

Research Article

Milling Tool Wear State Recognition by Vibration Signal Using a Stacked Generalization Ensemble Model

Yang Hui ^{1,2,3} Xuesong Mei,^{1,2,3} Gedong Jiang ^{1,2,3} Tao Tao,^{1,2,3} Changyu Pei,³ and Ziwei Ma³

¹State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an 710054, China

²Shaanxi Key Laboratory of Intelligent Robots, Xi'an Jiaotong University, Xi'an 710049, China

³School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, China

Correspondence should be addressed to Gedong Jiang; gdjiang@mail.xjtu.edu.cn

Received 26 July 2019; Accepted 23 September 2019; Published 3 November 2019

Guest Editor: Franco Concli

Copyright © 2019 Yang Hui 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.

Milling tool wear state recognition plays an important role in controlling the quality of milled parts and reducing machine tool downtime. However, the characteristics of milling process limit the accuracy and stability of tool condition monitoring (TCM) employing vibration signals. To improve this problem, this paper explores the use of vibration signals as sensing approach for recognizing tool wear states during milling operation by using the stacked generalization (SG) ensemble model. In this study, vibration signals collected during the milling process are analyzed through the time domain, frequency domain, and time-frequency domain to extract signal features. The support vector machine recursive feature elimination (SVM-RFE) algorithm is used to select the main features which are most relevant to tool wear states. The SG ensemble model based on SVM, decision tree (DT), naive Bayes (NB), and SG ensemble strategy is constructed to recognize tool wear states. The proposed method is experimental verified, and the results show that the recognition accuracy of the established SG ensemble model is 98.74% and the overall G-mean and AUC evaluation value of the model is 0.98 and 0.98, respectively. In addition, compared with other ensemble models and single models, the SG ensemble model based on vibration signals has better recognition accuracy and stability than other models.

1. Introduction

Milling is one of the most commonly used processes in today's industry and mechanical machining workshop for machining parts to precise sizes and shapes. As the direct executor, the milling tool that withstands high temperature and mechanical shock during the milling process will be worn gradually, which directly affect the surface quality of the workpiece and increase the rejection rate and production cost [1, 2], especially in the case of high-precision material cutting. To reduce this effect, the tool condition monitoring (TCM) system has been widely used and achieved remarkable results [3]. The data show that an effective and reliable TCM system can reduce downtime by 10–40% and improve workpiece quality [4, 5]. As an important part of tool monitoring system, tool wear state recognition mainly

includes direct and indirect methods. Compared with the direct method for directly monitoring the tool wear states by using optical equipment and machine vision technology, the indirect method recognizes the tool wear states mainly by establishing a classifier to depict the relationship between the sensor signal feature vector and the tool wear states. Therefore, the indirect method is of lower cost and more suitable for practical milling process; however, it requires high precision and stability of the classifier.

With the rapid development of machine learning technology, a large number of machine-learning modeling methods have been used to construct the classifier for tool wear state recognition. Based on computer vision method, García-Ordás et al. [6] adopted support-vector machine (SVM) classifier to monitor tool wear in edge profile milling operations. By extracting features from audible sound

signals in frequency domain, Kothuru et al. [7] built a multiclass support-vector machine (MSVM) to detect end milling tool wear under various cutting conditions. By categorizing tool wear into four states based on tool wear mechanism, tool wear rate, and tool life, Hu et al. [8] used v-SVM model to realize tool wear monitoring. Cao et al. [9] presented a new tool wear state recognition method by combining derived wavelet frames (DWFs) with convolutional neural network (CNN), and the effectiveness of the proposed model was demonstrated by experimental results from end milling. Considering the high accuracy, high computing speed, and excellent performance of deep learning algorithm, Chen et al. [10] applied a deep belief network (DBN) to predict the flank wear of a cutting tool in milling. To describe the time-variant transition probability of tool wear states and the state duration dependency, Zhu and Liu [11] proposed a hidden semi-Markov model (HSMM) for online tool wear estimation. By extracting features from the force signal in time domain, Kong et al. [12] proposed a Gaussian mixture-hidden Markov model (GMHMM) for tool wear estimation. Zhou and Xue [13] reviewed the state-of-the-art methods of TCM in milling processes, and the monitoring models such as artificial neural network (ANN), HMM, and SVM for the categorization of cutting tool states were discussed. By extracting the time and frequency domain features from the force signal, Karandikar et al. [4] built a naïve Bayes classifier to classify the tool wear states. Kong et al. [14] presented a Gaussian process regression (GPR) method to construct tool wear predictive model, and the experimental results show that the GPR model performed better than ANN and SVM.

These works [4, 6–14] solved the problem for recognizing tool wear states in the milling process. However, the models established in these studies are all based on single classifier. Moreover, no single classifier can claim to be superior to any other. Thus, the ensemble learning which significantly improves the generalization ability of a single classifier and gives better results than a single classifier [15] can be considered a viable alternative for obtaining the classifier for tool wear state recognition. The idea of ensemble learning method is to build a classifier predictive model by integrating multiple single classifiers [16]. Yu [17] proposed a discrete particle swarm optimization algorithm-based selective ANN ensemble (DPSOEN) approach to predict tool wear, and this method had better generalization performance than single ANN. Zheng and Liu [18] developed an ensemble learning model based on radial basis function (RBF), multilayer perceptron (MLP) neural network model, least mean square (LMS), and k^* model to predict wear loss. Yang et al. [19] proposed an effective coevolutionary particle swarm optimization-based selective neural network ensemble (E-CPSOEN) which enables tool wear prediction model, and the experimental results indicated that the E-CPSOEN was considerably better than these single ANN-based approaches. Yang et al. [20] established an integrated prediction model by using the trajectory similarity-based prediction (TSBP) and the differential evolution SVR (DE-SVR) algorithm to predict milling tool wear and life, and the results demonstrated that

this integrated prediction model was better than other four single algorithms (TSPB, SVR, PSO-SVM, and HMM). Javed et al. [21] presented an ensemble of summation wavelet extreme learning machine (SW-ELME) models with incremental learning scheme to predict tool wear, and this model can be used for estimating tool life span and giving confidence for decision-making.

These works [16–21] demonstrated the high accuracy of ensemble predictions in tool wear prediction. However, the ensemble strategies used in these studies are all based on bagging [22] or AdaBoost [23]. Unlike bagging and AdaBoost, stacked generalization (SG) [24] combines the yield which is produced by various base learners in the first level, and then, by utilizing a metalearner, it tries to combine the outcomes from these base learners in an ideal method to augment the generalization ability. It is well known that SG has been applied in the manufacturing areas, e.g., scheduling of flexible manufacturing systems [25], defect detection and classification in semiconductor units [26], and engine RUL prediction [27]. However, there are very few examples of SG applied to recognize tool wear states. Wang et al. [28] proposed a heterogeneous ensemble learning model based on SG ensemble strategy to realize tool wear state recognition. They used force sensor to depict the dynamic characteristics of the tool wear process. Compared with force sensor, vibration sensor is inexpensive and easy to install, so it is more suitable for practical application. However, the characteristics of milling processes limit the accuracy and stability of TCM employing vibration signals [13]. With the aim to improve this situation, a SG ensemble model based on SVM, decision tree (DT), naive Bayesian (NB), and SG ensemble strategy is investigated to realize milling tool wear state recognition. In the modeling process, time domain, frequency domain, and time-frequency domain features based on vibration signals are extracted, and main features are selected based on support-vector machine recursive feature elimination (SVM-REF) algorithm.

The rest of the paper is organized as follows. The modeling steps of SG ensemble model are introduced in Section 2. Section 3 presents the milling tool wear state recognition method based on the SG ensemble mode. In Section 4, the experimental setup is described. In Section 5, the results and discussions are given. Finally, the conclusions of this study are laid out in Section 6.

2. The Proposed SG Ensemble Model

In order to construct the SG ensemble model, it is necessary to guarantee that the individual classifiers are both accurate and diverse. In this paper, the SG ensemble model based on SVM, DT, and NB base classifiers is proposed. On the one hand, those base classifiers belong to the top 10 data mining algorithms identified by the IEEE [29], and they show good performances in classification accuracy and time for tool wear detection [4, 6, 30, 31]. On the other hand, SVM is a relatively new supervised learning method based on statistical learning theory and the structural risk minimization principle [32]. DT is a hierarchical model composed of decision rules that recursively split independent variables

into homogeneous zones [33]. NB is a classification system based on Bayes' theorem that assumes that all the attributes are fully independent given the output class, called the conditional independence assumption [34], so they are diverse in principle.

2.1. Establish the SG Ensemble Model. For SG ensemble model, SVM, DT, and NB are selected as the primary learner and SVM is used as the metalearner to reflect the relationship between the outputs of primary learner and tool wear states. In addition, the five-fold cross-validation method is used for model training. As shown in Figure 1, the steps to establish SG ensemble model are as follows:

Step 1. Divide sample dataset D into training dataset D_{training} and testing dataset D_{testing} . Then, randomly divide the training dataset D_{training} into subset $D_1, D_2, \dots, D_i (i = 5)$, and take D_i as subtesting dataset and the remaining subset $D^{(-i)} = D_{\text{training}} - D_i$ as subtraining dataset.

Step 2. Use the primary learner SVM to train the subtraining dataset $D^{(-i)}$, and output feature a_i by predicting the subtesting dataset D_i . In addition, the trained SVM is used to test the testing dataset D_{testing} and output feature b_i .

Step 3. Repeat Step 2 until the output feature sets $A_j = \{a_1, a_2, \dots, a_5\}$ and $B_j = \{b_1, b_2, \dots, b_5\}$ are obtained. Then, the features in feature set A_j are merged and formed into a column of feature set C_1 , and the features in feature set B_j are calculated by average value, and a column of feature set E_1 is obtained.

Step 4. By executing Step 1–3 for the primary learner DT and NB, the feature sets C_2 and C_3 are generated by training dataset D_{training} and the feature sets E_2 and E_3 are generated by testing dataset D_{testing} .

Step 5. Merge the feature sets C_1, C_2 , and C_3 with the real class label y_i to generate the new training dataset $M = \{(C_1, C_2, C_3), y\}$, and merge the feature sets C_1 and C_2 with C_3 to generate the new testing dataset $N = \{E_1, E_2, E_3\}$. Then, the metalearner SVM is trained by using the new training dataset M as input, and the final classification results are generated by inputting the new testing dataset N into the metalearner SVM.

2.2. Description of Various Classifiers and SG Ensemble Method

2.2.1. SVM. SVM is originally derived from the processing of the binary classification problem of linearly separable data. The mechanism of SVM is to find an optimal classification hyperplane that satisfies the classification requirements so that the hyperplane can maximize the blank area on its both sides while ensuring classification accuracy [32]. In our current problem, however, three wear classes are involved, for which, requires an application of multiclassification strategy. Thus, the one-versus-one method is adopted in this paper.

In this approach, if p is the number of classes, there are $A = p(p-1)/2$ binary SVM constructed, and each SVM is trained on data for two classes. For the training data from the j^{th} and k^{th} classes, the following binary classification problem is solved:

$$\begin{aligned} \text{Minimize } \phi(w) &= \frac{1}{2} \|w^{jk}\|^2 + C_j \sum_i^{y_i=j} \xi_i^{jk} + C_k \sum_i^{y_i=k} \xi_i^{jk}, \\ \text{Subject to } (w^{jk})^T \varphi(x_i) + b^{jk} &\geq 1 - \xi_i^{jk}, \quad \text{if } y_i = j, \\ (w^{jk})^T \varphi(x_i) + b^{jk} &\geq -1 + \xi_i^{jk}, \quad \text{if } y_i = k, \\ \xi_i^{jk} &\geq 0, \end{aligned} \quad (1)$$

where $\varphi(x_i)$ is a mapping function, (x_i, y_i) is the training data, w and b are the weighting factors, ξ_i^{jk} is the slack variable, and C is the penalty parameter.

The one-versus-one method is usually implemented using a ‘‘Max-Wins’’ voting strategy. If $\text{sign}((w^{jk})^T \varphi(x_i) + b^{jk})$ say x is in the j^{th} class, then vote for this class is added by one. Otherwise, the k^{th} class is increased by one. Then, x is designated to be in a class with the maximum number of votes.

Among the many common kernel functions that satisfy the condition, the radial basis function (RBF) kernel is a reasonable first choice in the practical application of SVM, and it is defined as follows:

$$K(x_i, x) = \exp \left\{ -\frac{|x - x_i|^2}{2\sigma^2} \right\}, \quad (2)$$

where $1/\sigma^2$ is recorded as γ .

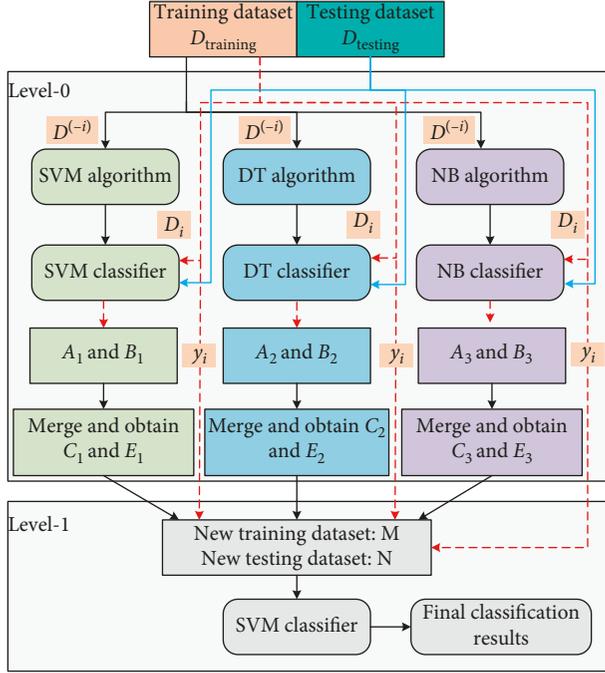
2.2.2. DT. DT can be defined as a map of the reasoning process. It describes a data set as a tree-like structure [33]. DT is typically built by recursive algorithm. At each step, data set is divided into different subsets by split attribute, which is the data set to be split in next step. The difference between different decision tree classification algorithms is the split attribute selection method. In this paper, information gain is chosen as the split attribute selection method.

The information gain of sample S for attribute x_i can be formulated as follows:

$$\text{Gain}(S, x_i) = \text{Entropy}(S, x_i) - \sum_{v \in V_i} \frac{|S_i(v)|}{S} \text{Entropy}(S, x_i), \quad (3)$$

where $\text{Entropy}(S, x_i)$ is the information gain between sample S and attribute x_i and $S_i(v)$ is the subset of value v of attribute x_i .

2.2.3. NB. NB is the most known and used classification method [34]. It is called naive as it assumes that all variables contribute towards classification and are mutually correlated. This assumption is called class-conditional independence.



$D^{(-i)}$: subtraining dataset D_i : subtesting dataset
 y_i : real class of D_i i : number of folds in the cross validation
 A_j : output feature set of D_i B_j : output feature set of $D_{testing}$
 C_j : merging A_j E_j : mean value of B_j
 $i = 1, 2, 3, 4, 5$ j : number of primary classifier, $j = 1, 2, 3$

FIGURE 1: The description of the SG ensemble model.

In addition, the NB method is not only easy to implement on various kinds of datasets but also it is quite efficient.

The basic idea of NB theorem is formulated as follows:

$$P(C_i | X) = \frac{P(X | C_i) \cdot P(C_i)}{P(X)}, \quad (4)$$

where $P(C_i | X)$ is the posterior probability of target class, $P(C_i)$ is called the prior probability of class, $P(X | C_i)$ is the likelihood which is the probability of predictor of given class, and $P(X)$ is the prior probability of predictor of class.

As $P(X)$ is the constant for all classes, only the product $P(X | C_i) \cdot P(C_i)$ needs to be maximized. The prior probabilities of the class are calculated as follows:

$$P(C_i) = \frac{\text{number of training samples of class } C_i}{m \text{ (} m \text{ is the total number of training samples)}}. \quad (5)$$

Using the conditional independence assumption between attributes, we can express

$$P(C_i | X) = \prod_{t=1}^n P(X_t | C_i), \quad (6)$$

where X_t are values for attributes in the sample X .

2.2.4. SG Ensemble Method. SG (also called stacking) is introduced by Wolpert in 1992 [22] as a well-known ensemble approach for combining multiple classifiers. It

attempts to give an accurate prediction even if the output of a certain base classifier is incorrect. So, it widely used in many fields. In addition, the method consists of training a set of learners (called the first-level learners) with the original training data. The outputs of the first-level learners are then used to train a second-level learner called the metalearner.

3. Milling Tool Wear State Recognition Based on SG Ensemble Model

Based on the proposed SG ensemble model, a milling tool wear state recognition method is constructed, whose structure is shown in Figure 2. Firstly, dynamic signals from vibration sensors are collected to depict the characteristic of the milling process. Secondly, due to the low signal-to-noise ratio (SNR) of sensing measurement, it is difficult to model the relationship between vibration signal and tool wear states in the milling process. To solve this problem, different features from time domain, frequency domain, and time-frequency domain are extracted, and the detailed features are listed in Table 1. Thirdly, not all features are essential since many of them are redundant or even irrelevant, which may reduce the performance of the algorithm. So, SVM-RFE algorithm [35] is used to select the main features which are most relevant to tool wear state from the high-dimension feature vectors. The description of SVM-RFE algorithm is shown in Algorithm 1. Fourthly, the SG ensemble model which is described in Section 2 is established to recognize milling tool wear states, and the performance of the model is evaluated. Finally, we used the milling tool wear state recognition model to judge the final tool wear state category.

4. Experiment

4.1. Experimental Setup. Experimental setup of milling tool wear monitoring is shown in Figures 3(a) and 3(b). A three-axis vertical machining center (VDL-1000, which is produced by Dalian Machine Tool Company) with a maximum speed of 8000 rpm is used for the milling tests. The cutting tool used for the experiment is a four-edge high-speed steel milling tool (diameter: 10 mm), and the workpiece material used for the current investigations is 45 heat-treatable steel whose main properties are shown in Table 2. The vibration signals are collected by a Kistler 8763 three-axis accelerometer (bandwidth: ± 75 kHz) mounted on the spindle, and near the cutting tool, the three directions of accelerometer are parallel to the three directions of machine tool. The NI-9234 is employed to collect signals and transport them to the computer. The sampling frequency is 4800 Hz. The measurement of milling tool flank wear uses a professional handheld digital microscope (Dino-Lite: AM3113T), polarization and high magnification capabilities up to 200 \times .

4.2. Design of Experiments. In the experiment, down milling with air-cooled is adopted during the whole process. The milling length is 100 mm for each cut, and milling direction is along the Y-axis of the machine tool. The other milling parameters are listed in Table 3. The measurement interval

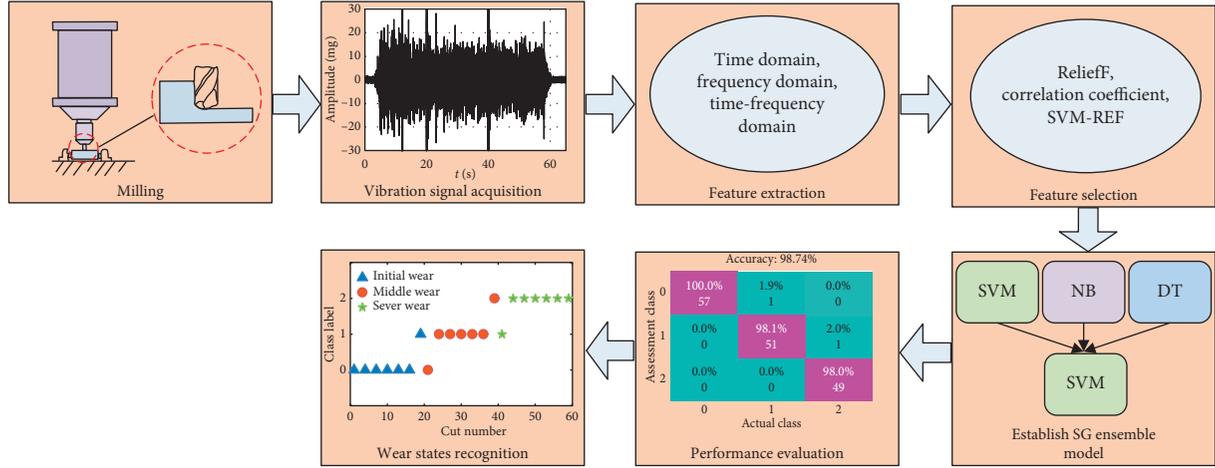


FIGURE 2: The structure of milling tool wear state recognition method.

TABLE 1: List of the extracted features.

Domain	Features	Expression
Time domain	Mean	$\mu_x = (1/N)\sum_{i=1}^N x_i$
	Peak	$x_p = \max(x_i)$
	Root mean square (RMS)	$x_{RMS} = \sqrt{(1/N)\sum_{i=1}^N x_i^2}$
	Variance	$x_{var} = \sum_{i=1}^N (x_i - \mu_x)^2 / (N - 1)$
	Kurtosis	$x_{Kur} = (1/N)\sum_{i=1}^N ((x_i - \mu_x)/\sigma)^4$
	Skewness	$x_{ske} = N \sum (x_i - \mu_x)^3 / (N - 1)(N - 2)\sigma^3$
Frequency domain	Amplitude of tooth passing frequency (TPF)	$x_A = iTPF$
Time-frequency domain	Energy percentage of IMF	$T_i = E_i / \sum_{i=1}^N E_i$

```

Input:  $X = [x_1, x_2, \dots, x_m]^T$  % Features matrix
        $Y = [y_1, y_2, \dots, y_m]^T$  % Class label
Out:  $r$  % Ranking coefficient
(1) begin
(2)    $s = [1, 2, \dots, k]$  % Current feature subset index
(3)    $r = []$  % Feature ranking
(4)   while ( $s \neq \emptyset$ )
(5)      $X_0 = X(:, s)$  % Obtain the current training sample
(6)      $\alpha = \text{SVM} - \text{train}(X, Y)$  % Train classifier
(7)      $c_i = \|w_i\|^2$  % Computational ranking criteria
(8)      $f = \arg \min(c)$  % Search for the lowest ranking feature attributes
(9)      $r = [s(f), r]$  % Update ranking feature index
(10)     $s = s(1 : f - 1, 1 : \text{length}(s))$  % Delete the lowest ranking feature attribute
(11)  end
(14) end

```

ALGORITHM 1: Description of the SVM-RFE algorithm.

for initial wear and sever wear of the milling tool is cutting 15 times, and for middle wear of the milling tool it is cutting 30 times. Because of manual measurement, each measurement is conducted three times to avoid any possible mistake during the procedure.

4.3. *Classification of Milling Wear States.* Generally, measurement criteria of tool flank wear include flank wear

bandwidth and flank wear area [36, 37]. Since the spiral milling tool has a certain spiral angle and microscope belongs to two-dimensional measurement, the change of the milling cutter angle would affect the measurement accuracy of tool wear dimension. So, the measurement of the flank wear bandwidth of the milling tool has higher accuracy and reliability. In this paper, the tool flank wear is described in the maximum flank wear bandwidth. As shown in Figure 4, the flank wear bandwidth of a new milling tool is $DL0 = 0.500$ mm

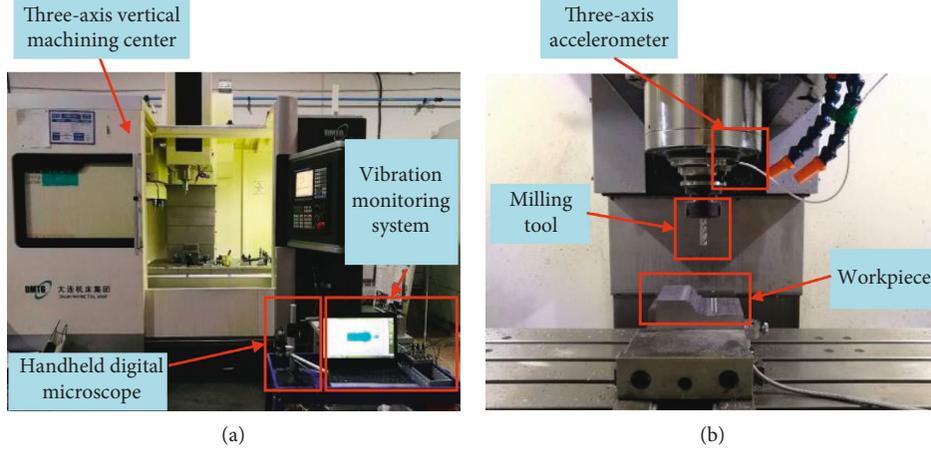


FIGURE 3: Scheme of milling tool wear experiment: (a) the whole tool wear condition monitoring system and (b) the relative position between the milling tool, the workpiece, and the three-axis accelerometer.

TABLE 2: The main properties of the 45 steel.

Material properties	Value
Size (mm × mm × mm)	100 × 100 × 80
Density (g/cm ³)	7.85
Elastic modulus (GPa)	210
Poisson ratio	0.269
Chemical composition	C 0.46, Si 0.23, and Mn 0.59

TABLE 3: List of the milling parameters.

Cutting speed (m/min)	Feed rate (mm/z)	Cutting depth (mm)	Cutting width (mm)
20	0.045	5	10

and the flank wear bandwidth of the worn milling tool is $DL1 = 0.214$ mm, so the tool flank wear of the milling tool is $L = DL0 - DL1 = 0.266$ mm.

Based on the experimental design in Section 4.2, the milling tool wear experiment is performed, and the milling tool flank wear value is recorded. Taking the cutting times as the horizontal axis and milling tool flank wear value as the vertical axis, the milling tool flank wear curve is obtained by polynomial fitting as shown in Figure 5. According to the milling tool flank wear curve and the tool life test standard (GB/T 16460-2016), the milling tool wear states are divided into three categories whose scopes are listed in Table 4.

5. Results and Discussion

5.1. Data Preprocessing. The experiment has collected the original vibration signals of 410 cutting processes. After removing the abnormal signals, 392 segments vibration signals with a length of 1024 points are randomly intercepted to establish the data set. For each cutting process, there are three mutually perpendicular milling vibration signals collected from the accelerometer. Thus, the number of the samples in the whole dataset is 3×392 . In addition, the min-

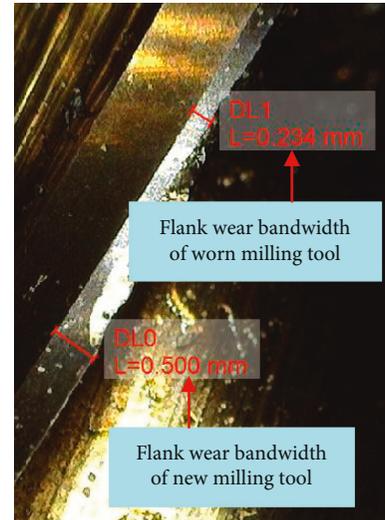


FIGURE 4: Measurement method of the tool flank wear.

max normalization method is used to standardize and normalize all input data to the range of (0-1). The calculation formula is expressed as follows:

$$X_{\text{Norm}} = \frac{(X - X_{\min})}{(X_{\max} - X_{\min})}. \quad (7)$$

5.2. Feature Extraction. Figure 6 depicts the waveforms of typical X -directional, Y -directional, and Z -directional vibration signals under three kinds of milling tool wear states. Clearly, the vibration signals present a certain periodic regularity. With the worsening of tool wear degree, amplitude of vibration signal becomes larger and larger; likewise, the periodic waveform becomes more and more obvious. So, vibration signals can reflect tool wear states. In addition, it can be found that the Z -directional vibration signals are larger than X -directional and Y -directional vibration signals. Because milling is an intermittent and nonstationary process, the relationship between vibration

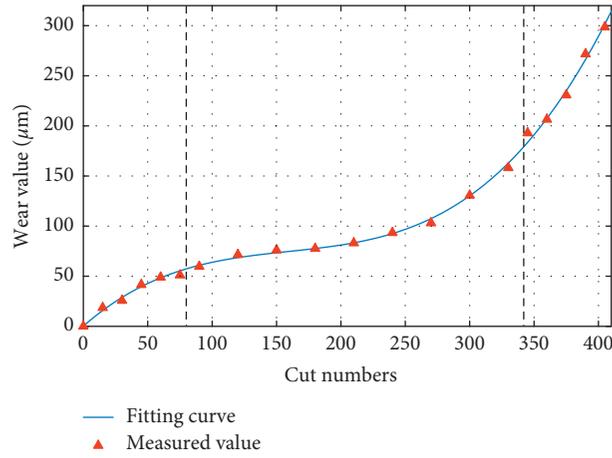


FIGURE 5: Milling tool flank wear curve.

TABLE 4: Tool wear categories and wear scope.

Tool wear category	Initial wear	Middle wear	Sever wear
Wear value	0–0.05 mm	0.05–0.2 mm	>0.2 mm
Class label	0	1	2

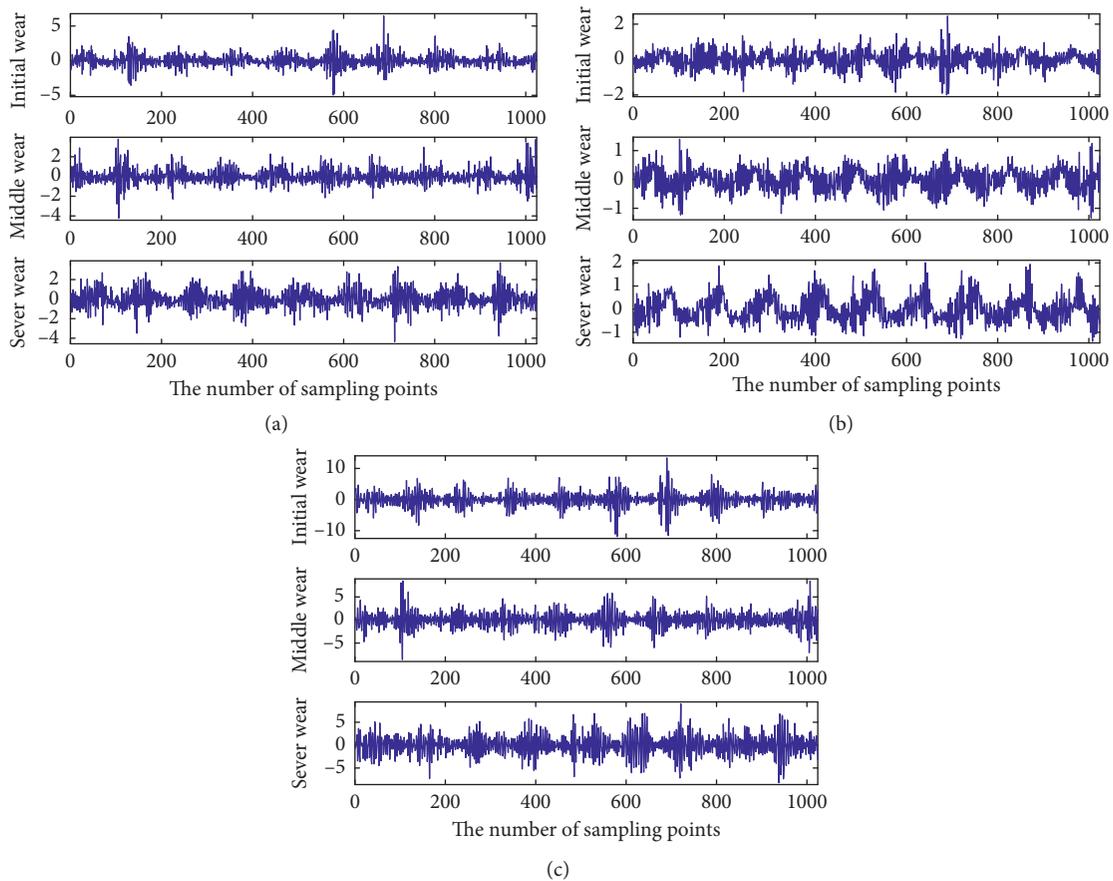


FIGURE 6: Vibration signal waveform in different milling tool wear states: (a) X-directional vibration signals, (b) Y-directional vibration signals, and (c) Z-directional vibration signals.

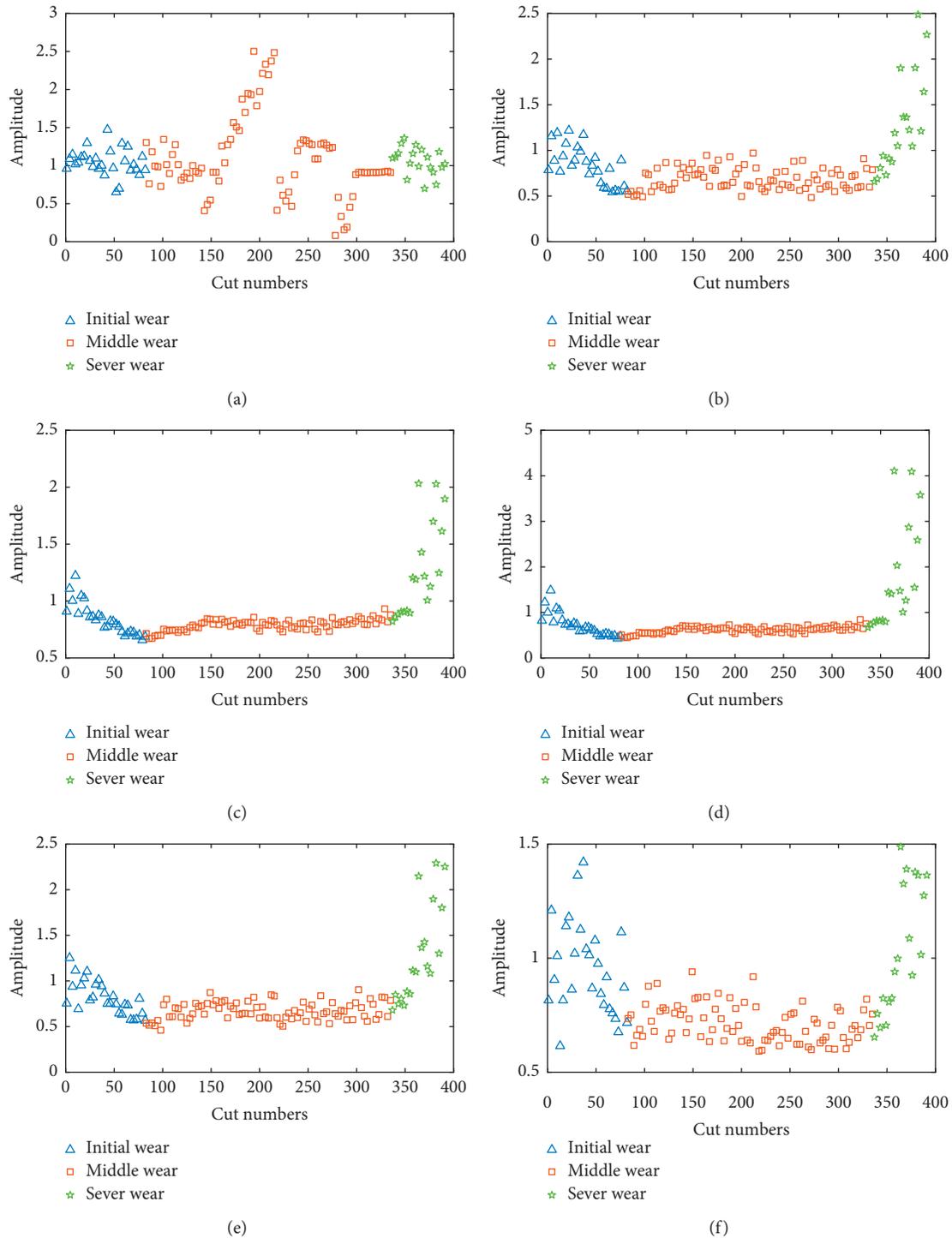


FIGURE 7: X-directional vibration signal feature extraction in time domain: (a) mean, (b) peak, (c) RMS, (d) variance, (e) kurtosis, and (f) skewness.

signals and tool wear cannot precisely be modeled. To reflect the tool wear process as accurately as possible, the time domain, frequency domain, and time-frequency domain features are extracted from the milling vibration signals as listed in Table 1.

5.2.1. Time Domain Features Analysis of Vibration Signals. In this section, six types of time-domain features are extracted from the milling vibration signals, including mean, peak, RMS, variance, kurtosis, and skewness. So, there are 18 time-domain features and all could be

obtained. Taking X -directional vibration signals as an example, the space distributions of time domain features of X -directional vibration signals are illustrated in Figure 7. It can be seen that these normalized features show a certain degree of clustering, which provides effective information for tool wear estimation. In addition, the space distribution of mean under three tool wear states are discrete (Figure 7(a)). Compared with mean feature, other features (Figures 7(b)–7(f)) are correlated with the tool wear states at a certain extent; however, this relationship is not obvious.

5.2.2. Frequency Domain Features Analysis of Vibration Signals. Based on the calculation equation of tool passing frequency (TPF) [38], in this paper, the TPF of milling tool is $f = (640 \times 4)/60 = 42.67$ Hz, where 640 r/min is the spindle rotating speed and 4 is the number of teeth of the cutting tool. Figure 8 depicts the spectrogram of vibration signals in three directions. It can be found that the amplitude of 1TPF of vibration signals in three directions increased from initial wear to sever wear, while other amplitudes of TPF of vibration signals did not. So, the amplitude of 1TPF of vibration signals is selected as frequency features.

Figure 9 illustrates the space distributions of frequency domain features of vibration signals under three tool wear states. Clearly, the amplitude of 1TPF of vibration signals in three directions increases with the increase in tool wear (cut numbers), which means that the amplitude of 1TPF of vibration signals can better characterize tool wear states.

5.2.3. Time-Frequency Domain Features Analysis of Vibration Signals. In the milling process, energy in each frequency band is different and changes with milling conditions [39]. So, in this paper, energy percentage of different frequency bands of vibration signals is used as time-frequency domain features to represent tool wear states. To extract the energy percentage of different frequency bands of vibration signals, the EEMD method [40, 41] is adopted and the IMFs of vibration signals are obtained. The calculation formula of energy percentage of IMF is listed in Table 1. Taking X -directional vibration signals as an example, the EEMD decompose results of vibration signals are given in Figure 10; there are only 10 IMFs shown in the figure because of limited space. In order to find the best IMFs which represented rich tool wear states, energy percentage of those IMFs are calculated, and the results shown in Figure 11(a). It can be found that the energy percentage of IMFs of vibration signals mainly focus on IMF1 to IMF7. Figures 11(b) and 11(c) illustrate the energy percentage of IMFs of Y -directional and Z -directional vibration signals. Based on the Figures 11(a)–11(c), we can get a conclusion that the percentage energy gives a good indication of the tool wear states and fully exhibits the disparities between selected eigenvalues.

5.3. Feature Selection. Based on the analysis above, 42 signal features are extracted from vibration signals for training and classification. Detailed information of extracted features of vibration signals in time domain, frequency domain, and time-frequency domain is depicted in Table 5. For lowering the dimension of the feature space and improving the computing speed of the model, a proper feature selection method is needed. By implementing feature selection algorithm, the complexity of modeling process could be reduced and new feature vectors are reconstructed. SVM-RFE is a sequential backward feature elimination method based on SVM and can give a feature ranking. This method has a very effective effect in gene selection, signal processing, and other fields [35]. Therefore, in this paper, SVM-RFE (in Algorithm 1) is used to find the features which can efficiently discriminate different tool wear states. In addition, the feature selection methods such as Pearson correlation and ReliefF are also used for performance comparison.

In the process of feature selection, a total of 392 samples which consist of 42 extracted features and one class feature are obtained. Randomly 198 samples are selected as training data and the remaining samples are selected as testing data first. Then, Pearson correlation, ReliefF, and SVM-RFE are used to get the rank of the 42 features, and three feature ranking lists are obtained. Finally, by using the obtained feature ranking lists, several nested feature subsets are constructed and inputted into the SVM model for model training. In order to obtain the optimal feature subset, the prediction accuracy of each test set is used as the criterion to evaluate the performance of the feature subset. In the SVM modeling process, the optimal parameter combination is provided by grid search algorithm. Figure 12 shows the classification accuracy as a function of the number of features selected by Pearson correlation, ReliefF, and SVM-RFE methods. It can be seen that the highest classification accuracy of SVM is 92.2%. To achieve this classification accuracy, 25 features are needed for Pearson correlation method, 18 features are needed for ReliefF method, and 12 features are needed for SVM-RFE method. So, SVM-RFE method is superior to Pearson correlation and ReliefF methods in feature selection. Therefore, the first 12 features are selected as the main features. The detailed information of selected main features is listed in Table 6.

5.4. SG Ensemble Model Training and Performance Evaluation. In this section, the SG ensemble model for tool wear state recognition is established. Then, performance evaluation indicators are applied to evaluate the model.

5.4.1. Model Training. Based on the selected main features of vibration signals, new samples have been generated with dimension 13×396 . Dividing the samples into an 80% training data set and an 20% testing data set, the training data set is used to construct the SG ensemble model for tool wear state recognition, and testing data set is used to verify the model. The performance of the SG ensemble model

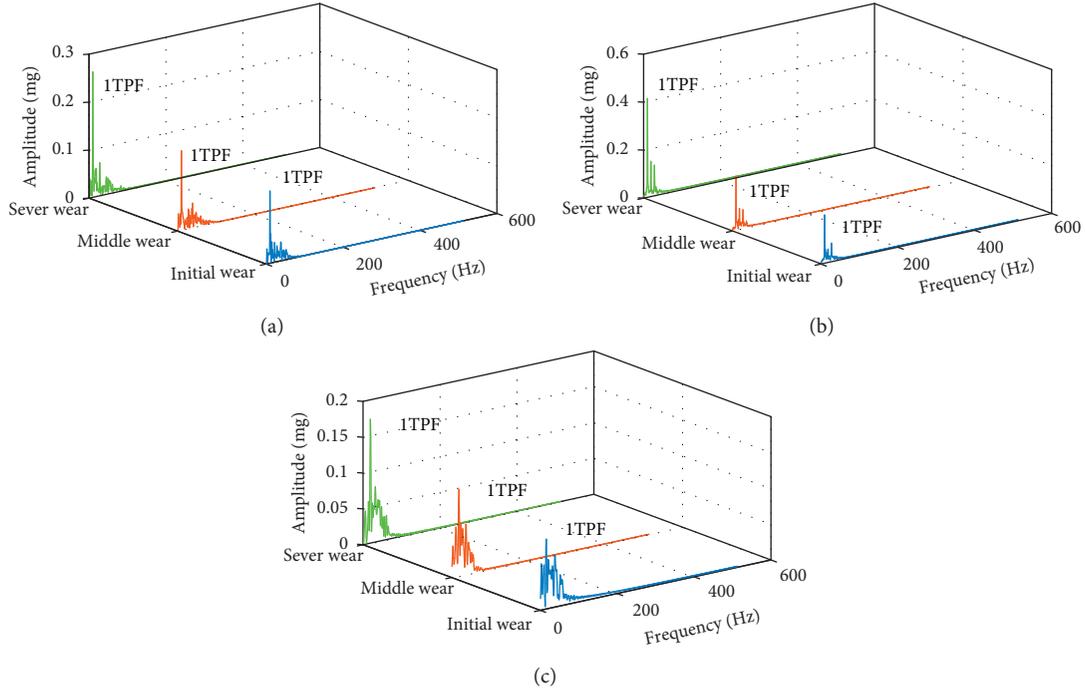


FIGURE 8: Spectrogram of vibration signals in different tool wear states: (a) X-directional vibration signals, (b) Y-directional vibration signals, and (c) Z-directional vibration signals.

depends on its modeling parameters. So, the grid search algorithm is used to find the optimal parameters of the model. When the number of iterations of the mesh search algorithm is 50, the optimal parameters are obtained, and it is given in Table 7. Based on those parameters, the SG ensemble model is established.

5.4.2. Performance Evaluation. In Figure 13, the confusion matrix [7] is used to evaluate the performance of the SG ensemble model. Each column of the matrix represents the instances in the actual class, while each row represents the instances in the recognized class. The number in the i^{th} row and j^{th} column represents samples whose target is the j^{th} class that was classified as i^{th} class. The evaluated misclassification (error) probability and the number of misclassifications are given in blue color, while the correct probability of each class and the number of correct classifications are given in pink color. Figure 13 shows that the 0th class is successfully recognized, with a success rate of 100%. Notwithstanding, some classes are not recognized so successfully, such as one of the 1st classes is misclassified as 0th class; thus, the rest is recognized correctly leading to 98.1% success rate. One of the 2nd classes is misclassified as 1st class, and the rest is recognized correctly leading to 98% success rate. Moreover, the overall classification accuracy of the SG ensemble model is 98.7%.

In this paper, another commonly used classification performance evaluating indicators, such as F -measure, G -mean, and area under the ROC curve (AUC) are adopted here. Also, the relevant calculation formulas are defined as follows:

$$\text{precision} = \frac{TP}{TP + FP},$$

$$\text{true positive rate} = \text{Recall} = \frac{TP}{TP + FN}, \quad (8)$$

$$\text{true negative rate} = \frac{TN}{TN + FP},$$

where TP is the number of true positives, TN is the number of true negatives, FN is the number of false negatives, and FP is number of false positives.

Assume k is the number of classes, P_i is the precision of i^{th} class, and R_i is the recall of i^{th} class. So, the F -measure is calculated as follows:

$$F\text{-measure} = \frac{2R_i P_i}{R_i + P_i}, \quad i = 1, 2, \dots, k. \quad (9)$$

For multiclass data, the G -mean is given as follows:

$$G\text{-mean} = \sqrt[k]{\prod_{i=1}^k R_i}. \quad (10)$$

Based on the above performance evaluation indicators, the performance evaluation results of the SG ensemble model are listed in Table 8. It can be seen that the model shows a better performance evaluation results, and all evaluation values are about 0.98.

5.5. Comparison with Other Ensemble Models. In order to demonstrate the advantages of the SG ensemble model as

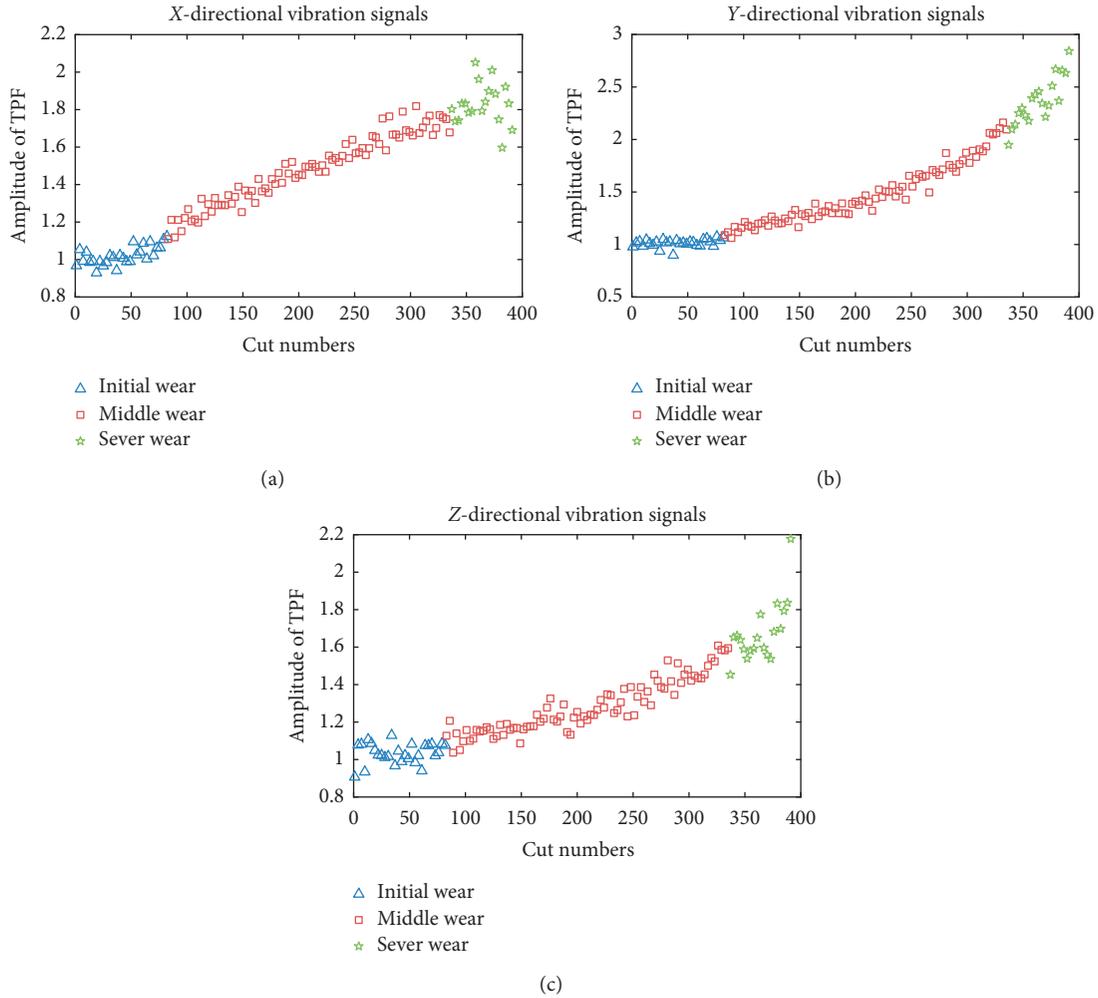


FIGURE 9: Vibration signal feature extraction in frequency domain: (a) X-directional vibration signals, (b) Y-directional vibration signals, and (c) Z-directional vibration signals.

discussed above, another kind of ensemble models such as Bagging-SVM and AdaBoost-SVM are adopted to recognize tool wear states, and training and testing for Bagging-SVM, AdaBoost-SVM, and SG ensemble model are conducted 50 times. By calculating the maximum, mean, and minimum of the recognition accuracy of each model, the final comparison results are obtained and shown in Figure 14. Clearly, the maximum, mean, and minimum of the recognition accuracy of proposed SG ensemble model are 98.3%, 97.2%, and 95.6%, respectively. So, the proposed model achieves better recognition accuracy. In addition, the performance evaluation index values of these models including AUC value and maximum deviation are given in Table 9. Clearly, the AUC value of AdaBoost-SVM and proposed SG ensemble model is larger than Bagging-SVM, but the difference in AUC value of them is small. The maximum deviation of Bagging-SVM, AdaBoost-SVM, and SG ensemble model are 13.4, 6.8, and 2.7, respectively, which means that the stability of SG ensemble model is better than other models. These above results prove that the

SG ensemble model has better performance than Bagging-SVM and AdaBoost-SVM.

5.6. Comparison with Other Single Models. In order to further verify the effectiveness and advancement of the proposed SG ensemble model, some other published methods are tested using the same set of training and testing data. Back-propagation neural network (BPNN) [42] and multiclass support-vector classification (MSVM) [7] are adopted to realize milling tool wear state recognition. In BPNN, the linear and sigmoid activation functions are designated for the hidden layer and output layer. The neural network parameters, namely, the number of neurons in hidden layer, learning rate, and maximum number of iterations are set to 25, 0.05, and 10000, respectively. In MSVM, the one-vs-one voting method is employed to achieve multiclassification, the RBF is used as kernel functions, and the grid search method is applied to obtain the optimal parameters. The recognition accuracy for tool wear states by using these

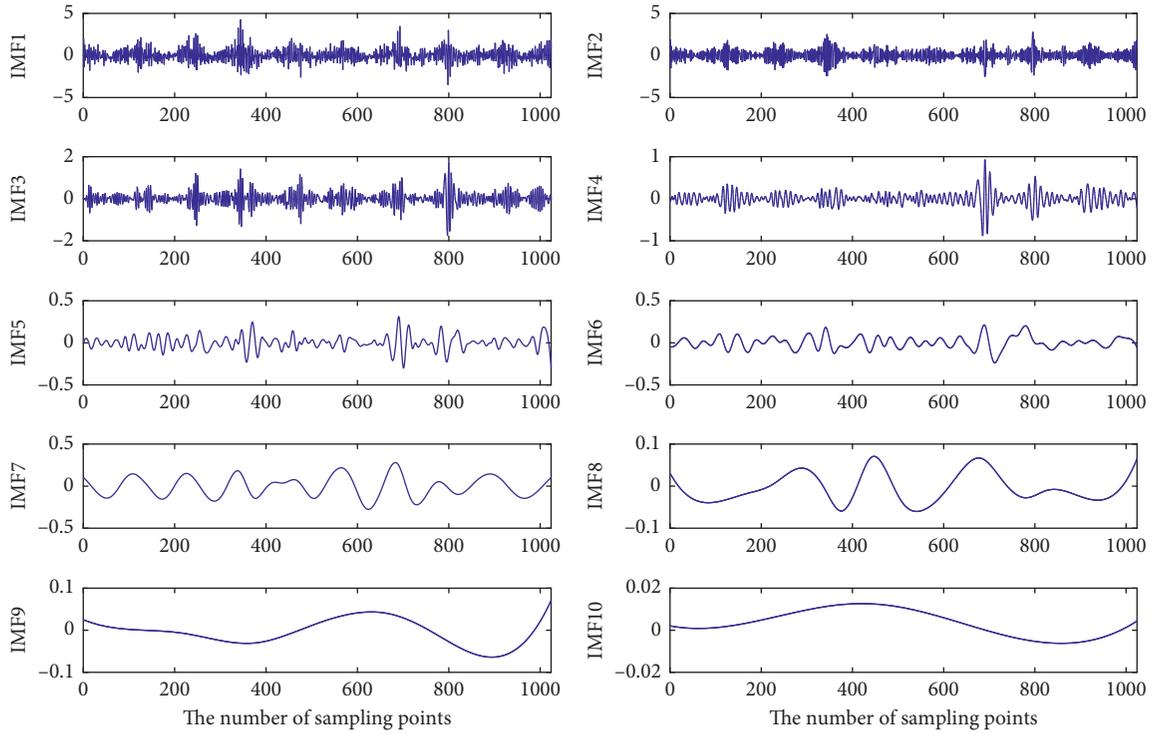
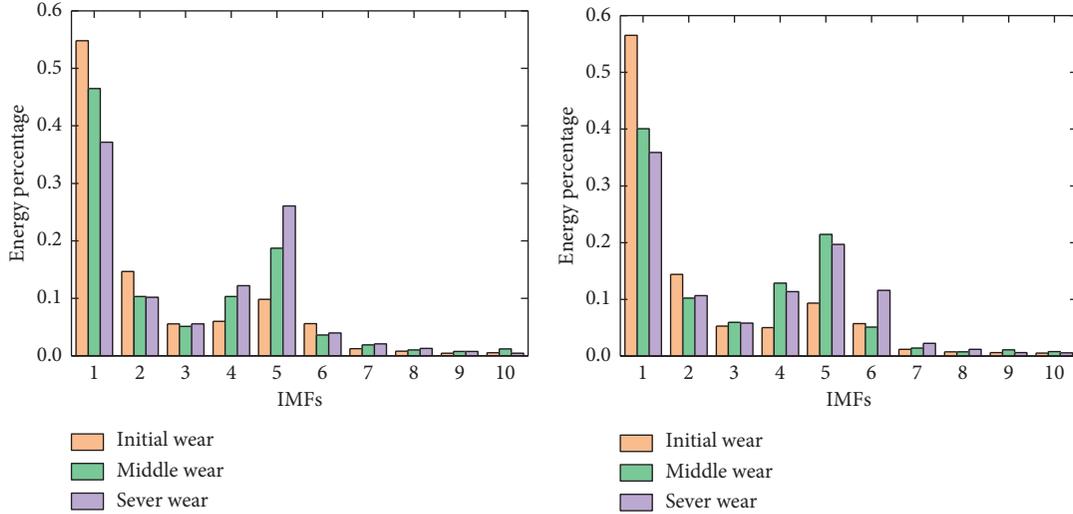


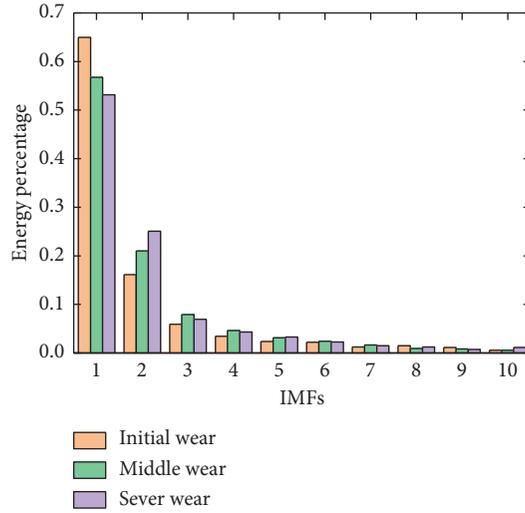
FIGURE 10: The EEMD decompose results of X-directional vibration signals.



(a)

(b)

FIGURE 11: Continued.



(c)

FIGURE 11: Energy percentage of IMFs of vibration signals: (a) X-directional vibration signals, (b) Y-directional vibration signals, and (c) Z-directional vibration signals.

TABLE 5: Extracted signal features.

Measured signal	Direction	Time domain	Frequency domain	Time-frequency domain	Total
Vibration signal	X/Y/Z	Mean, peak, RMS, variance, kurtosis, and skewness	Amplitude of TPF	IMF1-IMF7	42

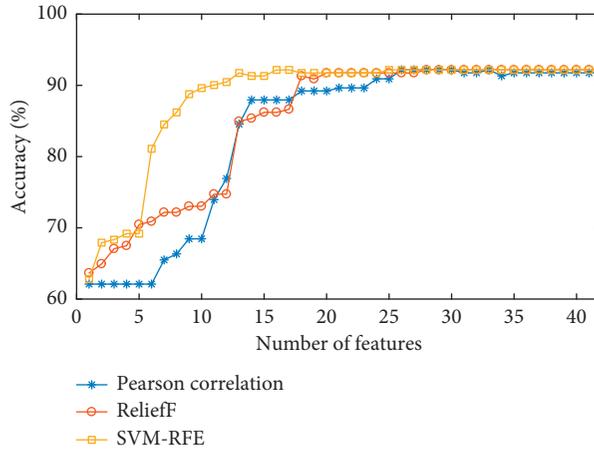


FIGURE 12: Classification accuracy as a function of the number of features.

TABLE 6: The detailed information of selected main features.

Signal	Features	Number of features	Total
X-directional vibration signal	Mean, RMS, amplitude of TPF	8, 2, 1	3
Y-directional vibration signal	Mean, peak, RMS, variance, skewness, amplitude of TPF, IMF1	11, 4, 10, 9, 7, 3, 5	7
Z-directional vibration signal	Peak, RMS	6, 12	2

TABLE 7: The optimal value of model parameters.

Initial parameters	Value or range	Optimal value
Penalty parameter (C)	[0.01, 0.05, 0.1, 0.2, 0.5, 2, 5, 10, 20, 50, 100]	2
Kernel function parameter (γ)	[0.01, 0.05, 0.1, 0.2, 0.5, 1, 5]	10
MaxDepth (De)	[3, 4, 5, 6, 7, 8, 9, 10]	5
Min samples split (Ss)	2	2
Min samples leaf (Sl)	1	1

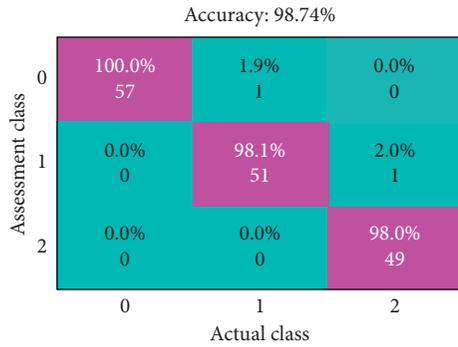


FIGURE 13: Confusion matrix of the SG ensemble model with overall classification accuracy 98.7%.

TABLE 8: Performance evaluation results of the SG ensemble model.

Class label	Precision	Recall	F-measure	G-mean	AUC
0	0.98	1	0.99		
1	0.98	0.98	0.98	0.98	0.98
2	1	0.98	0.99		

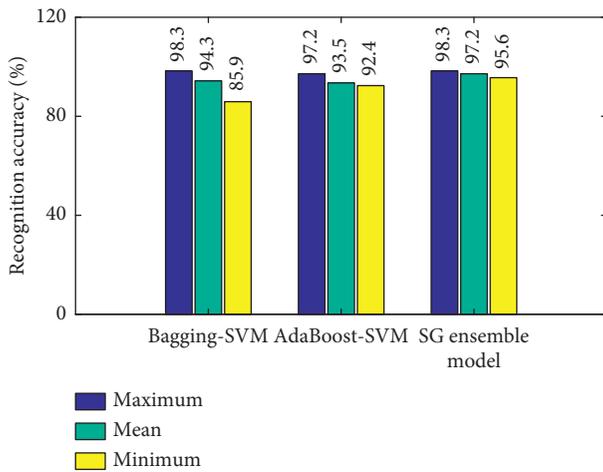


FIGURE 14: Comparison results of other ensemble models.

single models are illustrated in Figure 15. The predicted results reveal that the proposed SG ensemble model has better recognition accuracy than MSVM and BPNN.

TABLE 9: Performance comparison with other ensemble models.

Name	AUC value	Maximum deviation (%)
Bagging-SVM	0.917	13.4
AdaBoost-SVM	0.945	6.8
Proposed model	0.951	2.7

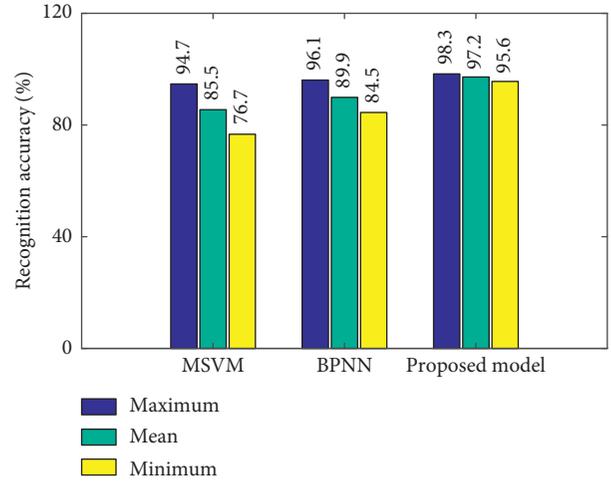


FIGURE 15: Comparison results of other single models.

TABLE 10: Performance comparison with other single models.

Name	AUC value	Maximum deviation (%)
MSVM	0.801	19.5
BPNN	0.851	18
Proposed model	0.951	2.7

Moreover, the evaluation index values (in Table 10) indicate that the proposed SG ensemble model has better stability than MSVM and BPNN.

6. Conclusion

In this paper, a SG ensemble model based on vibration signal features is proposed for milling tool wear state recognition. A total of 42 vibration signal features from time domain, frequency domain, and time-frequency domain are analyzed and extracted. Among them, 12 main features which are most relevant to tool wear states are selected by using SVM-RFE algorithm. Based on this, the SVM, DT, NB, and SG ensemble strategy is applied to establish the SG ensemble model. The experimental results demonstrate that the recognition accuracy of the established SG ensemble model is 98.7%. Comparison results reveal that the SG ensemble model has relatively higher accuracy and stability than other ensemble models (Bagging-SVM and AdaBoost-SVM) and other single models (MSVM and BPNN). Therefore, the proposed SG ensemble model is especially propitious to improve the accuracy and stability of TCM employing vibration signals.

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 article.

Acknowledgments

This research was supported by the National Key Research and Development Project of China (no. 2018YFB1701200).

References

- [1] N. Ghosh, Y. B. Ravi, A. Patra et al., "Estimation of tool wear during CNC milling using neural network-based sensor fusion," *Mechanical Systems and Signal Processing*, vol. 21, no. 1, pp. 466–479, 2007.
- [2] K. Zhu and B. Vogel-Heuser, "Sparse representation and its applications in micro-milling condition monitoring: noise separation and tool condition monitoring," *The International Journal of Advanced Manufacturing Technology*, vol. 70, no. 1–4, pp. 185–199, 2014.
- [3] D. E. Dimla and P. M. Lister, "On-line metal cutting tool condition monitoring," *International Journal of Machine Tools and Manufacture*, vol. 40, no. 5, pp. 739–768, 2000.
- [4] J. Karandikar, T. McLeay, S. Turner, and T. Schmitz, "Tool wear monitoring using naïve Bayes classifiers," *The International Journal of Advanced Manufacturing Technology*, vol. 77, no. 9–12, pp. 1613–1626, 2015.
- [5] L. N. L. de Lacalle, A. Lamikiz, J. Muñoa, and J. A. Sánchez, "Quality improvement of ball-end milled sculptured surfaces by ball burnishing," *International Journal of Machine Tools and Manufacture*, vol. 45, no. 15, pp. 1659–1668, 2005.
- [6] M. T. García-Ordás, E. Alegre, V. González-Castro, and R. Alaiz-Rodríguez, "A computer vision approach to analyze and classify tool wear level in milling processes using shape descriptors and machine learning techniques," *The International Journal of Advanced Manufacturing Technology*, vol. 90, no. 5–8, pp. 1947–1961, 2017.
- [7] A. Kothuru, S. P. Nooka, and R. Liu, "Application of audible sound signals for tool wear monitoring using machine learning techniques in end milling," *The International Journal of Advanced Manufacturing Technology*, vol. 95, no. 9–12, pp. 3797–3808, 2018.
- [8] M. Hu, W. Ming, Q. An, and M. Chen, "Tool wear monitoring in milling of titanium alloy Ti-6Al-4 V under MQL conditions based on a new tool wear categorization method," *The International Journal of Advanced Manufacturing Technology*, pp. 1–12, 2019.
- [9] X.-C. Cao, B.-Q. Chen, B. Yao, and W.-P. He, "Combining translation-invariant wavelet frames and convolutional neural network for intelligent tool wear state identification," *Computers in Industry*, vol. 106, pp. 71–84, 2019.
- [10] Y. Chen, Y. Jin, and G. Jiri, "Predicting tool wear with multi-sensor data using deep belief networks," *The International Journal of Advanced Manufacturing Technology*, vol. 99, no. 5–8, pp. 1917–1926, 2018.
- [11] 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, 2018.
- [12] D. Kong, Y. Chen, and N. Li, "Force-based tool wear estimation for milling process using Gaussian mixture hidden Markov models," *The International Journal of Advanced Manufacturing Technology*, vol. 92, no. 5–8, pp. 2853–2865, 2017.
- [13] Y. Zhou and W. Xue, "Review of tool condition monitoring methods in milling processes," *The International Journal of Advanced Manufacturing Technology*, vol. 96, no. 5–8, pp. 2509–2523, 2018.
- [14] D. Kong, Y. Chen, and N. Li, "Gaussian process regression for tool wear prediction," *Mechanical Systems and Signal Processing*, vol. 104, pp. 556–574, 2018.
- [15] N. C. Oza and K. Tumer, "Classifier ensembles: select real-world applications," *Information Fusion*, vol. 9, no. 1, pp. 4–20, 2008.
- [16] J. Heinermann and O. Kramer, "Machine learning ensembles for wind power prediction," *Renewable Energy*, vol. 89, pp. 671–679, 2016.
- [17] J. Yu, "Online tool wear prediction in drilling operations using selective artificial neural network ensemble model," *Neural Computing and Applications*, vol. 20, no. 4, pp. 473–485, 2011.
- [18] R. G. Zheng and X. M. Liu, "Effect of load and velocity on wear behaviour of Cu based self-lubricating composite," *Materials Research Innovations*, vol. 18, no. sup2, pp. S2–S12, 2014.
- [19] W.-A. Yang, W. Zhou, W. Liao, and Y. Guo, "Prediction of drill flank wear using ensemble of co-evolutionary particle swarm optimization based-selective neural network ensembles," *Journal of Intelligent Manufacturing*, vol. 27, no. 2, pp. 343–361, 2016.
- [20] Y. Yang, Y. Guo, Z. Huang et al., "Research on the milling tool wear and life prediction by establishing an integrated predictive model," *Measurement*, vol. 145, pp. 178–189, 2019.
- [21] K. Javed, R. Gouriveau, X. Li, and N. Zerhouni, "Tool wear monitoring and prognostics challenges: a comparison of connectionist methods toward an adaptive ensemble model," *Journal of Intelligent Manufacturing*, vol. 29, no. 8, pp. 1873–1890, 2018.
- [22] L. Breiman, "Bagging predictors," *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996.
- [23] G. Rätsch, T. Onoda, and K.-R. Müller, "Soft margins for AdaBoost," *Machine Learning*, vol. 42, no. 3, pp. 287–320, 2001.
- [24] C. Lemke and B. Gabrys, "Meta-learning for time series forecasting and forecast combination," *Neurocomputing*, vol. 73, no. 10–12, pp. 2006–2016, 2010.
- [25] P. Priore, B. Ponte, J. Puente, and A. Gómez, "Learning-based scheduling of flexible manufacturing systems using ensemble methods," *Computers & Industrial Engineering*, vol. 126, pp. 282–291, 2018.
- [26] B. M. Haddad, S. Yang, L. J. Karam, J. Ye, N. S. Patel, and M. W. Braun, "Multifeature, sparse-based approach for defects detection and classification in semiconductor units," *IEEE Transactions on Automation Science and Engineering*, vol. 15, no. 1, pp. 145–159, 2016.
- [27] S. K. Singh, S. Kumar, and J. P. Dwivedi, "A novel soft computing method for engine RUL prediction," *Multimedia Tools and Applications*, vol. 78, no. 4, pp. 4065–4087, 2019.
- [28] G. Wang, Y. Yang, and Z. Li, "Force sensor based tool condition monitoring using a heterogeneous ensemble learning model," *Sensors*, vol. 14, no. 11, pp. 21588–21602, 2014.

- [29] X. Wu, V. Kumar, J. Ross Quinlan et al., "Top 10 algorithms in data mining," *Knowledge and Information Systems*, vol. 14, no. 1, pp. 1-37, 2008.
- [30] Y. Kuo and K. P. Lin, "Using neural network and decision tree for machine reliability prediction," *The International Journal of Advanced Manufacturing Technology*, vol. 50, no. 9-12, pp. 1243-1251, 2010.
- [31] M. Hassan, A. Damir, H. Attia, and V. Thomson, "Benchmarking of pattern recognition techniques for online tool wear detection," *Procedia CIRP*, vol. 72, pp. 1451-1456, 2018.
- [32] V. N. Vapnik, *Statistical Learning Theory*, Wiley-Interscience, New York, NY, USA, 1998.
- [33] A. J. Myles, R. N. Feudale, Y. Liu, N. A. Woody, and S. D. Brown, "An introduction to decision tree modeling," *Journal of Chemometrics*, vol. 18, no. 6, pp. 275-285, 2004.
- [34] D. Soria, J. M. Garibaldi, F. Ambrogi, E. M. Biganzoli, and I. O. Ellis, "A "non-parametric" version of the naive Bayes classifier," *Knowledge-Based Systems*, vol. 24, no. 6, pp. 775-784, 2011.
- [35] G. S. Chen and Q. Z. Zheng, "Online chatter detection of the end milling based on wavelet packet transform and support vector machine recursive feature elimination," *The International Journal of Advanced Manufacturing Technology*, vol. 95, no. 1-4, pp. 775-784, 2018.
- [36] R. W. Maruda, G. M. Krolczyk, E. Feldshtein, P. Nieslony, B. Tyliczszak, and F. Pusavec, "Tool wear characterizations in finish turning of AISI 1045 carbon steel for MQCL conditions," *Wear*, vol. 372-373, pp. 54-67, 2017.
- [37] J. C. Su, C. K. Huang, and Y. S. Tarn, "An automated flank wear measurement of microdrills using machine vision," *Journal of Materials Processing Technology*, vol. 180, no. 1-3, pp. 328-335, 2006.
- [38] S. Orhan, A. O. Er, N. Camuşcu, and E. Aslan, "Tool wear evaluation by vibration analysis during end milling of AISI D3 cold work tool steel with 35 HRC hardness," *NDT & E International*, vol. 40, no. 2, pp. 121-126, 2007.
- [39] C. L. Zhang, B. Li, B. Q. Chen, H. R. Cao, Y. Y. Zi, and Z. H. He, "Weak fault signature extraction of rotating machinery using flexible analytic wavelet transform," *Mechanical Systems and Signal Processing*, vol. 64-65, pp. 162-187, 2015.
- [40] P. Wolszczak, G. Litak, and M. Dziuba, "Monitoring of drilling conditions using the Hilbert-Huang transformation," *MATEC Web of Conferences*, vol. 148, p. 16003, 2018.
- [41] P. Wolszczak, K. Łygas, and G. Litak, "Monitoring of cutting conditions with the empirical mode decomposition," *Advances in Science and Technology Research Journal*, vol. 11, no. 1, pp. 96-103, 2017.
- [42] G. C. M. Patel, A. K. Shettigar, and M. B. Parappagoudar, "A systematic approach to model and optimize wear behaviour of castings produced by squeeze casting process," *Journal of Manufacturing Processes*, vol. 32, pp. 199-212, 2018.

