)$$

M_j and $\mathbf{w}_j^{C_l}$ denote the same thing as in Eq. (13).

4. The proposed intelligent method

In the proposed method, the constructed DWF is employed to process input vibration signals. It can completely decompose the original signal into multiple frequency bands of different scales and different center frequencies. CNN possesses powerful abilities to intelligently extract features from images and has achieved many results in the field of image classification. Therefore, this paper proposes a tool wear state identification method combining DWF and CNN, which is mainly made up of the following four steps, as shown in Fig. 8.

Step 1). Machine tool spindle vibration signal acquisition, including analogical filtering, digital sampling, anti-aliasing filtering, data segmentation, etc.

Step 2). Decomposition of the original signal via the proposed DWFs. DWFs are used to decompose the original signal into frequency bands of different bandwidths and different center frequencies. The characteristics of each frequency band are more obvious with the change of tool wear process. Considering that the contact frequency of the cutting edge with the workpiece is much lower than the sampling frequency of the vibration signal, the information related to the tool wear state is mainly contained in the low frequency band. Therefore, only the low frequency band of each layer is reconstructed, which effectively filters out the high frequency noise, and also reduces the input dimension of the CNN model, which is beneficial to model training.

Step 3). Two-dimensional signal matrix construction. The 1-D reconstructed sub-signals are spliced into a 2-D matrix, aiming at better convolution operations and retaining more information of the original signal. Hilbert envelope demodulation spectrum is very effective in exploiting the periodic impact components of one dimensional signals. Therefore, the Hilbert envelope demodulation spectra (HEDS) of reconstructed signals in each frequency band are also spliced into the 2-D signal matrix. The information of the sub-reconstruction signal used in the signal matrix is shown in Table 1 and the stacking method of reconstructed sub-signals is shown in Fig. 8.

Step 4). Tool wear state classification based on CNN model. The 2-D signal matrices are randomly divided into 3 groups, named as train

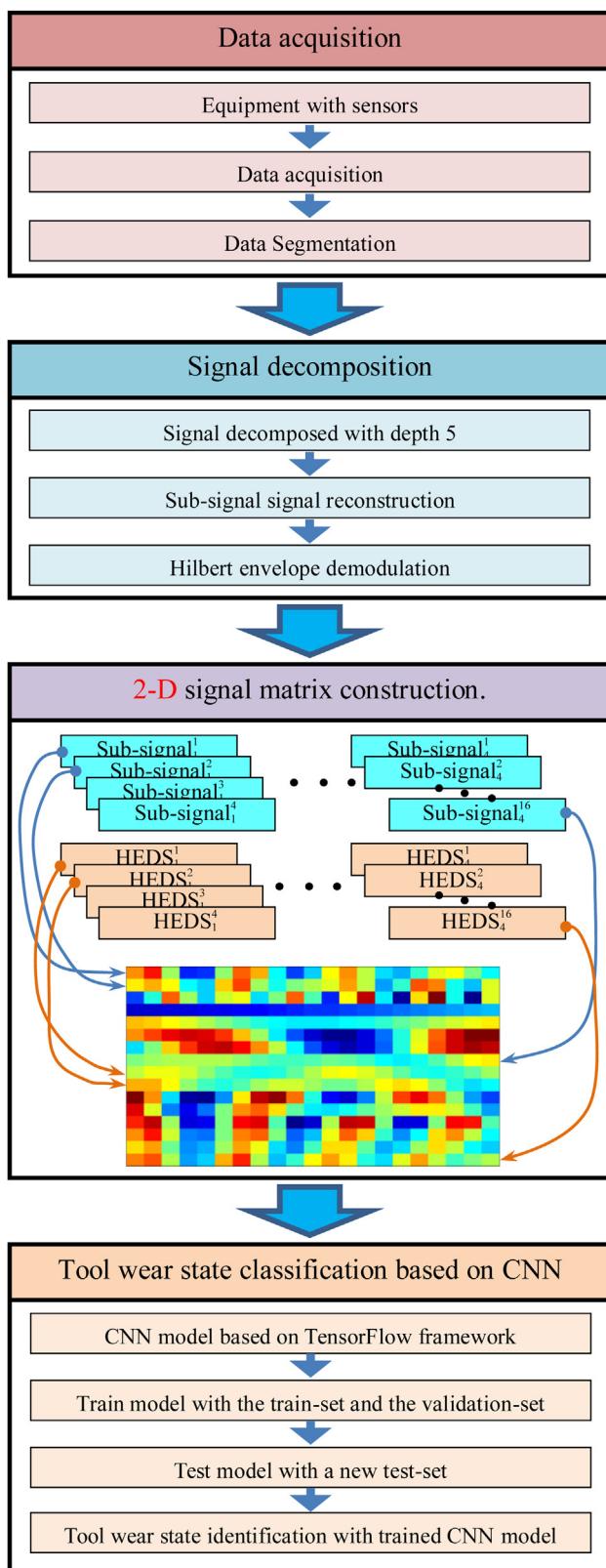


Fig. 8. Flow chart of the proposed method.

set, validation set and test set. Train the CNN network with train set and validation set. After the training process is completed, the test set that does not involve the training process is used for test.

Compared with traditional data processing, by reconstructing intermediate-scale sub-signals, DTCWT constructs a more detailed

sub-band reconstruction signal sequence when the decomposition depth is the same. The 2-D signal matrix retains more original information, and CNN can realize high-level feature representation autonomously, thereby reducing the reliance on the prior knowledge of operator.

5. Experimental investigation

5.1. experiment and data acquisition

In order to verify the effectiveness of the proposed method, a series of S45C steel plane milling experiments were carried out. Experiments are conducted on VMC650E /850E machining center (produced by Shenyang Machine Tool Company (Shenyang, China)), as shown in Fig.9.

The workpiece was an S45C steel brick as shown in Fig.10(a), and the dimension is 200 mm × 150 mm×200 mm, which is much larger the cutting parameters. As the experiment progresses, the dimension of the workpiece does not change significantly, the stiffness of the workpiece itself remains substantially constant, and the vibration characteristics are substantially unchanged. Thus all 25 cutting process can be completed on the same workpiece, reducing the complexity of the experimental operation.

Uncoated four teeth carbide end mill with a diameter of 20 mm is used, and the overhang is set to 60 mm while the overall length is 100 mm. Accelerometer was mounted on the spindle housing and workpiece to measure the real-time vibration signals, as shown in Fig. 10(a). The sampling frequency is 12,800 Hz. The NI PXIe-1078 chassis was employed to collect signals and transport them to the computer. Images of the cutting edges are acquired with an industrial camera after each cutting process is completed, as shown in Fig. 10(b). Each time the image is acquired, the tool is in the same machine coordinate position and the locating element is used to position the cutting edge to ensure that the position of the tool in the distance coordinate system remains constant each time the image is acquired. Experiments are conducted by straightly milling the brick. In order to better simulate the actual production process, the cutting process consist of sixteen working step with different machining parameters as listed in Table 2. From the start of the experiment to the end of the tool life, a total of 25 repeated cutting processes were performed, with the cumulative cutting time of the tool in each process being approximately 12 min.

5.2. Data segmentation and data set

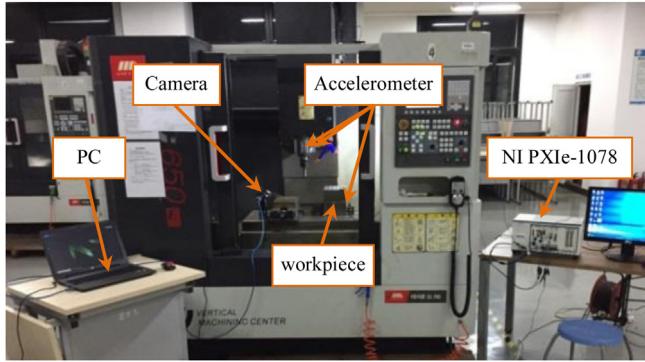
The original signal is shown in Fig. 11. The preliminary analysis shows that the data quality of the spindle Ydirection is the highest, which can clearly distinguish the cutting process and the idling process. Compared to the signals of the other two measuring points, the amplitude of the signal in the Y direction of the main axis is larger and the SNR is higher. In order to simplify the signal acquisition and analysis process, only the signal in the direction of the main axis is analyzed.

The experiment collected the original signals of 25 cutting processes, taking the last two steps of the fifth, tenth, fifteenth, twentieth, and twenty-fifth cutting processes as five wear states. Although the 15th step is down milling and the 16th step is up milling, the cutting parameters are the same. In order to improve the generalization ability of the neural network model, both steps are used to establish the data set. The original data of the two steps each contains about 5×10^5 data points, and respectively, 300 data segments with a length of 128000 are randomly intercepted to establish the data set. Then 3000 samples were randomly segmented into train set, validation set and test set, and the size ratio was 4:1:1.

Table 1

The composition of the 2-D vibration signal matrix.

Bandwidth (Hz)	center frequency (Hz)
1600	800, 1600, 2400, 3200
800	400, 800, 1200, 1600, 2000, 2400, 2800, 3200
400	200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, 2000, 2200, 2400, 2600, 2800, 3000, 3200,
200	100, 300, 500, 700, 900, 1100, 1300, 1500, 1700, 1900, 2100, 2300, 2500, 2700, 2900, 3100

**Fig. 9.** Milling experiment setup.

The Fourier transform is performed on the sample data, then the correlation between the normalized spectrums of the samples is compared. Suppose the signals to be compared are $x_1(t)$ and $x_2(t)$ and their spectrums are denoted $\hat{x}_1(t)$ as and $\hat{x}_2(t)$, the correlation coefficient between these two signals is expressed as

$$\text{Corr}(\hat{x}_1, \hat{x}_2) = \frac{\text{Cov}(\hat{x}_1, \hat{x}_2)}{(\text{Var}(\hat{x}_1))^{1/2} (\text{Var}(\hat{x}_2))^{1/2}} \quad (27)$$

where $\text{Var}[\cdot]$ is the variance of the spectrum and $\text{Cov}(\cdot, \cdot)$ is the covariance of the two spectra. The Min-Max scaling method is used for normalization in this paper, which is defined as:

$$X_{\text{norm}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (28)$$

where X_{norm} denotes the new data after normalization, and X is the original data. 100 samples were taken for statistical analysis of each tool wear state, and the results are shown in Fig.12.

The color and number of the blocks in the figure represent the average of the spectral correlation coefficients of the samples in different tool wear states, and the average of the spectral correlation coefficients between the samples in the same tool wear state. For example, the block in the upper left corner represent the average of 10,000 spectral correlation coefficients between the 100 samples in tool wear state 1 and the 100 samples in tool wear state 5, and the block in the lower left corner represents the average of 10,000 spectral correlation coefficients between the 100 samples in tool wear state 1. As shown in the figure, the spectra of the samples in the same tool wear state have a strong correlation. And it is stronger than that of the samples in different tool wear state. This proves the feasibility of the data segmentation method in this paper.

5.3. Two-dimensional signal matrix construction

The size of the 2-D signal matrix should be minimized while retaining sufficient information. Three types of matrices constructed in this paper: only the reconstructed sub-signal, only the Hilbert envelope spectrum of the sub-signal and composed of the first two. And the influence of the number of matrix columns on the accuracy of the CNN model is tested. Considering that the ratio of the train set, validation set, and test set is 4:1:1, we implemented a 6-fold cross-validations experiment. The overall data set was randomly divided into 6 equal parts, one was taken as a test set, and one of the remaining five was randomly taken as a validation set. In this way, a total of 6 trials were performed, and each sample was used for training and was identified as a test sample

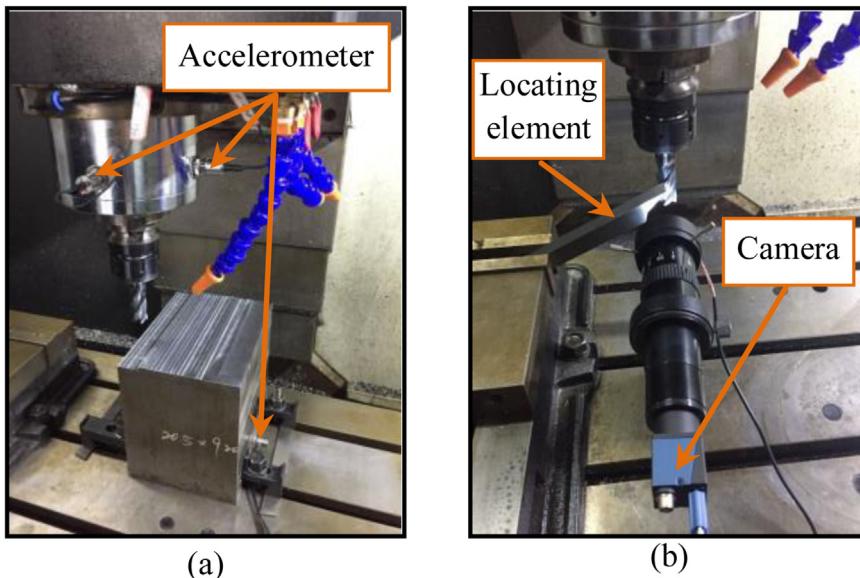
**Fig. 10.** Experiment setup: (a) Acceleration data acquisition and (b) Cutting edge image acquisition.

Table 2

Machining parameter.

working step	radial cutting depth (mm)	feed rate (mm/rev)	spindle speed (r/min)
1-4	1	0.08	800
5-8	2.5	0.06	800
9-12	1	0.08	1200
13-16	2.5	0.06	1200

throughout the cross-validation process. After the experiment is completed, the average test accuracy and the range difference of each type of 2-D signal matrix are shown in Fig. 13.

It can be seen from Fig. 13 that when the number of columns is not more than 250, the performance of the 2-D signal matrix of the three components is significantly improved, and the effect of the third signal matrix (Sub-Signal+HDDS) is significantly better than the former two; After more than 300, the performance of each signal matrix is no longer significantly improved, and even fluctuates. The authors believe that the performance of the third signal matrix is stable when the number of columns is 250, achieving high average test accuracy and low range. Although the average accuracy of the first and second signal matrices is very high when the number of columns exceeds 300, but the range is larger than the third signal matrix when the number of columns is 250. The corresponding validation accuracy curve of the neural network model using three different signal matrices as inputs is shown in Fig. 14. The three signal matrices are Sub-Signal with 300 columns, HEDS with 300 columns, and Sub-Signal+ HEDS with 250 columns, respectively. It can be seen that although the model can achieve high validation accuracy regardless of which signal matrix is used as the input, the convergence process is more stable when the third signal matrix is used.

In summary, considering the identification accuracy, stability and calculation speed of the algorithm, the signal matrix containing the sub-signal and its Hilbert envelope spectrum with 250 columns is a proper selection. The typical signal matrices are shown in Fig. 15.

5.4. Convolutional neural network structure selection and model training

The modeling capability of a CNN model is related to the depth and width of the network. As the depth deepens, the network can better approximate the objective function with increased nonlinearity and obtain better feature representation. However, it also increases the complexity of the network, which makes the network more difficult to optimize and tend to occur overfitting phenomenon. In this research, the employer CNN network is consists of two convolutional layers, two max-pooling layers and one full connection layer. The size of the filters in the first convolutional layer is set to 3×3 , and it is set to 2×2 in the second convolutional layer. And the size of pooling window is set to 2×2 . After the first convolutional layer, there is an additional layer to drop twenty percent of the nodes in order to prevent over-fitting. Finally, in the output layer, softmax activation is chosen for the classification to represent the categorical distribution. In this study, Adam optimizer is adopted to minimize the categorical cross entropy.

An experiment was conducted to optimize the width of the network. After 30 epochs, the test accuracy of different width models all reached 98%, but the test loss values are different, as shown in Fig. 16. As shown in Figs. 17 and 18, the number of nodes in the fully connected layer has a significant impact on the convergence of the model during training. Therefore, we propose to employ 9 filters in the first convolutional layer, 11 filters in the second convolutional layer, and 35 nodes in the full connection layer.

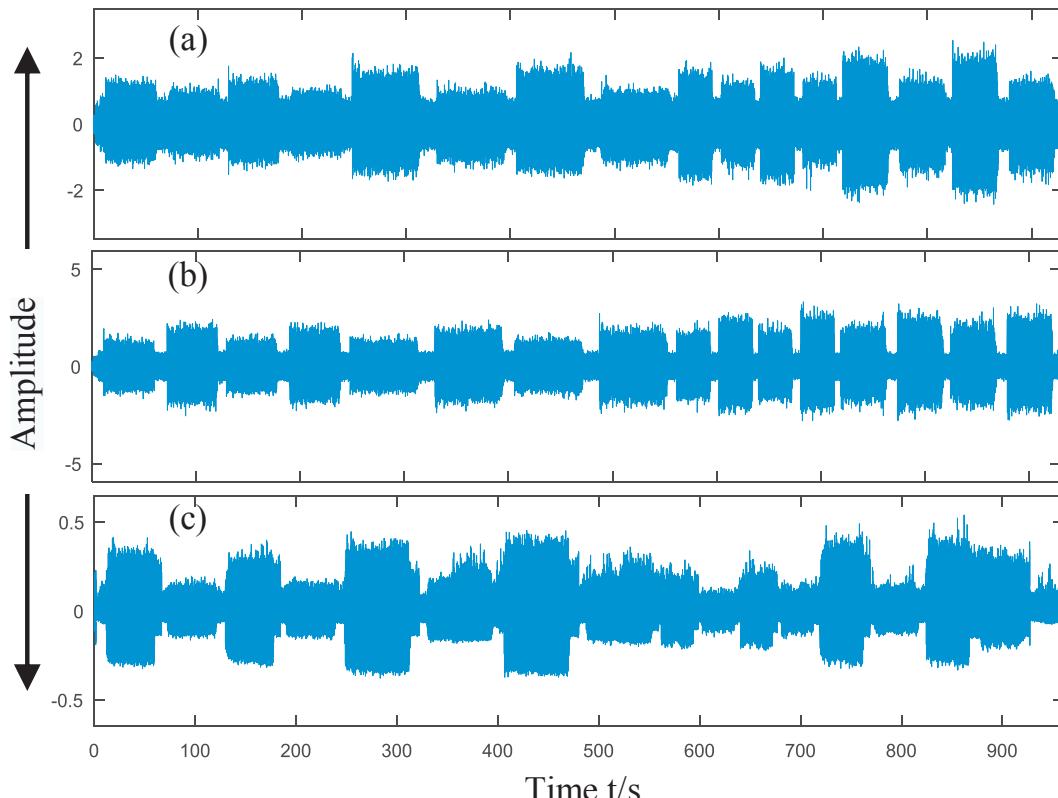


Fig. 11. Original signal: (a) X direction of the spindle; (b) Y direction of the spindle; (c) X direction of the workpiece.

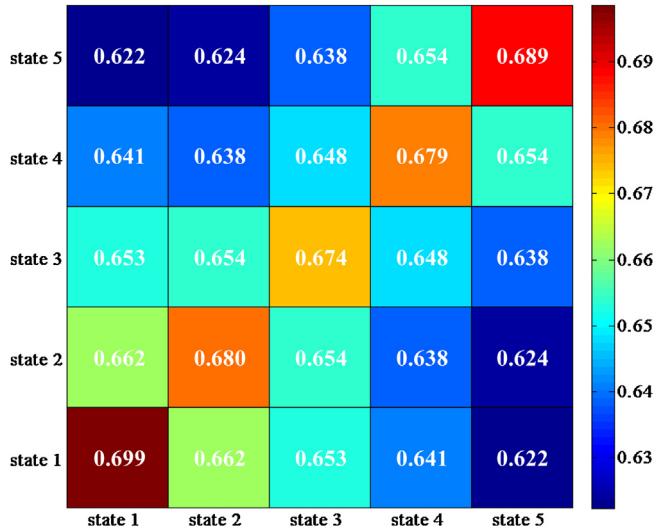


Fig. 12. Average spectral correlation coefficient between samples of different tool wear states.

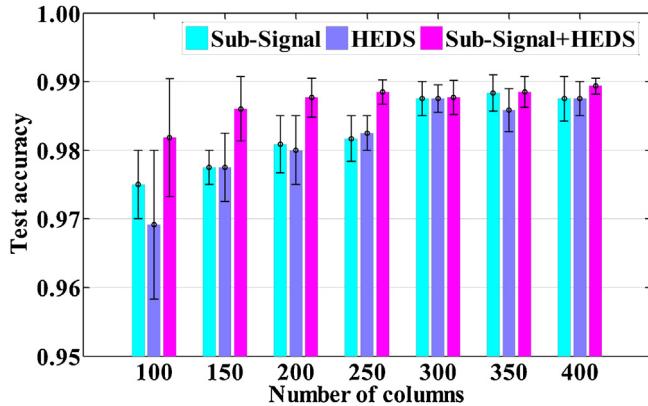


Fig. 13. Performance of different 2-D signal matrices.

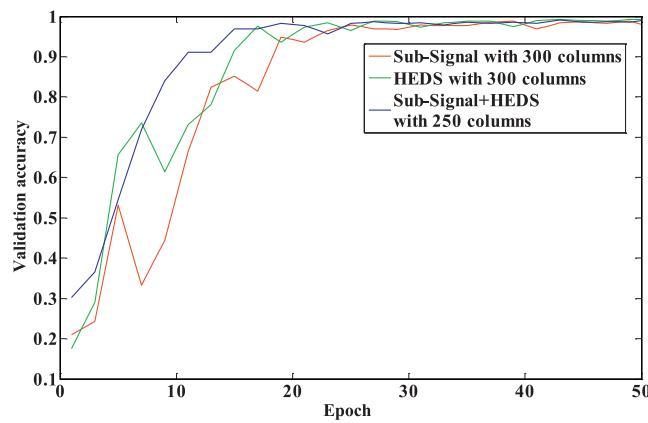


Fig. 14. Accuracy rate curves of different signal matrix.

6. Results and discussion

6.1. Experiment result

Confusion matrix is an effective visualization tool to estimate the performance of classification algorithm. Each column of the

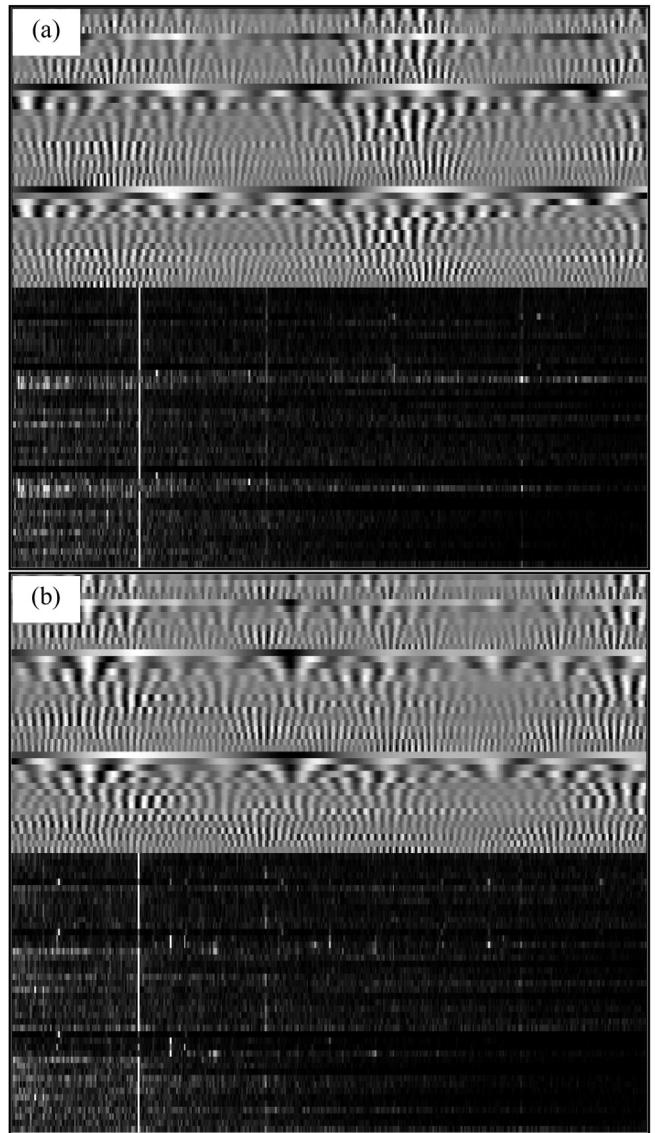


Fig. 15. Image of signal matrix: (a) tool wear state 1.

matrix represents the instances in a predicted tool wear state (Output), while each row represents the instances in an actual tool wear state (Target). A test set with 1000 samples is input into the trained CNN model, and the resulting confusion matrix is as shown in Fig. 19. It can be seen that the recognition accuracy of the state 5 that near failure reaches 99.5%, the recognition accuracy of the state 2 that in the stable wear phase also reaches 98%, and the overall recognition accuracy reaches 98.7%.

6.2. Modeling performance using different methods for 2-D signal matrix construction

There are several methods to convert a 1-D signal frame into a 2-D matrix. In order to evaluate the performances of the proposed method, other methods are used to build dataset, briefly described as follows.

A. STFT [56]. The short-time Fourier transform is performed on the original vibration signal of 12,800 data points (1 s). Using the Hamming window function, the window width is set to 512 and the window overlap length is set to 511. Then compress the size of the spectrogram to 256*256 by downsampling operations, as shown in Fig. 20.

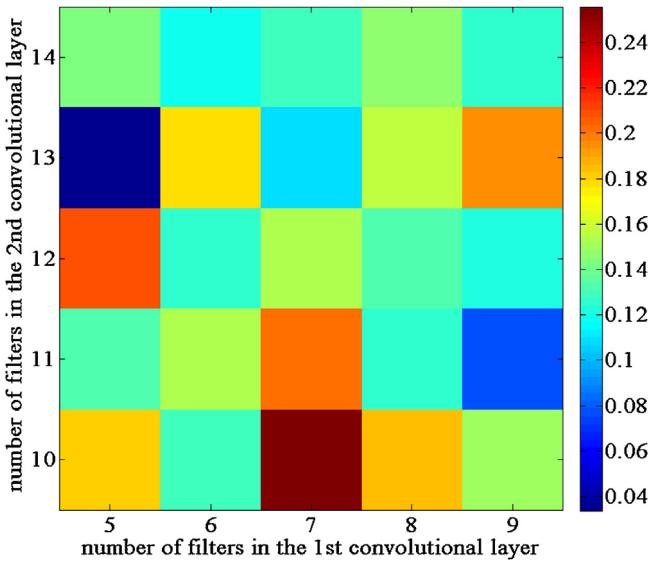


Fig. 16. Test loss value of CNN model with different number of filters in the two convolutional layers.

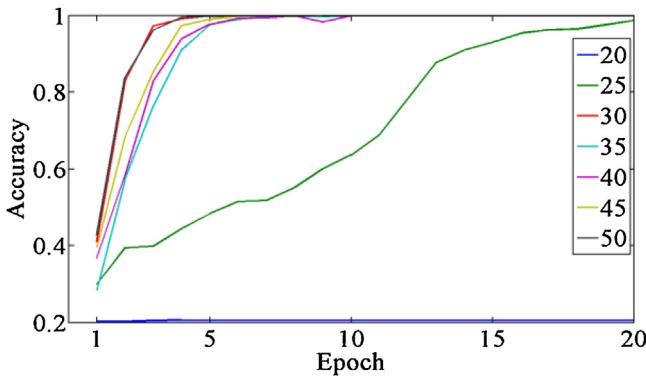


Fig. 17. Accuracy rate curves of CNN model with different number nodes in the full connection layer.

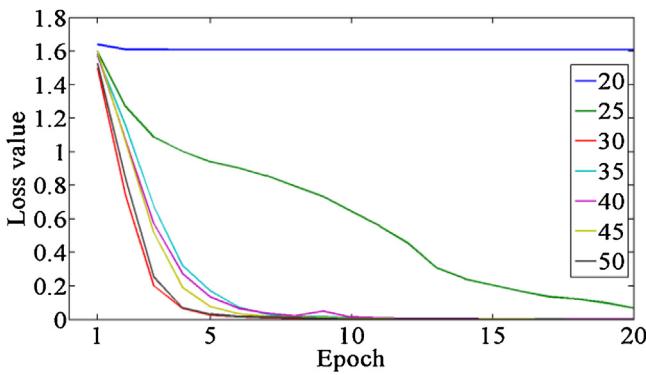


Fig. 18. Loss value curves of CNN model with different number nodes in the full connection layer.

B. CWT. The continuous wavelet transform is performed on the original vibration signal of 2560 data points (0.2 s). Use the db6 wavelet function and set the transform scale to 256. Then compress the size of the spectrogram to 256*256 by down-sampling operations, as shown in Fig. 21.

C. DCWPD + FOLD [57,58]. One-dimensional signal folding is a simple method to convert signals from one-dimensional form to 2-

	state1	state2	state3	state4	state5
Output	0.995	0.005	0.000	0.000	0.000
state1	0.005	0.985	0.010	0.000	0.000
state2	0.000	0.005	0.980	0.015	0.000
state3	0.000	0.000	0.005	0.990	0.005
state4	0.000	0.000	0.000	0.005	0.995
state5	0.000	0.000	0.000	0.005	0.995

Fig. 19. Confusion matrix.

D form. Combined with CNN neural network, it has achieved good results in the field of bearing fault identification. Suppose the size of the folded 2-D signal matrix is $m \times m$, the value of a pixel located in the j th column of the i th row is:

$$P(i,j) = L((i-1) \times m + j) \quad (29)$$

where $i = 1, 2, \dots, m; j = 1, 2, \dots, m$. $L(k)$ is the normalized amplitude of the original vibration signal. However, unlike the experimental data of bearing failure, the influence of tool wear on the spindle vibration is very weak, and the noise pollution is more serious. Experiments show that the original signal is directly folded into a 2-D matrix as the input of the CNN model, and the CNN model is difficult to converge. Therefore, this method is enhanced by the proposed DCWPD: firstly, the original signal is decomposed into 3 layers; then the low frequency band [0, 1600 Hz] is reconstructed; finally, the reconstructed sub-signal is folded to obtain 2-D signal matrix. Set the size of the matrix to 100×100 , as shown in the following Fig. 22.

D. DCWPD + PLOT [59]. Decompose the original vibration signal of 1280 data points (0.1 s) by the proposed DCWPD with a decomposition depth 3. Then plot the sub-signal of low frequency band [0, 1600 Hz] in a gray-scale map, and the size is set to 256*256, as shown in Fig. 23. The same original signal segment samples are processed by the above methods, and data sets are respectively established to train the 2-D convolutional neural network.

At the same time, the segment with length of 1920 (0.15 s) is extracted from the reconstructed sub-signal of [0, 1600 Hz] frequency band, and the 1-D form is used to establish the data set to train the 1-D convolutional neural network for comparative study.

In the corresponding references, the scholars optimized the structure and hyper-parameters of the CNN model for the signal matrix data set constructed by them, and finally obtained a high recognition accuracy. But the test accuracy of the model structure given in the reference is far less than that given in the publication when applied to the data set in this paper. Experiments also show that the optimized neural network containing double convolutional layers of the proposed tool wear state method is not suitable for these new data sets. Referring to the model structure given in the references, from simplifying the structure to further adding

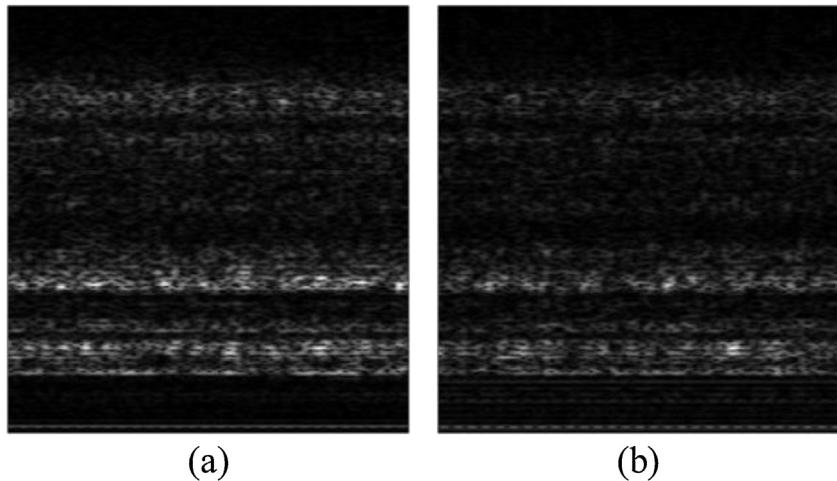


Fig. 20. STFT time-frequency spectra of samples: (a) tool wear state 1 and (b) tool wear state 5.

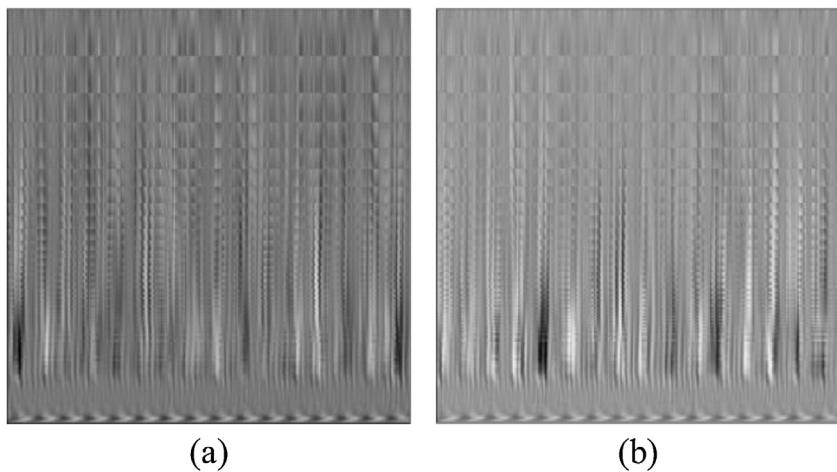


Fig. 21. CWT time-frequency spectra of samples: (a) tool wear state 1 and (b) tool wear state 5.

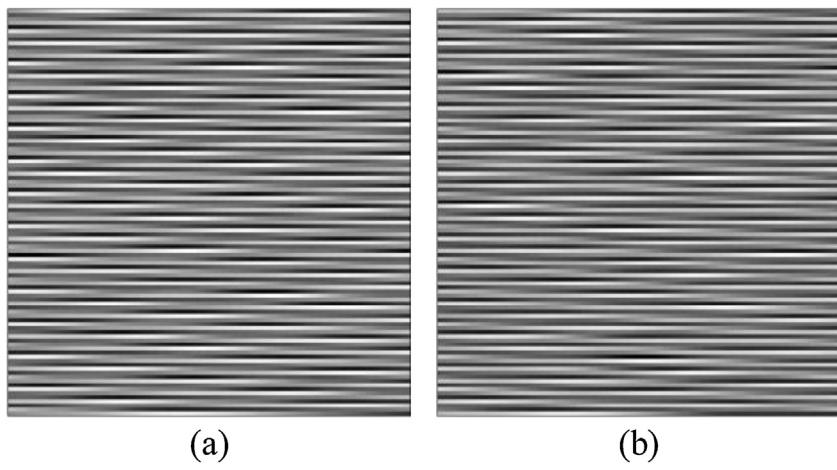


Fig. 22. Folded matrices of sample signals: (a) tool wear state 1 and (b) tool wear state 5.

hidden layers and neurons, experiments are carried out to find a suitable network structure, and those achieved better results were detailed in [Table 3](#), where FN denotes the number of filters and FS

denotes the size of the filter. Compared with other methods, the proposed method achieves better recognition accuracy and the structure of convolutional neural network is more concise.

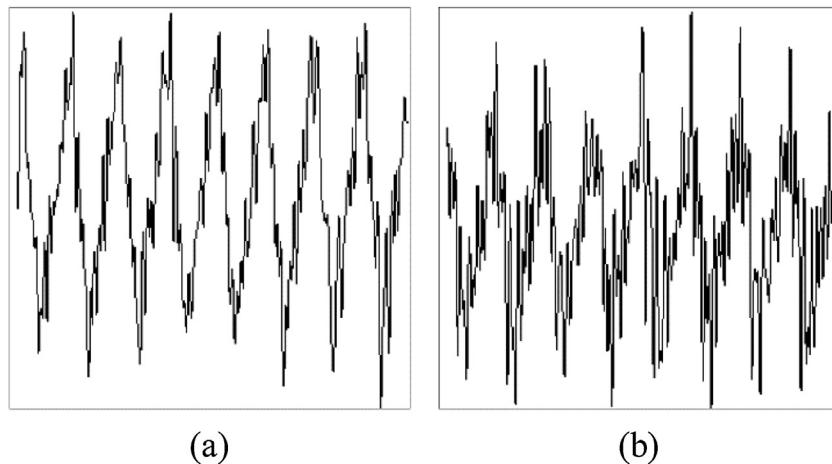


Fig. 23. Gray-scale maps: (a) tool wear state 1 and (b) tool wear state 5.

Table 3
Classification results for different diagnosis.

Method	CNN configuration						Test Accuracy	
	Input Size	C1	C2	C3	C4	Full Connected		
STFT + 2-D CNN	256 × 256	FN 8 FS 4 × 4	FN 13 FS 2 × 2	FN 13 FS 2 × 2	--	50	0.005	93.7%
CWT + 2-D CNN	256 × 256	FN 10 FS 4 × 4	FN 15 FS 2 × 2	FN 15 FS 2 × 2	--	50	0.01	94.3%
DCWPD + FOLD + 2-D CNN	100 × 100	FN 16 FS 5 × 5	FN 32 FS 3 × 3	FN 64 FS 3 × 3	--	64	0.0025	83.5%
DCWPD + PLOT + 2-D CNN	256 × 256	FN 10 FS 5 × 5	FN 15 FS 5 × 5	FN 15 FS 8 × 8	FN 20 FS 5 × 5	80	0.0075	93.4%
DCWPD + 1-D CNN	1920 × 1	FN 60 FS 5 × 1	FN 80 FS 5 × 1	FN 120 FS 5 × 1	FN 120 FS 4 × 1	64	0.005	93.5%
Proposed method	88 × 300	FN 9 FS 3 × 3	FN 11 FS 2 × 2	--	--	35	0.001	98.7%

7. Conclusions

In this paper, we propose an on-machine tool wear state recognition based on spindle vibration using a DTCWPT enhanced CNN, aiming at promoting recognition accuracy, generalization and computational efficiency on the premise of less expert knowledge requirements. After normalization, reconstructed sub-signal sequences of different scales and their Hilbert envelope demodulation spectra are sequentially stacked to form a 2-D signal matrix. The 2-D signal matrix is used to train and test the CNN model. The major findings of this work can be summarized as follows:

- (1) The DWFs and matrix stacking method proposed in this paper effectively filters out the high frequency noises within the original signals, and the features of the signal matrix are more sparse. The experimental results show that the proposed method is superior to the other four methods for converting 1-D temporal signal into 2-D signal matrix.
- (2) The 2-D signal matrix retains more information than a single reconstructed sub-signal. Compared with the deep one-dimensional convolutional neural network trained with a single sub-signal data set, the proposed method achieves higher test accuracy.
- (3) The effects of different configurations of neural network models were experimentally studied in this paper. Different sample formats require different network structures, and the model established by the method proposed in this paper is the simplest.

The proposed tool wear state recognition method for end milling process can also be extended to other machining process such as drilling and turning. The challenge is then how to investigate its applications to more complicated machining process such as milling complex surfaces on a five-axis machining center. Meanwhile, it is worthwhile to embed and integrate the proposed method into automated manufacturing environment in a completely unsupervised manner.

Acknowledgments

This research is supported financially by National Natural Science Foundation of China (No. 51605403), Natural Science Foundation of Guangdong Province, China (No. 2015A030310010), Natural Science Foundation of Fujian Province, China (No. 2016J01012).

References

- [1] Dan Li, J. Mathew, Tool wear and failure monitoring techniques for turning—a review[J], Int. J. Mach. Tools Manuf. 30 (4) (1990) 579–598.
- [2] S. Kurada, C. Bradley, A review of machine vision sensors for tool condition monitoring[J], Comput. Ind. 34 (1) (1997) 55–72.
- [3] Karandikar Jaydeep, Mcleay Tom, Turner Sam, et al., Tool wear monitoring using naïve Bayes classifiers[J], Int. J. Adv. Manuf. Technol. 77 (9–12) (2015) 1613–1626.
- [4] Najafi Babak, Hakim Hossein, A comparative study of non-parametric spectral estimators for application in machine vibration analysis[J], Mech. Syst. Signal Process. 6 (6) (1992) 551–574.
- [5] Fu Pan, A.D. Hope, G.A. King, Intelligent tool condition monitoring in milling operation[J], Milling Machines (1998).

- [6] S.C. Lin, R.J. Lin, Tool wear monitoring in face milling using force signals[J], Wear 198 (1-2) (1996) 136–142.
- [7] Drouillet Cyril, Karandikar Jaydeep, Nath Chandra, et al., Tool life predictions in milling using spindle power with the neural network technique[J], J. Manuf. Process. 22 (2016) 161–168.
- [8] Y. Sanchez, F.J. Trujillo, L. Sevilla, et al., Indirect Monitoring Method of Tool Wear Using the Analysis of Cutting Force During Dry Machining of Ti Alloys[J], 13(2017), pp. 623–630.
- [9] A.I. Azmi, Monitoring of tool wear using measured machining forces and neuro-fuzzy modelling approaches during machining of GFRP composites[J], Adv. Eng. Softw. 82 (2015) 53–64.
- [10] Nakai Mauricio Eiji, Aguiar Paulo Roberto, Guillard Hildo, et al., Evaluation of neural models applied to the estimation of tool wear in the grinding of advanced ceramics[J], Expert Syst. Appl. 42 (20) (2015) 7026–7035.
- [11] Niaki Farbod Akhavan, Michel Martin, Mears Laine, State of health monitoring in machining: extended Kalman filter for tool wear assessment in turning of IN718 hard-to-machine alloy[J], J. Manuf. Process. 24 (2) (2016) 361–369.
- [12] Niaki Farbod Akhavan, Ulutan Durul, Mears Laine, In-process tool flank wear estimation in machining gamma-prime strengthened alloys using Kalman filter[C], North American Manufacturing Research Conference, (2015).
- [13] Yan Ruqiang, Li Xiang, Chen Zhangwei, et al., Improving calibration accuracy of a vibration sensor through a closed loop measurement system[J], IEEE Instrum. Meas. Mag. 19 (1) (2016) 42–46.
- [14] Hsieh Wan Hao, Lu Ming Chyuan, Chiou Shean Juinn, Application of backpropagation neural network for spindle vibration-based tool wear monitoring in micro-milling[J], Int. J. Adv. Manuf. Technol. 61 (1-4) (2012) 53–61.
- [15] Ratava Juhu, Lohtander Mika, Varis Juha, Tool condition monitoring in interrupted cutting with acceleration sensors[J], Robot. Comput. Manuf. (2016) 47.
- [16] Chen Xuefeng Du Zhaohui, Zhang Han, et al., Weighted low-rank sparse model via nuclear norm minimization for bearing fault detection[J], J. Sound Vib. 400 (2017) 270–287.
- [17] Quintana Guillen, Ciurana Joaquim, Chatter in machining processes: a review [J], Int. J. Mach. Tools Manuf. 51 (5) (2011) 363–376.
- [18] He Shuilong, Liu Yikun, Chen Jinglong, et al., Wavelet transform based on inner product for fault diagnosis of rotating machinery[J], Mech. Syst. Signal Process. 70 (2016) 1–35.
- [19] Cui Lingli, Huang Jinfeng, Zhang Feibin, et al., HVS RMS localization formula and localization law: localization diagnosis of a ball bearing outer ring fault[J], Mech. Syst. Signal Process. 120 (2019) 608–629.
- [20] Sun Weifang, Chen Binqiang, Yao Bin, et al., Complex wavelet enhanced shape from shading transform for estimating surface roughness of milled mechanical components[J], J. Mech. Sci. Technol. 31 (2) (2017) 823–833.
- [21] Chen Binqiang, Zhang Zhiusuo, Zi. Yanyang, et al., Novel ensemble analytic discrete framelet expansion for machinery fault diagnosis[J], J. Mech. Eng. 50 (17) (2014) 77.
- [22] Wang Yanxue, Zhengjia He, Yanyang Zi, Enhancement of signal denoising and multiple fault signatures detecting in rotating machinery using dual-tree complex wavelet transform[J], Mech. Syst. Signal Process. 24 (1) (2010) 119–137.
- [23] R. Teti, K. Jemielniak, G. O'Donnell, et al., Advanced monitoring of machining operations[J], CIRP Ann. Manuf. Technol. 59 (2) (2010) 717–739.
- [24] Zhiqiang Ge, Song Zhihuan, Gao Furong, Review of recent research on database-based process monitoring[J], Ind. Eng. Chem. Res. 52 (10) (2013) 3543–3562.
- [25] A. Al-Habaibeh, N. Gindy, A new approach for systematic design of condition monitoring systems for milling processes[J], Journal of Materials Processing Tech 107 (1) (2000) 243–251.
- [26] Kirby Mr E. Daniel, Zhang Ms Zhe, C. Joseph, et al., Development of an accelerometer-based surface roughness prediction system in turning operations using multiple regression techniques[J], J. Ind. Technol. 20 (4) (2004).
- [27] L.B. Jack, A.K. Nandi, Genetic algorithms for feature selection in machine condition monitoring with vibration signals[J], Iee Proc. - Vis. Image Signal Process. 147 (3) (2000) 205–212.
- [28] Shi Dongfeng, N. Gindy Nabil, Tool wear predictive model based on least squares support vector machines[J], Mech. Syst. Signal Process. 21 (4) (2007) 1799–1814.
- [29] K. Patra, A.K. Jha, T. Szalay, et al., Artificial neural network based tool condition monitoring in micro mechanical peck drilling using thrust force signals[J], Precis. Eng. 48 (2017) 279–291.
- [30] Corne Raphael, Nath Chandra, Mansori Mohamed El, et al., Study of spindle power data with neural network for predicting real-time tool wear/breakage during inconel drilling[J], J. Manuf. Syst. 43 (2017) 287–295.
- [31] Sun Wenjun, Shao Siyu, Zhao Rui, et al., A sparse auto-encoder-based deep neural network approach for induction motor faults classification[J], Measurement 89 (ISFA) (2016) 171–178.
- [32] Yumak Onder, Ertunc H. Metin, Tool wear condition monitoring in drilling processes using fuzzy logic[C], International Conference on Neural Information Processing, (2006) , pp. 508–517.
- [33] Cuka Besmir, Kim Dong Won, Fuzzy logic based tool condition monitoring for end-milling[J], Robot. Comput. Manuf. (2017) 47.
- [34] Liang Shuang Yu Jinsong, Tang Diyin, et al., A weighted hidden Markov model approach for continuous-state tool wear monitoring and tool life prediction[J], Int. J. Adv. Manuf. Technol. 91 (1-4) (2016) 1–11.
- [35] N. Bhat Nagaraj, Dutta Samik, K. Pal Surjya, et al., Tool condition classification in turning process using hidden Markov model based on texture analysis of machined surface images[J], Measurement 90 (2016) 500–509.
- [36] Dutta Samik, K. Pal Surjya, Sen Ranjan, On-machine tool prediction of flank wear from machined surface images using texture analyses and support vector regression[J], Precis. Eng. 43 (2016) 34–42.
- [37] Viharos Zsolt János, Kis Krisztián Balázs, Support Vector Machine (SVM) based general model building algorithm for production control[J], Ifac Proc. Vol. 44 (1) (2011) 14103–14108.
- [38] Krizhevsky Alex, Sutskever Ilya, E. Hinton Geoffrey, ImageNet classification with deep convolutional neural networks[C], International Conference on Neural Information Processing Systems, (2012) , pp. 1097–1105.
- [39] Zhu Yingying, Zhang Chengquan, Zhou Duoyou, et al., Traffic sign detection and recognition using fully convolutional network guided proposals[J], Neurocomputing 214 (2016) 758–766.
- [40] Kim Yoon, Convolutional neural networks for sentence classification[J], Eprint Arxiv (2014) K.
- [41] Palaz Dimitri, Collobert Ronan, Magimai-Doss Mathew, Estimating phoneme class conditional probabilities from raw speech Signal using convolutional neural networks[C], INTERSPEECH, (2013) .
- [42] Sun Weifang, Yao Bin, Zeng Nianyin, et al., An intelligent gear fault diagnosis methodology using a complex wavelet enhanced convolutional neural network[J], Materials 10 (7) (2017) 790.
- [43] Fuan Wang, Hongkai Jiang, Haidong Shao, et al., An adaptive deep convolutional neural network for rolling bearing fault diagnosis[J], Meas. Sci. Technol. 28 (9) (2017).
- [44] Janssens Olivier, Slavkovikj Viktor, Vervisch Bram, et al., Convolutional neural network based fault detection for rotating machinery[J], J. Sound Vib. 377 (2016) 331–345.
- [45] Jing Luyang, Wang Taiyong, Zhao Ming, et al., An adaptive multi-sensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox[J], Sensors 17 (2) (2017) 414.
- [46] Wang Shuhui, Xiang Jiawei, Zhong Yongteng, et al., Convolutional neural network-based hidden markov models for rolling element bearing fault identification[J], Knowledge Based Syst. (2017).
- [47] Wang Yanxue, He Zhengjia, Zi Yanyang, Enhancement of signal denoising and multiple fault signatures detecting in rotating machinery using dual-tree complex wavelet transform [J], Mech. Syst. Signal Process. 24(1) (2010) 119–137.
- [48] W. Selesnick Ivan, G. Baraniuk Richard, N.C. Kingsbury, The dual-tree complex wavelet transform[J], Signal Processing Magazine IEEE 22 (6) (2005) 123–151.
- [49] Yan Ruqiang, X. Gao Robert, Chen Xuefeng, Wavelets for fault diagnosis of rotary machines: a review with applications[J], Signal Processing 96 (5) (2014) 1–15.
- [50] I.W. Selesnick, R.G. Baraniuk, N.C. Kingsbury, The dual-tree complex wavelet transform[J], IEEE Signal Process. Mag. 22 (6) (2005) 123–151.
- [51] Yang Weixin, Jin Lianwen, Tao Dacheng, et al., DropSample : a new training method to enhance deep convolutional neural networks for large-scale unconstrained handwritten Chinese character recognition[J], Pattern Recognit. 58 (4) (2016) 190–203.
- [52] S. Ji, M. Yang, K. Yu, 3D convolutional neural networks for human action recognition[J], IEEE Trans. Pattern Anal. Mach. Intell. 35 (1) (2013) 221–231.
- [53] Bengio Yoshua, Learning deep architectures for AI[J], Found. Trendsâ® Mach. Learn. 2 (1) (2009) 1–127.
- [54] Lecun Yann, Bengio Yoshua, Hinton Geoffrey, Deep learning[J], Nature 521 (7553) (2015) 436.
- [55] Boureau Y. Lan, Ponce Jean, Lecun Yann, A theoretical analysis of feature pooling in visual recognition[C], International Conference on Machine Learning, (2010) , pp. 111–118.
- [56] Marinescu Julian, A. Axinte Dragos, A critical analysis of effectiveness of acoustic emission signals to detect tool and workpiece malfunctions in milling operations[J], Int. J. Mach. Tools Manuf. 48 (10) (2008) 1148–1160.
- [57] Wen Long, Li Xinyu, Gao Liang, et al., A new convolutional neural network-based data-driven fault diagnosis method[J], Ieee Trans. Ind. Electron. 65 (7) (2018) 5990–5998.
- [58] Hoang Duy Tang, Kang Hee Jun, Rolling element bearing fault diagnosis using convolutional neural network and vibration image[J], Cogn. Syst. Res. (2018).
- [59] Fu Yang, Zhang Yun, Gao Yuan, et al., Machining vibration states monitoring based on image representation using convolutional neural networks[J], Eng. Appl. Artif. Intell. 65 (2017) 240–251.