



Tool wear state recognition based on GWO–SVM with feature selection of genetic algorithm

Xiaoping Liao¹ · Gang Zhou² · Zhenkun Zhang² · Juan Lu^{1,3} · Junyan Ma²

Received: 22 December 2018 / Revised: 16 April 2019 / Accepted: 20 May 2019 / Published online: 14 June 2019
© Springer-Verlag London Ltd., part of Springer Nature 2019

Abstract

Tool wear is an important consideration for Computerized Numerical Control (CNC) machine tools as it directly affects machining precision. To realize the online recognition of tool wear degree, this research develops a tool wear monitoring system using an indirect measurement method which selects signal characteristics that are strongly correlated with tool wear to recognize tool wear status. The system combines support vector machine (SVM) and genetic algorithm (GA) to establish a nonlinear mapping relationship between a sample of cutting force sensor signal and tool wear level. The cutting force signal is extracted using time domain statistics, frequency domain analysis, and wavelet packet decomposition. GA is employed to select the sensitive features which have a high correlation with tool wear states. SVM is also applied to obtain the state recognition results of tool wear. The gray wolf optimization (GWO) algorithm is used to optimize the SVM parameters and to improve prediction accuracy and reduce internal parameters' adjustment time. A milling experiment on AISI 1045 steel showed that when comparing with SVM optimized by commonly used optimization algorithms (grid search, particle swarm optimization, and GA), the proposed tool wear monitoring system can accurately reflect the degree of tool wear and achieves strong generalizability. A set of vibration signals are adopted to verify the presented research. Results show that the proposed tool wear monitoring system is robust.

Keywords Tool wear monitoring · Feature selection · Support vector machine · Parameter optimization

1 Introduction

Batch precision machining stimulated by high productivity levels and inexpensive high-quality products is competitive in manufacturing industries. Cutting tools are important components of machining that affect processing efficiency and product quality. As such, the development of tool condition monitoring (TCM) systems is desirable so as to monitor the

cutting process online and recognize tool wear state. With such systems, machining parameters can be timely adjusted to extend tool life and preserve product quality in mass production.

Over the last few decades, a few studies on machining process monitoring have been conducted [1, 2]. Two types of techniques, namely, direct and indirect sensing methods, can be used to measure tool wear. Direct sensing techniques use digital cameras (e.g., toolmaker's microscope) to measure tool conditions (e.g., flank wear width). However, direct measurements can only be employed in intermittent operations. On the contrary, indirect sensing techniques can continuously measure auxiliary in-process signals (e.g., cutting force [3, 4], vibration [5, 6], acoustic emission [7], and motor current [8]) to infer tool wear because in-process signals can be measured in real time; thus, indirect sensing techniques are suitable for real-time monitoring [9]. In indirect methods, signal features extracted from the original collected data of measurable signals include the time domain (TD) [10], frequency domain (FD) [11], time–frequency domain (TFD) [12], and fusion features [13].

✉ Juan Lu
lujuan3623366@163.com

¹ Guangxi Key Laboratory of Manufacturing Systems and Advance Manufacturing Technology, Guangxi University, Nanning 530004, China

² Department of Mechanical Engineering, Guangxi University, Nanning 530004, China

³ Department of Mechanical and Marine Engineering, Beibu Gulf University, Qinzhou 535011, China

However, each signal feature is not always correlated with tool wear. A subset of features that best predicts tool wear must be selected instead. The feature selection methods for tool wear can be divided according to whether they are related to subsequent tool wear model features. Specifically, feature selection methods may be filter and wrapper or embedded methods. Filter methods are independent of the subsequent tool wear model. Such techniques also use the statistical performance of all training data to directly evaluate and select features. Principal component analysis is a common filter method [14]. Model-independent feature selection methods are computationally fast, but a large deviation exists between feature evaluation and the performance of the tool wear model. Therefore, such a performance may not be optimized. Features can be selected by wrapper or embedded methods that consider the tool wear model [15]. Wrapper approaches generally employ optimization algorithms (such as genetic algorithm (GA) [16] and ant colony algorithm [17]) to search for optimal features with high correlation with the tool model. In the present study, GA, which is generally effective for the rapid global search of large, nonlinear, and poorly understood spaces [18], is used for the further selection of input features that are relevant to tool wear states. Similarly, Kaya et al. [19] employed GA to reduce the dimensionality of a feature set by selecting the features that correlate best with tool conditions, and they also demonstrated efficacy experimentally.

Using selected signal characteristics as inputs, a predictive model can be built, and tool wear must be estimated in real time. Data-driven approaches use learning algorithms, and experimental data build predictive models so that an in-depth understanding of underlying physical processes is not a prerequisite. Artificial neural network (ANN) [20, 21] and support vector machine (SVM) [22, 23] are the most widely applied data-driven approaches for tool wear monitoring. Relative to ANN, SVM is a powerful learning model that is free from the three problems of training efficiency, testing efficiency, and overfitting [24–26]. Besides, researchers can apply these models on different products obtained in different batches because the insensitive region of SVM absorbs small-scale random fluctuations in random response [24]. The internal parameter values of SVM greatly affect predictive model accuracy. Certain optimization algorithms are used to optimize SVM parameters and thereby obtain the best parameter combination and reduce parameter adjustment time. García proposed a particle swarm optimization (PSO)–SVM-based model and successfully used it to predict tool flank wear [27]. Li and Zhang used GA to optimize internal SVM parameters. The experimental results indicated that GA–SVM can effectively track the tool wear trend and achieve tool wear and tool life monitoring [28, 29]. Zhang et al. proposed a pattern recognition approach for intelligent fault diagnosis of rotating machinery by using the SVM optimized by the ant colony algorithm. The simulation results revealed that the model

has strong generalizability and high prediction accuracy [30]. The gray wolf optimization (GWO) algorithm is an emerging group intelligent optimization method that mimics the leadership level and hunting mechanism of gray wolves in nature. GWO also has a simple structure, easy implementation, and good global performance, all of which are beneficial to engineering implementation [31]. However, GWO algorithm has never been employed to optimize internal SVM parameters for tool wear monitoring through a complex and stochastic process. Therefore, SVM optimized by the GWO algorithm (GWO–SVM) is developed in the current work to build a prediction model of tool wear.

In this work, a TCM system is developed by using an indirect measurement method that is based on GWO–SVM. First, the cutting force signal of the machining process is collected via a cutting process test. Second, the main TD, FD, and wavelet domain features of the signal samples that must be evaluated are extracted. In addition, feature selection (dimensionality reduction) is performed by GA, which avoids the processing complexity of high-dimensional nonlinear feature data and weakens the noise component characteristics of the signal. GWO–SVM is employed to evaluate the dimensionality reduction features and obtain the tool wear grade. Moreover, the robustness of the tool wear monitoring system is further validated with the vibration signals from ref. [32].

2 Methodologies

The tool wear monitoring system used in this study is illustrated in Fig. 1. The input is the cutting force signal of a real-time milling process, and the output is the wear state of the milling cutter (initial, normal, and severe wear). The specific implementation process is divided into three modules: feature extraction, feature selection, and tool wear state classification modules. Their details are introduced in the following subsections.

2.1 Feature extraction module

During the process of machining, the dynamic cutting force data generated are always large. Such huge data can affect the calculation speed, and the online wear monitoring system must reflect the result quickly. To improve the calculation speed, the feature extraction module significantly reduces the dimension of raw data (in TD and FD). The module also aims to maintain the relevant information about tool wear conditions in the extracted features. Signal characteristics are extracted using TD analysis, FD analysis, and four-layer wavelet packet decomposition. Table 1 presents the 28 corresponding feature values extracted, of which S401–S416 represent the 16 corresponding proportions of wavelet energy.

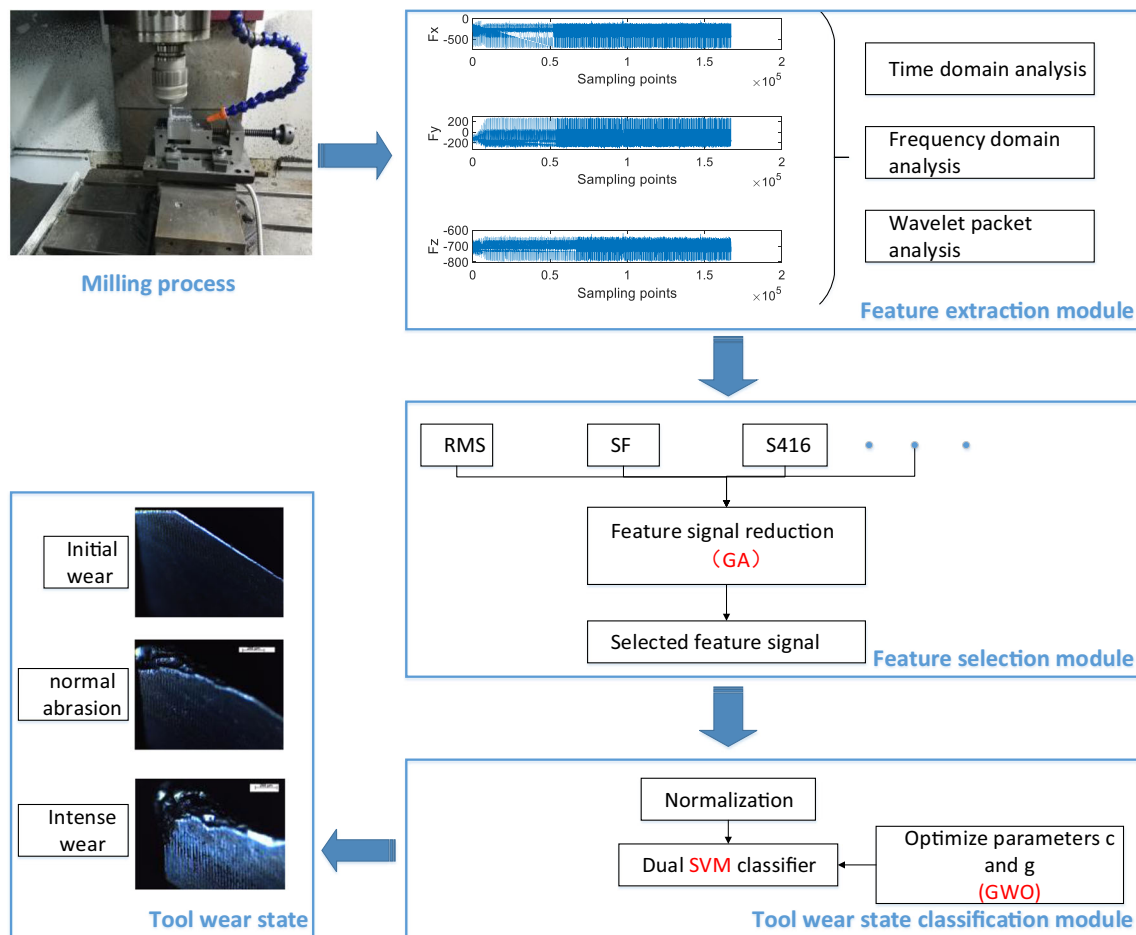


Fig. 1 Tool wear state recognition model

2.2 Feature selection module

Multiple types of features extracted through feature extraction module can result in many potential choices. Thus, combining them into a classification module is not efficient. For such a

reason, GA is used for feature selection. The GA process includes replication, hybridization, mutation, competition, and selection. Its advantages are simplicity, randomness, and compatibility with other algorithms. Detailed descriptions of the GA feature selection can be found in [32]. The fitness function set in

Table 1 Statistical parameters and mathematical expressions in the TFD

Feature	Description	Feature	Description
Mean	$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$	Waveform factor	$S = x_{rms} / \frac{1}{n} \sum_{i=1}^n x_i $
Root mean square	$x_{rms} = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$	Peak factor	$C = x_{peak} / x_{rms}$
Peak	$x_{peak} = \frac{1}{m} \sum_{j=1}^m x_j$	Center of gravity frequency	$FC = \frac{\int_0^{+\infty} f s(f) df}{\int_0^{+\infty} s(f) df}$
Square root amplitude	$x_r = \left(\frac{1}{n} \sum_{i=1}^n \sqrt{ x_i } \right)^2$	Mean square frequency	$MSF = \frac{\int_0^{+\infty} f^2 s(f) df}{\int_0^{+\infty} s(f) df}$
Variance	$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$	Frequency variance	$VF = \frac{\int_0^{+\infty} (f - FC)^2 s(f) df}{\int_0^{+\infty} s(f) df}$
Skewness	$C_s = \frac{1}{n} \sum_{i=1}^n \frac{(x_i - \bar{x})^3}{\sigma^3}$	Wavelet packet decomposition	S401–S416
Kurtosis	$K = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4 / x_{rms}^4$		

GA feature selection module is $fitval = 1 - |r|$, and the closer the value of “fitval” is to 0, the higher the correlation between the feature and the tool wear. r is the Pearson correlation coefficient, and the expression is as follows:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \tag{1}$$

In Formula (1), r represents the correlation coefficient between variables x and y .

2.3 Tool wear state classification module

2.3.1 SVM classification algorithm

SVM is a common method of classification and pattern recognition. Proposed by Vladimir N. Vapnik and Alexey Ya, SVM aims to transform a signal into a high-dimensional feature space and then solve a binary problem in a hyper plane [33].

Taking the soft edge problem as an example to illustrate the SVM classification calculation method, this study provides the specific calculation formula. When the input data are linearly separable (hard edge problem), they can be solved by taking a small λ value. The specific calculation method is as follows:

$$\left[\frac{1}{n} \sum_{i=1}^n \max(0, 1 - y_i(\vec{w} \vec{x}_i - b)) \right] + \lambda \|\vec{w}\|^2 \tag{2}$$

The initial calculation is to change Formula (2) into a constrained optimization problem with a differentiable objective function. For $i \in \{1, \dots, n\}$, variable $\zeta_i = \max(0, 1 - y_i(w x_i - b))$ is introduced. ζ_i satisfies $y_i(w x_i - b) \geq 1 - \zeta_i$ and is the minimum non-negative number. Thus, the optimization problem becomes

$$\text{minimize } \frac{1}{n} \sum_{i=1}^n \zeta_i + \lambda \|\vec{w}\|^2 \tag{3}$$

The given constraint is $y_i(w x_i - b) \geq 1 - \zeta_i$, and $\zeta_i \geq 0$.

The given Lagrangian dual problem is solved by the following simplification:

$$\begin{aligned} &\text{minimize } f(c_1, \dots, c_n) = \\ &\sum_{i=1}^n c_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i c_i (x_i y_j) y_j c_j \\ &\text{s.t. } \begin{cases} \sum_{i=1}^n c_i y_i = 0 \\ 0 \leq c_i \leq \frac{1}{2n\lambda} \end{cases} \end{aligned} \tag{4}$$

The dual maximization problem is a quadratic function problem of the c_i linear constraint. This problem can be effectively solved by the quadratic programming algorithm. Thus, c_i has the following definition:

$$\vec{w} = \sum_{i=1}^n c_i y_i \vec{x}_i \tag{5}$$

When \vec{x}_i is on the correct side of the edge line, $c_i = 0$; when \vec{x}_i is on the edge line, $0 \leq c_i \leq (2\eta\lambda)^{-1}$. Given the definition, \vec{w} can be written as a linear combination of support vectors. Offset b can also be calculated by \vec{x}_i .

$$y_i(\vec{w} \vec{x}_i - b) = 1 \leftrightarrow b = \vec{w} \vec{x}_i - y_i \tag{6}$$

where $y_i = \pm 1$ and $y_i^{-1} = y_i$.

Nonlinear problems can be transformed into linear problems in a high-dimensional space by nonlinear function mapping and processing. Suppose that a nonlinear classification model must be trained. After mapping point \vec{x}_i to a high-dimensional space on the model, the nonlinear classification model can be changed into a linear model, that is, $\vec{x}_i \rightarrow \varphi(\vec{x}_i)$. Then, a kernel function can be defined.

$$k(\vec{x}_i, \vec{x}_j) = \varphi \vec{x}_i \varphi \vec{x}_j \tag{7}$$

The classification vector \vec{w} in the converted space can be expressed as follows:

$$\vec{w} = \sum_{i=1}^n c_i y_i \varphi(\vec{x}_i) \tag{8}$$

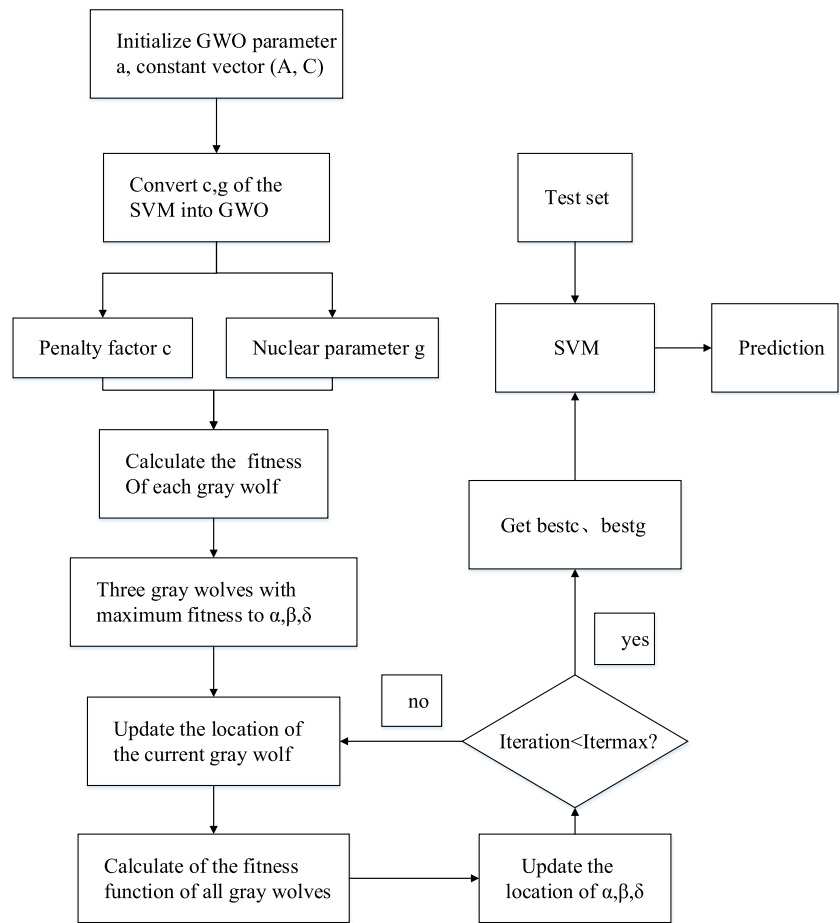
where c_i can be solved by the following formula.

$$\begin{aligned} &\text{minimize } f(c_1 \dots c_n) \\ &= \sum_{i=1}^n c_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i c_i (\varphi(\vec{x}_i) \varphi(\vec{x}_j)) y_j c_j \\ &= \sum_{i=1}^n c_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i c_i k(\vec{x}_i, \vec{x}_j) y_j c_j \\ &\text{s.t. } \begin{cases} \sum_{i=1}^n c_i y_i = 0 \\ 0 \leq c_i \leq \frac{1}{2n\lambda} \end{cases} \end{aligned} \tag{9}$$

Coefficient c_i can be solved according to the aforementioned method. b can also be solved according to the point $\varphi(\vec{x}_i)$ on the edge line in the transformed space.

$$\begin{aligned} b &= \vec{w} \varphi(\vec{x}_i) - y_i = \left[\sum_{k=1}^n c_k y_k \varphi(\vec{x}_k) \varphi(\vec{x}_i) \right] - y_i \\ &= \left[\sum_{k=1}^n c_k y_k k(\vec{x}_k, \vec{x}_i) \right] - y_i \end{aligned} \tag{10}$$

Fig. 2 GWO of SVM parameters



The classification of the new points after the conversion space is calculated.

$$\vec{z} \rightarrow \text{sgn}(\vec{w}\varphi(\vec{z}) - b) = \text{sgn}\left(\left[\sum_{i=1}^n c_i y_i k(\vec{x}_i, z)\right] - b\right) \tag{11}$$

c_i and b are two important SVM parameters. This study uses a radial basis function as a kernel function, which is presented below.

$$k(\vec{x}_i, \vec{x}_j) = \exp\left(\frac{-\|\vec{x}_i - \vec{x}_j\|^2}{2g^2}\right) \quad g > 0 \tag{12}$$

This study focuses on the optimum combination of penalty factor c and nuclear parameter g of the final model [34].

2.3.2 GWO-SVM

To improve the prediction accuracy of SVM, this study uses the GWO algorithm in optimizing penalty factor c and nuclear parameter g in SVM. Mirjalili (2014) first proposed GWO, which is a relatively new algorithm [35]. The GWO algorithm

simulates the hunting behavior process under the gray wolf hierarchy. The algorithm can be divided into three stages: encirclement, pursuit, and attack. The mathematical model is described as follows.

(1) Surround

The first step in the hunting process is the encirclement of prey, which can be described as follows:

$$\vec{D} = |\vec{C} * \vec{X}_p(t) - \vec{X}(t)| \tag{13}$$

$$\vec{X}(t+1) = \vec{X}_p(t) - \vec{A} * \vec{D},$$

where \vec{D} and \vec{A} are the weighting coefficients, \vec{X}_p is the position vector of the prey, t is the number of iterations, and \vec{X} is the position of the wolf.

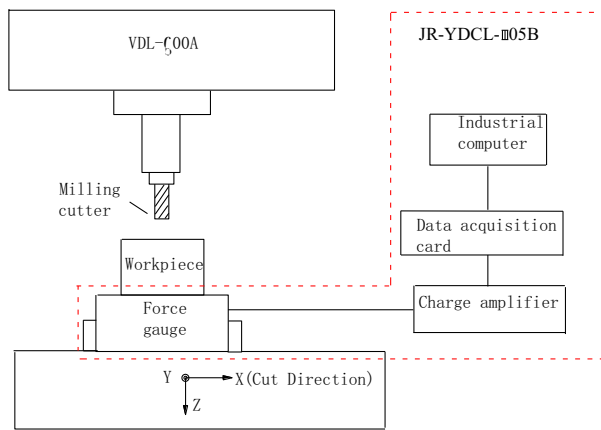
In Formula (13), coefficients \vec{A} and \vec{C} are respectively updated as follows:

$$\vec{A} = 2\vec{a} * r_1 - \vec{a}, \tag{14}$$

$$\vec{C} = 2r_2,$$



Fig. 3 Milling processing platform



where $|\vec{a}|$ is a linearly decreasing constant and $|\vec{a}| \in (2, 0)$; r_1, r_2 are constants, and $r_1, r_2 \in (0, 1)$.

(2) Hunt

Gray wolves can identify the position of their prey and surround it. The hunting process is usually dominated by α , and β and δ are responsible for the assistance. In an abstract space, the best position of the prey is unknown. Thus, when simulating the hunting behavior of gray wolves in mathematics, the position of α is assumed to be the best candidate solution. Moreover, β and δ can further understand the potential of the prey. In terms of location, the three best solutions that are currently available should be prepared. Other search units (including ω) can also be asked to update their location by search. The formula is expressed as follows:

$$\begin{aligned} \vec{D}_\alpha &= |\vec{C} * \vec{X}_\alpha - \vec{X}|, \vec{D}_\beta = |\vec{C} * \vec{X}_\beta - \vec{X}|, \\ \vec{D}_\delta &= |\vec{C} * \vec{X}_\delta - \vec{X}|, \vec{X}_1 = \vec{X}_\alpha - \vec{A}_1 * (\vec{D}_\alpha), \\ \vec{X}_2 &= \vec{X}_\beta - \vec{A}_2 * (\vec{D}_\beta), \vec{X}_3 = \vec{X}_\delta - \vec{A}_3 * (\vec{D}_\delta) \\ \vec{X}(t+1) &= (\vec{X}_1 + \vec{X}_2 + \vec{X}_3) / 3, \end{aligned} \tag{15}$$

where $\vec{X}(t+1)$ is the optimal solution for the current iteration.

(3) Attack

The attack is the last step in the hunting process. The purpose is to capture the target, that is, to complete GWO. According to the formula, $\vec{A} = 2\vec{a} * r_1 - \vec{a}$. When \vec{a} is linearly descending from 2 to 0, \vec{A} is decremented, that is, $\vec{A} \in (-2a, 2a)$. The judgment condition is $|\vec{A}|$. When $|\vec{A}| < 1$, the attack is initiated (GWO finds the optimal solution); when $|\vec{A}| > 1$, the wolf group diverges (seeking the new optimal solution).

The process of using GWO to optimize the SVM algorithm is depicted in Fig. 2.

The specific steps are as follows.

- Step 1 Initialization: The parameter \vec{a} of the GWO algorithm and the constant vectors \vec{A} and \vec{C} are initially assigned. The range of values is also agreed upon.
- Step 2 Parameters c (penalty factor) and g (nuclear parameter) of SVM are used as the 2D coordinates of the individual positions of the wolves.
- Step 3 The SVM model is trained through the training set samples to calculate the fitness. The ranks of the wolves (α , β , δ , and ω) are classified according to the fitness values. The formula for calculating the fitness value is as follows:

$$\text{Fit} = 1 - \frac{y_t}{y_t + y_f} \tag{16}$$

Table 2 Basic information of workpiece material

Material grade	AISI1045
Material name	High quality carbon steel
Hardness	HRC30
Size	80 × 50 × 40 (mm)
Performance	A commonly used steel with good plasticity, toughness, and high strength

Table 3 Level of machining variable values

Cutting amount	Symbol	Unit	Value
Spindle speed	V	m/min	47.12
Feed rate	f	mm/r	0.2
Axial depth of cut	ap	mm	1.0
Radial depth of cut	ae	mm	5.0

Table 4 GA parameter settings

Parameter	Value
Population size	28
Elite count	2
Crossover rate	0.8
Mutation rate	0.001

where y_c is the correct number of classifications and y_f is the number of classification errors.

- Step 4 The position of the wolves is updated by the position update formula.
- Step 5 The fitness update of individuals is calculated, and the optimal fitness value of the current algebra is recorded as Fit_best . If $Fit_{best} > Fit_{\alpha}$ (the fitness of α), then Fit_{α} is updated to Fit_{best} , and the corresponding position is recorded. If $Fit_{\beta} < Fit_{best} < Fit_{\alpha}$, then Fit_{best} is assigned to β . The corresponding position is also updated to β . If $Fit_{\delta} < Fit_{best} < Fit_{\beta}$, then Fit_{best} and the corresponding position are updated to δ . The position of α is the optimal position for the population that must be iterated to the current algebra.
- Step 6 Whether the number of iterations reaches the set threshold or the global optimal position meets the minimum limit is determined. If the condition is satisfied, then the loop is terminated, and the optimal parameters $bestc$ and $bestg$ are returned. Otherwise, the process returns to Step 4.
- Step 7 SVM classification models are established through $bestc$ and $bestg$.

3 Milling experiment and feature selection

3.1 Experimental design

This experiment uses VDL-600A CNC machining center, JR-YDCL-III05B piezoelectric three-measurement system, and a

metallographic microscope to build a test platform (Fig. 3). The tool is a two-flute tungsten carbide-coated solid carbide milling cutter with a diameter of 10 mm. The material of the workpiece is carbon steel (AISI 1045). AISI 1045 has a good processing performance and wide application range. AISI 1045 is also mainly used to manufacture moving parts with high strength and high surface quality, such as turbine impellers and compressor pistons. The basic information of AISI 1045 is presented in Table 2.

Table 3 shows that the cutting parameters in the experiment are fixed and that the flank wear value (VB) is measured after a complete 200-mm cutting distance using a metallographic microscope. The average VB of the two blades of the milling cutter is considered the tool wear value. The sampling frequency of the force gauge is set to 20,000 Hz, and the cutting force data generated in the first 50 mm distance are collected during each pass. The cutting force signal includes three directions: X, Y, and Z. A total of 108 sets of data are obtained.

According to the feature extraction method of the statistical signal and wavelet packet analysis in the TFD, the number of extracted features is up to $28 \times 3 = 84$ (the force has three directions). In addition, the extracted feature directly used as an input can cause a dimensional disaster, and the calculation amount is large. The obtained mathematical models often fail to meet the target requirements at the same time. Therefore, obtaining a wear state recognition model with good calculation performance requires the selection of the most relevant feature signals from the model.

3.2 Feature selection based on GA

This study uses MATLAB GA toolbox for feature selection. The specific GA parameters are set and are shown in Table 4.

According to the parameter settings in Table 4, the force signals in the three directions of x , y , and z are selected. When GA reaches the termination condition, the selected feature index value and the optimal and average fitness values are returned.

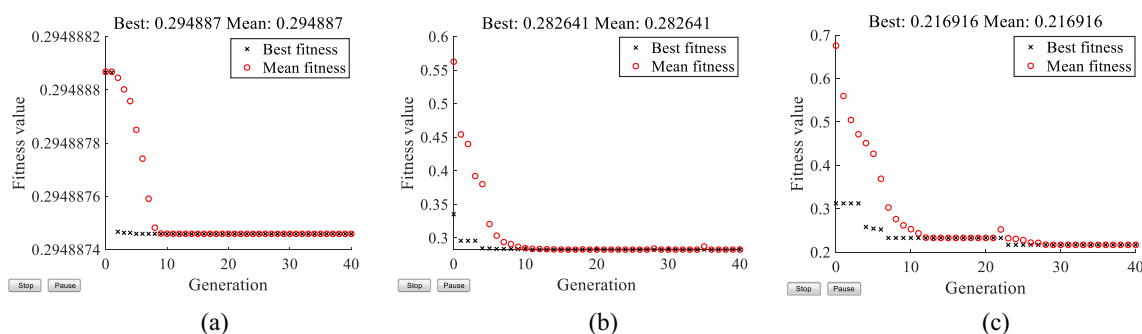


Fig. 4 a–c GA feature selection

Table 5 Selected features by GA (force signals)

Signal	Signal features				
F_x	σ^2	S402	S410	S413	S416
F_y	σ^2	S402	S406	S410	S416
F_z	σ^2	S413	S414	S416	

Figure 4a, b, and c represent the feature selection process of F_x , F_y , and F_z by GA, respectively. In Fig. 4a, the cycle is terminated when the iteration is 40. Here, the optimal and average fitness values are 0.295. The correlation between the selected features and the wear amount is high, and the absolute value of the Pearson coefficient is close to 0.7. Figure 4b illustrates that the average fitness value is 0.283 after 10 iterations from 0.56 and that the whole cycle is terminated when iterating to the 40th generation of the population. The final optimal fitness value is 0.283, which indicates that the Pearson coefficient value $|r|$ between the selected F_y eigenvalues and the tool wear is also above 0.7. In Fig. 4c, the average and best fitness values converge to the optimal in the 27th generation. Both values equal to

0.217, that is, the correlation between the selected feature and the wear value is above 0.78.

The feature selection program is repeated 30 times in the repeatability test to find the features related to tool wear and avoid large random errors in a single operation. The features with high frequency of occurrence are selected from Table 5. The number of signal features selected is 14, which is significantly lower than the previous number of 84 feature vectors.

3.3 Feature correlation analysis

To verify the validity of the GA feature selection, this study plots the correlation between the feature and the wear value (Fig. 5). The abscissa is the tool wear value, and the ordinate is the corresponding feature value. Given the space limitations, only 3 of the 14 features selected on the basis of GA are used. The correlation of these features with wear is illustrated in Fig. 5a, c, and e. Figure 5b, d, and f present the correlations between the feature and wear values that are not selected by GA.

Among the characteristic signals of F_x , Fig. 5a is the correlation diagram between the characteristic $F_x_{\sigma^2}$ (variance of the F_x signal) and the wear value selected by

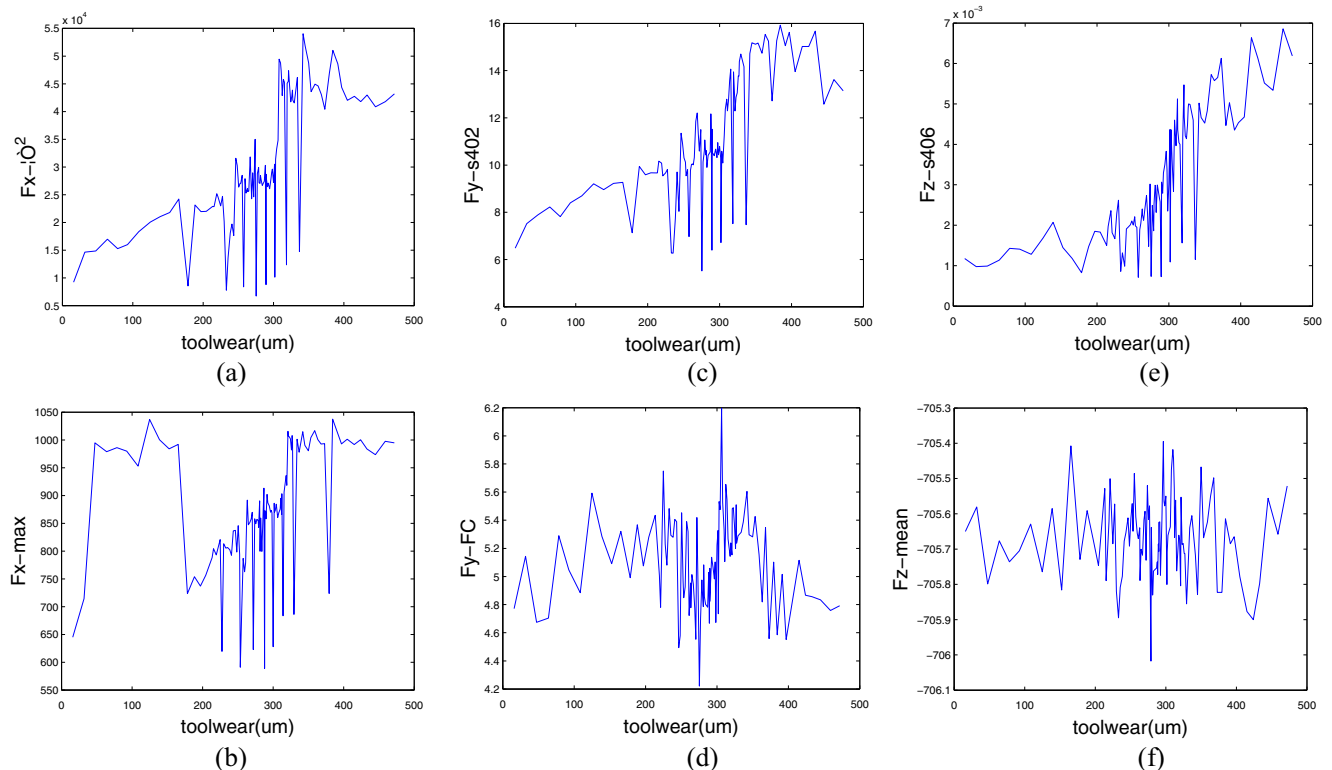


Fig. 5 a–f Analysis of the correlation between the features and tool wear

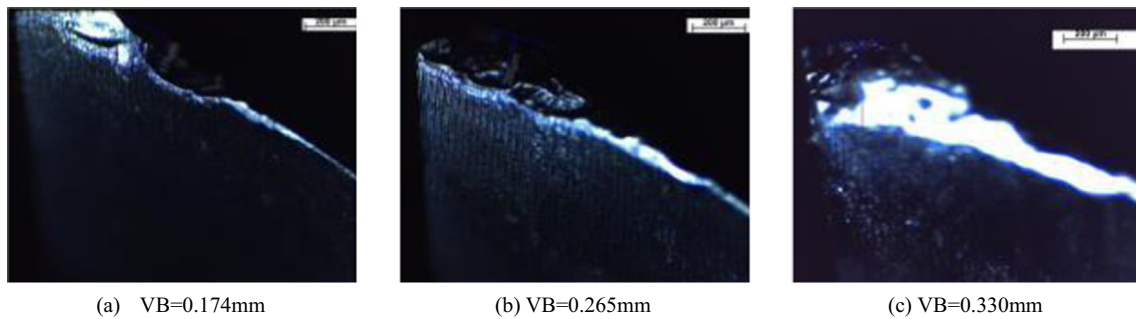


Fig. 6 Wear process of cutting tool

GA. In the diagram, Fx_{σ^2} increases significantly as the wear value increases. The correlation is evident. The graph in Fig. 5b presents the correlation curve between the Fx_{max} (maximum Fx signal) and the wear value that is not selected by GA. When the wear value increases, no evident variation law of Fx_{max} exists. The correlation of Fx_{max} with the wear value is not as great as that of Fx_{σ^2} (selected feature).

The correlation between the eigenvalue of Fy signal and the wear value is illustrated in Fig. 5c and d. Figure 5c presents the selected feature Fy_{s402} (the ratio of the energy of the s402 band to the total energy after reconstructing the wavelet packet decomposition of the Fy signal), which increases as the wear value increases. By contrast, Fig. 5d shows that the unselected feature Fy_{FC} (FC of the Fy signal) exhibits random fluctuations when the wear value changes. Such fluctuations are minimal.

Among the Fz signal characteristics, the selected and unselected features are displayed in Fig. 5e and f, respectively. The selected feature Fz_{s406} has a significant positive correlation with wear value, whereas Fz_{mean} and Fz signal are

average (not selected). The characteristic has no obvious regular change trend when wear value increases. Therefore, the correlation between Fz_{mean} and wear value is not large.

Comparing other selected features with the unselected features also follows the same rules, which explain the validity of the GA feature selection to a certain extent.

4 Experimental results and analysis

The tool wear state recognition model takes the signal feature extracted in Section 3.3 as the input. The output is the corresponding tool wear state.

During the milling process, 108 sets of data are obtained. The tool wear process is presented in Fig. 6. Different wear stages are divided by tool wear values. The first 16 sets are the initial wear stage, the 17–87 sets are the normal wear stage, and the remaining 21 sets are the severe wear stage. Figure 7 illustrates two marked points where the tool wear rate changes from the previous data.

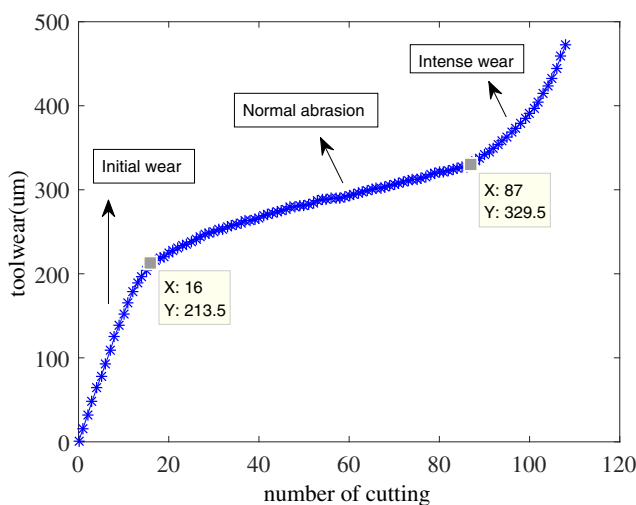


Fig. 7 Wear state division (108 sets)

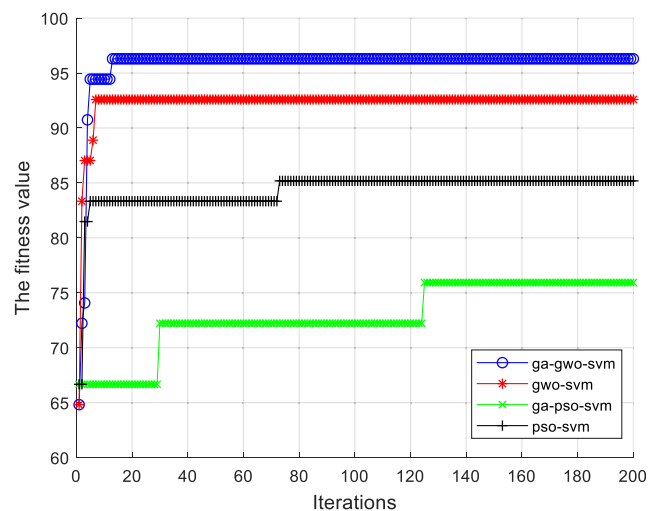


Fig. 8 Iterative process comparison

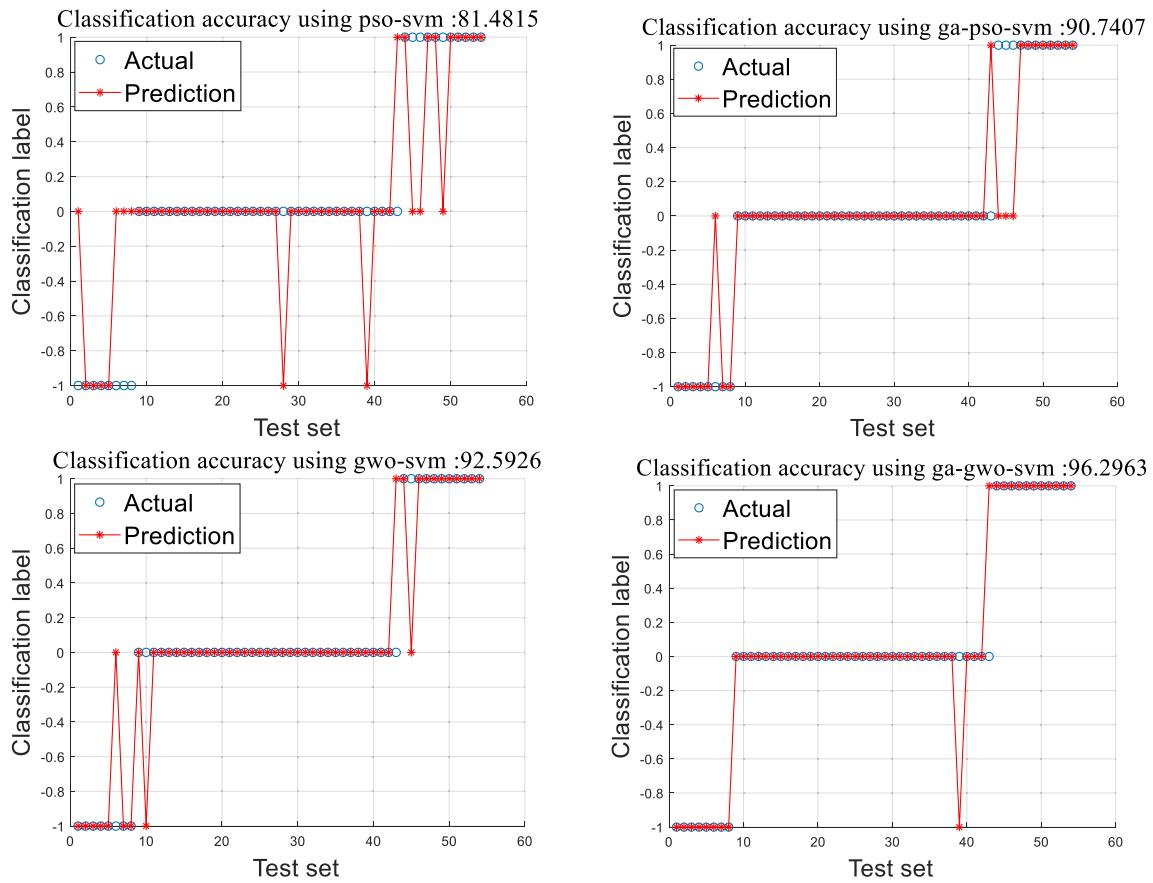


Fig. 9 Classification accuracy for the test set

The training and test sets, including their data, are made separate. To compare the effectiveness of the GA feature selection method used in this study (Section 3 of this paper) and the optimized SVM effects of PSO and GWO, we analyzed the optimization process of the four algorithms and the accuracy of the classification results, as presented in Figs. 8 and 9, respectively. The relevant parameters of PSO and GWO are set, as presented in Table 6.

In Fig. 8, PSO–SVM represents a classification algorithm that optimizes SVM parameters using PSO with all extracted features as inputs. GWO–SVM represents a classification algorithm that uses all extracted features as inputs and GWO to optimize SVM parameters. GA–PSO–SVM is a classification algorithm that optimizes SVM parameters by PSO and uses the features selected by GA as inputs. After GA feature selec-

tion, GA–GWO–SVM uses the features as inputs, and GWO optimizes the SVM parameters. The convergence speed and final fitness value of GA–GWO–SVM are better than those of GWO–SVM. However, the convergence speed of GA–PSO–SVM and the fitness value on the training set are worse than those of PSO–SVM. In addition, the optimization effect of the GWO algorithm is higher than that of PSO.

In Fig. 9, the classification accuracy of the four algorithms in the three wear stages of the test set is presented. The accuracy of the classification algorithm using GA feature selection is greater than that of the classification algorithm with all features as inputs, namely, GA–PSO–SVM > PSO–SVM, GA–GWO–SVM > GWO–SVM. GWO–SVM has a better effect than PSO–SVM and benefits from random walks and few parameters. The accuracy of PSO–SVM is 81.48%, that of GWO–SVM is 92.59%, and the accuracy of GA–PSO–SVM is 90.74%.

The classification accuracy of GA–GWO–SVM is 96.29%. Table 7 and Fig. 9 compare the classification effects of various algorithms on the test set; here, GS stands for grid search. SVM and GA–SVM divide the points of all test sets in the second category (normal abrasion), and the classification effect is extremely poor. Among other classification algorithms,

Table 6 Parameter settings of the optimization algorithms

	Iteration	Group	cu, gu	cl, gl	C1	C2	γ
GWO	200	20	100,1000	0.1,0.01	–	–	–
PSO	200	20	100,1000	0.1,0.01	1.5	1.7	0.1–1000

Table 7 Comparison of classification methods (force signals)

Method	Initial wear	Normal abrasion	Intense wear	Classification accuracy
SVM	0	100%	0	64.815%
GS-SVM	50.000%	91.43%	81.82%	83.333%
GA-GS-SVM	75.000%	97.140%	81.820%	90.741%
GA-SVM	0	100%	0	64.815%
GA-GA-SVM	62.500%	94.290%	81.820%	87.0370%
PSO-SVM	50.000%	91.429%	72.727%	81.482%
GA-PSO-SVM	87.500%	97.140%	72.727%	90.741%
GWO-SVM	87.500%	94.290%	90.900%	92.590%
GA-GWO-SVM	100%	94.290%	100%	96.296%

GA-GWO-SVM has the best classification effect, and the classification accuracy rates in the three wear stages are 100%, 94.29%, and 100%. The final total classification accuracy is 96.30%, which can result in the superior effect of online tool wear recognition.

5 Method validation with vibration signals

The milling dataset from Ref. [32] is implemented to validate the tool wear monitoring system. The data set has force, vibration, and current signals in the milling data set. The vibration signals are selected to verify the robustness of the proposed method. In the experiment, the stainless steel (HRC52) is machined by a two-flute ball nose cutter in a high speed milling machine (Röders Tech RFM760) with a spindle speed of up to 42,000 rpm. Three Kistler piezo accelerometers are mounted on the workpiece to measure the machine tool vibrations of the cutting process in the *x*, *y*, and *z* directions.

The 315 sets of tool wear values are divided into three stages, namely, Stage I, Stage II, and Stage III (Fig. 10). The

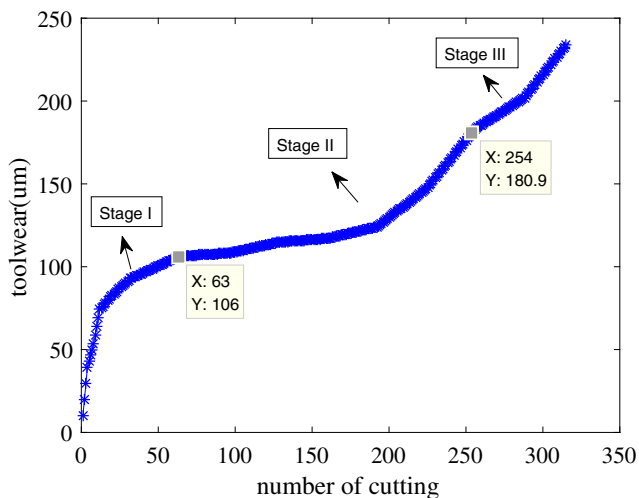


Fig. 10 Wear state division (315 sets)

tool wear rates differ in each stage. Based on the proposed tool wear monitoring system in this paper, the vibration signals are filtered by time domain, frequency domain, and Wavelet analysis, and the 84 features are obtained. Then 28 of the 84 features, which are show in Table 8, are selected by GA after running the feature selection program 30 times. Finally, the selected features are fed into GWO-SVM and other models to recognize tool wear states.

The results of tool wear states recognition based on vibration signals are presented in Table 9. It is seen from Table 9 that the classification accuracy of the models with feature selection by GA is higher than the models without feature selection, and the running time of models with feature is less than the models without feature selection except SVM model. SVM model saves time at the expense of accuracy because its parameters are not optimized. In addition, GA-GWO-SVM and GWO-SVM have the highest classification accuracy. Comprehensive consideration of classification accuracy and running time, the proposed GA-GWO-SVM outperforms other algorithms.

6 Conclusions

Efficiently and accurately monitoring tool wear state during machining provides guidance for the adjustment of machining parameters to ensure the stability of machining quality. In this study, an automated tool wear condition monitoring scheme for the milling process is developed using a cutting force sen-

Table 8 Selected features by GA (vibration signals)

Signal	Signal features					
<i>V_x</i>	\bar{x}	σ^2	x_r	x_{rms}	x_{peak}	VF
	C_S	S402	S414	S415	S416	
<i>V_y</i>	σ^2	x_{peak}	x_{rms}	x_r	K	C_S
	FC	MSF	S412	S415	S416	
<i>V_z</i>	σ^2	x_r	x_{rms}	x_{peak}	S	FC

Table 9 Comparison of classification methods (vibration signals)

Method	Classification accuracy	Running time (s)
GA–GWO–SVM	100%	13.736
GA–GS–SVM	99.3631%	7.327
GA–GA–SVM	98.7261%	24.560
GA–PSO–SVM	98.7261%	67.840
GWO–SVM	100%	22.705
GS–SVM	98.7261%	12.764
GA–SVM	98.0892%	31.151
PSO–SVM	98.0892%	111.171
SVM	97.4522%	2.052

sor. The scheme includes three modules, namely, feature extraction, feature selection, and prediction. Feature extraction is implemented by TD analysis, FD analysis, and wavelet packet decomposition to obtain comprehensive information. Feature selection employs GA to reduce the complexity of the prediction model. The SVM model optimized by GWO (i.e., GWO–SVM) for generalizability is proposed to predict tool wear state.

GA searches the correlation between signal and wear features and selects signal features that are highly correlated with wear features from the extracted features by TD analysis, FD analysis, and wavelet packet decomposition. GA greatly reduces the input dimension of the prediction model and improves the prediction accuracy. In Table 7, the comparison between the systems with and without feature selection based on GA indicates a 4% increase in system accuracy because of the use of GA. The results of the comparison of the proposed method and the SVM model optimized by certain common optimization techniques reveal the superiority of the GWO–SVM model (Table 7). Therefore, the GWO–SVM model outperforms other models (SVM, GA–GS–SVM, and GA–PSO–SVM) in terms of prediction accuracy and time.

Vibration signals taken from Ref. [32] are used to verify the robustness of the tool wear condition monitoring scheme. Table 9 shows that the proposed tool wear monitoring scheme maintains its superiority in prediction accuracy. Therefore, the proposed tool wear condition monitoring scheme based on GA–GWO–SVM technique is beneficial in improving recognition accuracy of tool wear state, and provides an effective guide for decision-making in the machining process.

Funding information This study was financially supported by the National Natural Science Foundation of China (No. 51665005), Innovation Project of Guangxi Graduate Education (YCBZ2017015), High-level Research Project of Qinzhou University (Grant no. 16PYSJ06), the Project of Guangxi Colleges and Universities Key Laboratory Breeding Base of Coastal Mechanical Equipment Design, Manufacturing and Control (GXLH2016ZD-06), and Guangxi Key Laboratory of manufacturing systems and advanced manufacturing technology.

References

- Ye Y, Wu M, Ren X, Zhou J, Li L (2018) Hole-like surface morphologies on the stainless steel surface through laser surface texturing underwater. *Appl Surf Sci* 462:847–855
- Lauro CH, Brandão LC, Baldo D, Reis RA, Davim JP (2014) Monitoring and processing signal applied in machining processes—a review. *Measurement* 58:73–86
- Wang G, Yang Y, Xie Q, Zhang Y (2014) Force based tool wear monitoring system for milling process based on relevance vector machine. *Adv Eng Softw* 71:46–51
- Zhu K, Liu T (2018) Online tool wear monitoring via hidden semi-Markov model with dependent durations. *IEEE Trans Ind Inf* 14(1): 69–78
- Pai PS, Mello GD (2015) Vibration signal analysis for monitoring tool wear in high speed turning of Ti-6Al-4V. *JEMS* 22(6):652–660
- Xie Z, Li J, Lu Y (2018) An integrated wireless vibration sensing tool holder for milling tool condition monitoring. *Int J Adv Manuf Technol* 95(5–8):2885–2896
- Duspara M, Sabo K, Stoic A (2014) Acoustic emission as tool wear monitoring. *Tehnicki Vjesnik-Technical Gazette* 21(5):1097–1101
- Salgado DR, Cambero I, Herrera JM, García Sanz-Calcedo J, García Sanz-Calcedo AG, Núñez López PJ, García Plaza E (2014, July) A tool wear monitoring system for steel and aluminium alloys based on the same sensor signals and decision strategy. *Int Mater Sci Forum* 797:17–22
- Kong D, Chen Y, Li N, Tan S (2017) Tool wear monitoring based on kernel principal component analysis and v-support vector regression. *Int J Adv Manuf Technol* 89(1–4):175–190
- Kong D, Chen Y, Li N (2017) Force-based tool wear estimation for milling process using Gaussian mixture hidden Markov models. *Int J Adv Manuf Technol* 92(5–8):2853–2865
- Kothuru A, Nooka SP, Liu R (2018) Application of audible sound signals for tool wear monitoring using machine learning techniques in end milling. *Int J Adv Manuf Technol* 95(9–12):3797–3808
- Zhang M, Liu H, Li B (2014) Face milling tool wear condition monitoring based on wavelet transform and Shannon entropy. *Appl Mech Mater* 541–542:1419–1423
- Aghazadeh F, Tahan A, Thomas M (2018) Tool condition monitoring using spectral subtraction and convolutional neural networks in milling process. *Int J Adv Manuf Technol* 98(9–12):3217–3227
- Chandrashekar G, Sahin F (2014) A survey on feature selection methods. *Comput Electr Eng* 40(1):16–28
- Zhang B, Katinas C, Shin YC (2018) Robust tool wear monitoring using systematic feature selection in turning processes with consideration of uncertainties. *J Manuf Sci Eng* 140(8):081010
- Stein G, Chen B, Wu AS, & Hua KA (2005). Decision tree classifier for network intrusion detection with GA-based feature selection. In *Proceedings of the 43rd annual southeast regional conference-volume 2 ACM*. pp. 136–141
- Liao TW (2010) Feature extraction and selection from acoustic emission signals with an application in grinding wheel condition monitoring. *Eng Appl Artif Intell* 23(1):74–84
- Beasley D, Bull DR, Martin RR (1993) An overview of genetic algorithms: part 1, fundamentals. *Univ Comput* 15(2):56–69
- Kaya B, Oysu C, Ertunc HM (2011) Force-torque based on-line tool wear estimation system for CNC milling of Inconel 718 using neural networks. *Adv Eng Softw* 42:76–84
- Klaic M, Murat Z, Staroveski T, Brezak D (2018) Tool wear monitoring in rock drilling applications using vibration signals. *Wear* 408–409:222–227
- Kurek J, Kruk M, Osowski S, Hoser P, Wiczorek G, Jegorowa A, Górski J, Wilkowski J, Śmietańska K, Kossakowska J (2016) Developing automatic recognition system of drill wear in standard

- laminated chipboard drilling process. *Bull Polish Acad Sci Tech Sci* 64(3):633–640
22. Wang GF, Yang YW, Zhang YC, Xie QL (2014) Vibration sensor based tool condition monitoring using ν support vector machine and locality preserving projection. *Sensors Actuators A Phys* 209: 24–32
 23. Pandiyan V, Caesarendra W, Tjahjowidodo T, Tan HH (2018) In-process tool condition monitoring in compliant abrasive belt grinding process using support vector machine and genetic algorithm. *J Manuf Process* 31:199–213
 24. Aich U, Banerjee S (2014) Modeling of EDM responses by support vector machine regression with parameters selected by particle swarm optimization. *Appl Math Model* 38(11–12):2800–2818
 25. Tao X, & Tao W (2010). Cutting tool wear identification based on wavelet package and SVM. In 2010 8th world congress on intelligent control and automation (WCICA), IEEE. 5953–5957
 26. Nie P, Xu H, Liu Y, Liu X, & Li Z (2011). Aviation tool wear states identifying based on EMD and SVM. In 2011 second international conference on digital manufacturing and automation (ICDMA), IEEE. 246–249
 27. García-Nieto PJ, García-Gonzalo E, Vilán JV, Robleda AS (2016) A new predictive model based on the PSO-optimized support vector machine approach for predicting the milling tool wear from milling runs experimental data. *Int J Adv Manuf Technol* 86(1–4): 769–780
 28. Jia L, Jian-ming Z, Xiao-Jing B, & Lei WL (2011). Research on GA-SVM tool wear monitoring method using HHT characteristics of drilling noise signals. In 2011 international conference on consumer electronics, communications and networks (CECNet), IEEE. 635–638
 29. Zhang KF, Yuan HQ, Nie P (2015) A method for tool condition monitoring based on sensor fusion. *J Intell Manuf* 26(5):1011–1026
 30. Zhang X, Chen W, Wang B, Chen X (2015) Intelligent fault diagnosis of rotating machinery using support vector machine with ant colony algorithm for synchronous feature selection and parameter optimization. *Neurocomputing* 167:260–279
 31. Zhou Z, Zhang R, Wang Y, Zhu Z, Zhang J (2018) Color difference classification based on optimization support vector machine of improved grey wolf algorithm. *Optik* 170:17–29
 32. Li X, Lim BS, Zhou JH, Huang S, Phua SJ, Shaw KC, & Er MJ. (2009). Fuzzy neural network modelling for tool wear estimation in dry milling operation. In Annual conference of the prognostics and health management society. 1–11
 33. Cortes C, Vapnik V (1995) Support-vector networks. *Mach Learn* 20(3):273–297
 34. Smola AJ, Schölkopf B (2004) A tutorial on support vector regression. *Stat Comput* 14(3):199–222
 35. Mirjalili S, Mirjalili SM, Lewis A (2014) Grey wolf optimizer. *Adv Eng Softw* 69:46–61

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.