ORIGINAL ARTICLE

Tool wear prediction method based on bidirectional long short-term memory neural network of single crystal silicon micro-grinding

Chengxi She[1,2] · Kexin Li[1] · Yinghui Ren[1] ⬤ · Wei Li[1] · Kun Shao[3]

Abstract
Although many efforts have been made on off-line tool wear prediction, the on-line intelligent prediction of tool wear prediction based on indeterminate factor relationship has not been addressed. In this paper, a tool wear prediction model is built based on bidirectional long short-term memory neural network (BiLSTM) to deal with these challenges. The acoustic emission (AE) signal and tool wear image are selected as indicators to characterize the wear behavior of micro-grinding tool. The BiLSTM model is constructed with the input of 4-dimensional feature vector, which is composed of medium frequency energy ratio of the AE signal, initial tool cross-sectional area, micro-grinding depth and micro-grinding length, and the output of the loss of tool cross-sectional area. Two derived models of genetic algorithm-optimized BiLSTM (GA-BiLSTM) and long short-term memory neural network (LSTM) are developed to compare the accuracy of the BiLSTM model. Two machine learning models of back propagation neural network (BP) and algorithm optimized BP neural network (GA-BP) are developed to compare the stability and superiority of the BiLSTM model. The micro-grinding experiment is conducted by the electroplated diamond micro-grinding tool to verify the feasibility using the proposed methods and the results show that the average prediction accuracy of the BiLSTM model reached 92.08% while the accuracies of other models from GA-BiLSTM to GA-BP are separately 87.2%, 88.6%, 84.4%, and 85.8%. The BiLSTM model provides a novel wear characterization and prediction method that combines AE signals and visual images using small-sample and multi-sourced heterogeneous data. It undoubtedly promotes sustainable manufacturing and provide theoretical basis for independent decision-making in precision intelligent manufacturing.

Keywords Micro-grinding · Tool wear prediction · Neural network · BiLSTM · Small sample · Multi-sourced heterogeneous data

1 Introduction

The wear of micro-grinding tool is a crucial factor that affects the processing efficiency, quality, and dimensional accuracy of micro-structure workpieces [1]. The tool wear prediction can effectively evaluate the tool service life and then guide tool replacement strategies to reduce cost and control quality [2, 3]. According to previous reports, the state signals generated during the machining process, such as cutting force signals, vibration signals, acoustic emission (AE) signals, and so on, or tool static geometry and morphology characteristic such as tool diameter and radius of cutting edge can be used to predict the tool wear by theoretical physical model or empirical model based on determinate factor relationship [4]. However, it appears challenging to achieve the micro-grinding tool wear prediction online integrating the processing state signals and tool static characteristic based on indeterminate factors relationship. These processing state signals and tool static characteristics exhibit multi-source heterogeneous data type. And collecting signal data during the machining process requires a large amount of cost [5, 6]. Therefore, it is more significant to predict the tool wear online using small-sample and multi-sourced heterogeneous data with the development of artificial intelligence technology.

✉ Yinghui Ren
  rebecca_ryh@hnu.edu.cn

[1] College of Mechanical and Vehicle Engineering, Hunan University, Changsha 410082, China

[2] School of Management & Engineering, Nanjing University, Nanjing 210093, China

[3] AVIC Manufacturing Technology Institute, Beijing 100024, China

The characterization of tool wear is essentially a problem of pattern state recognition, including acquiring the feature signals during the machining process, transforming the signals into dimensionality reduction, extracting the feature information that has a certain mapping relationship with the tool wear state, and forming feature vectors [7]. The characterization of tool wear can be mainly divided into direct and indirect ways. The direct way is to directly measure and observe the geometric shape, size, and morphological characteristics of the tool using machine vision or optical sensors. Banda et al. [8] illustrated the application and the significance of machine vision systems in tool wear monitoring and tool performance optimization. Zhang et al. [9] established an in-situ monitoring system based on machine vision to detect tool wear behavior. The indirect way is to detect the status information generated in the machining process, such as force, acceleration, and AE signals [10]. Zhou et al. [11] demonstrated the basic principles and key technologies of AE signal used for monitoring in sawing. Wang et al. [12] proposed an effectively AE signal processing method to monitor the damage mechanisms of monocrystalline silicon in ultra-precision grinding. Zhou et al. [13] presented a monitoring method to predict the tool wear using force and acceleration signals and it is effective in tool wear prediction with maximum error around 8 μm.

The prediction methods of tool wear proposed can be divided into two categories, which are physical model driven methods and data-driven methods [14, 15]. The physical-model-driven method is to build an explicit mathematical model for tool wear prediction through physical laws or mathematical models [16]. Li et al. [17] presented a new physics-informed meta-learning framework to predict tool wear under varying wear rates and validated the effectiveness of the method presented through a milling machining experiment. Qiang et al. [18] proposed a physics-informed transfer learning (PITL) framework to predict tool wear under variable working conditions and the experiment results showed highly prediction accuracy under variable conditions. The data-driven method is to collect tool wear status information through a variety of sensors and then use machine learning or deep learning methods to build a prediction model using the data after information processing and recognition [19]. Tool wear prediction is an important research field in intelligent manufacturing. Data-driven method is recognized as an effective means to achieve accurate prediction of tool wear, which provides a significant research direction for on-line tool wear monitoring problems [20, 21]. Wei et al. [22] established a genetic algorithm (GA-BP) neural network prediction model to predict the tool wear condition during high-speed milling and the relative error is within 5%. Chen et al. [23] proposed a tool wear prediction method based on back propagation neural network and compared the results with a long short-term memory (LSTM) model. The results

show that the error was reduced by 29% and 25% as the input eigenvalues increased, respectively.

Bidirectional long short-term memory neural network (BiLSTM) is one of the data-driven tool wear prediction methods. As a derivative model of recurrent neural network (RNN), it inherits the advantages of RNN that can effectively represent the characteristics of time series data and avoids the problem that RNN is prone to produce gradient disappearance [24]. Tool wear is accumulated with the processing time, so it is necessary to mine the deep time series characteristics of data for tool wear prediction. Wu et al. [25] proposed a tool wear prediction model based on singular value decomposition and BiLSTM with cutting force signal. Ma et al. [26] established a wear prediction model based on convolutional BiLSTM and convolutional bidirectional gated recurrent unit using cutting force signals.

It is challenging to integrate multi-source heterogeneous data to characterize wear characteristics and predict micro-grinding wear, such as multi-source heterogeneous data fusion, mixed prediction model construction, and small sample cost prediction. To deal with the challenge, a novel prediction method based on BiLSTM is proposed to characterize and predict the wear combining AE signals and visual images using such data. The micro-channel grinding experiments are conducted by the electroplated diamond micro-grinding tool. The prediction results of BiLSTM model are compared with the experimental test results to verify the feasibility of BiSLTM model. Two derived models of GA-BiLSTM and LSTM are compared to verify the accuracy of BiLSTM model. Two machine learning models of BP and GA-BP are compared to verify the stability and superiority of BiLSTM model. Several abbreviations and notations are used in the paper and are described separately in Table 1.

## 2 Tool wear prediction model

2.1 Principle of signal characterization

The state of cutting tools in actual machining process is influenced by various factors. The machine vision technology is widely used to observe the morphology and features of micro-grinding tool nowadays, which has the characteristics of high accuracy and convenience [27]. Moreover, the morphology of micro-grinding tool is shown in Fig. 1a. According to the wear form of micro-grinding tools and the changes in morphology, the wear situation can be directly characterized by the loss of the cross-sectional area of the micro-grinding head. The loss of the cross-sectional area of the micro-grinding head refers to the difference in the cross-sectional area of the grinding head before and after grinding, as shown in Fig. 1b. The calculation formula is as shown in Eq. (1).

Table 1 List of used abbreviations and notations

Abbreviation/notation	Meaning of abbreviation/notation
BiLSTM	Bidirectional long short-term memory neural network
LSTM	Long short-term memory neural network
GA-BiLSTM	Genetic algorithm optimized bidirectional long short-term memory neural network
BP	Back propagation neural network
GA-BP	Genetic algorithm optimized back propagation neural network
AE	Acoustic emission
PITL	Physics-informed transfer learning framework
RNN	Recurrent neural network
C_0	The edge contour before micro grinding
C_i	The edge contour after machining
i	The number of machining channel micro grooves
φ_i	The cross-sectional area loss of the grinding head
$h^{(t)}$	The output of the network at time t
$X^{(t)}$	The input of the network at time t
$\boldsymbol{\omega}$	The corresponding parameters of network nodes
$g^{(t)}$	The output network calculated at time t
$\boldsymbol{W}_f, \boldsymbol{W}_i, \boldsymbol{W}_C, \boldsymbol{W}_O$ $\boldsymbol{b}_f, \boldsymbol{b}_i, \boldsymbol{b}_C, \boldsymbol{b}_O$	The parameters in LSTM structure
n	The number of hidden neurons
T	The length of real-time sequential data
N	The sample size
X_k	The input in group k of BiLSTM model
Y_k	The output in group k of BiLSTM model
l	The tool hanging length
v	The spindle speed
f	The feed rate
L	The micro groove length
d	The micro groove depth

Fig. 1 Wear characterization method for micro-grinding tool. **a** The morphology of micro-grinding tool. **b** The measuring method for the loss of the cross-sectional area
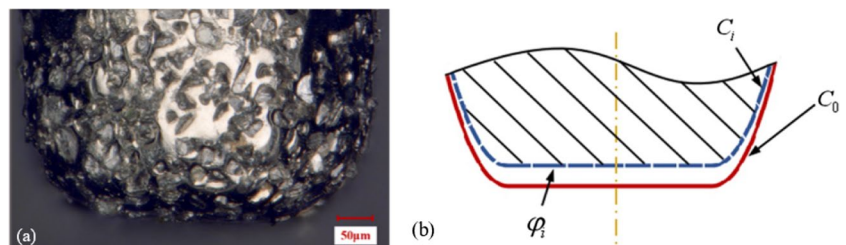
$$\varphi_i = C_0 - C_i \tag{1}$$

where the C_0 represents the edge contour before micro-grinding. The C_i represents the edge contour after machining i channel micro-grooves. The φ_i represents the cross-sectional area loss of the grinding head after machining i channel micro-grooves. The coordinate unit is μm^2.

Another way to characterize the wear of micro-grinding tool is AE signal. The AE signals during the grinding process originate from the interaction between abrasive particles and materials and the crystal structure during material removal, which rapidly release energy, generating transient elastic waves that manifest as AE signals [28]. Due to the differences in AE sources, signal time and frequency domain analysis can be used to distinguish different material removal behaviors. The process of AE detection is shown in Fig. 2. Firstly, the AE sensor is attached to the unmachined surface, which is used to convert the AE signal generated during the processing into a voltage signal. Secondly, the voltage signal can be amplified, filtered, and A/D converted through a preamplifier. Therefore, the AE signal can be quantitatively analyzed by the data acquisition system to derive the actual
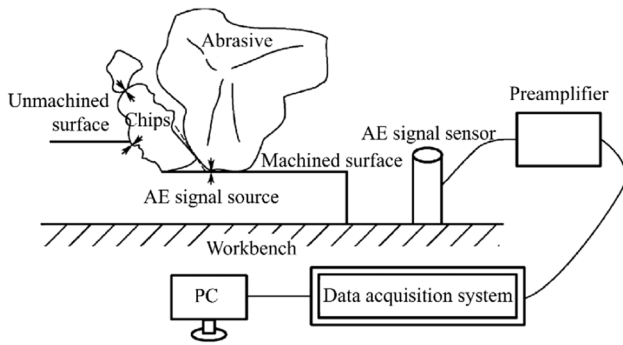
Fig. 2 The process of AE signal detection [30]

processing status based on the relevant feature parameters. It was found that there is a certain mapping relationship between the characteristics of AE signals and the external processing environment and the inherent properties of the workpiece itself [29]. However, the AE signals collected are influenced by external environment and the inherent characteristics of the machine tool, and the signal energy fluctuates up and down through spectrum analysis [30]. Therefore, it is more reasonable for characterizing the wear of micro-grinding tool using energy ratio instead of energy value.

The experimental group with a micro-groove depth of 200 μm was taken as an example. The signal time domain characteristics are extracted by root mean square method and the frequency bands are divided by 4-layer wavelet packet decomposition and fast Fourier transform. The signal in the sampling interval is reconstructed through these signal processing methods, and the energy proportion diagram was drawn for each frequency band through energy proportion calculation, as shown in Fig. 3a. It can be seen that the signal energy in the sampling interval mainly comes from the low

and medium frequency bands and they show a negative correlation relationship. With the grinding length accumulating, the total amount of material removal increases, the proportion of medium frequency band energy continuously increases and tends to stabilize, and the energy of low frequency band continuously decreases and tends to stabilize. Compare the proportion of medium frequency band energy with the corresponding radial wear of the micro-grinding tool, as shown in Fig. 3b. There is a positive correlation between the proportion of medium frequency band energy and the corresponding radial wear of the micro-grinding tool. Therefore, the proportion of medium frequency band energy is the most suitable characterization of micro-grinding tool wear.

2.2 Model design

Long short-term memory neural network (LSTM) is an improved model of RNN, which has the same input and output as RNN [31]. LSTM has the ability to extract temporal features and expand them according to time dimensions [32], as shown in Eq. (2).

$$h^{(t)} = f\left(h^{(t-1)}, X^{(t)}; \boldsymbol{\omega}\right) = f\left(f\left(h^{(t-2)}, X^{(t-1)}; \omega\right), X^{(t)}; \boldsymbol{\omega}\right)$$
$$= \cdots = g^{(t)}\left(X^{(t)}, X^{(t-1)}, X^{(t-2)}, \cdots, X^{(2)}, X^{(1)}\right) \quad (2)$$

where the $h^{(t)}$ represents the output of the network at time t, which corresponds to the signal features extracted at time t. The $X^{(t)}$ represents the input of the network at time t, which corresponds to the signal data collected at the time t. The $\boldsymbol{\omega}$ Represents the corresponding parameters of network nodes. The function $g^{(t)}$ represents the output network calculated at time t based on all historical time series signals.

As shown in Fig. 4, the structure of a single LSTM unit body includes a forgetting gate, an input gate, and an
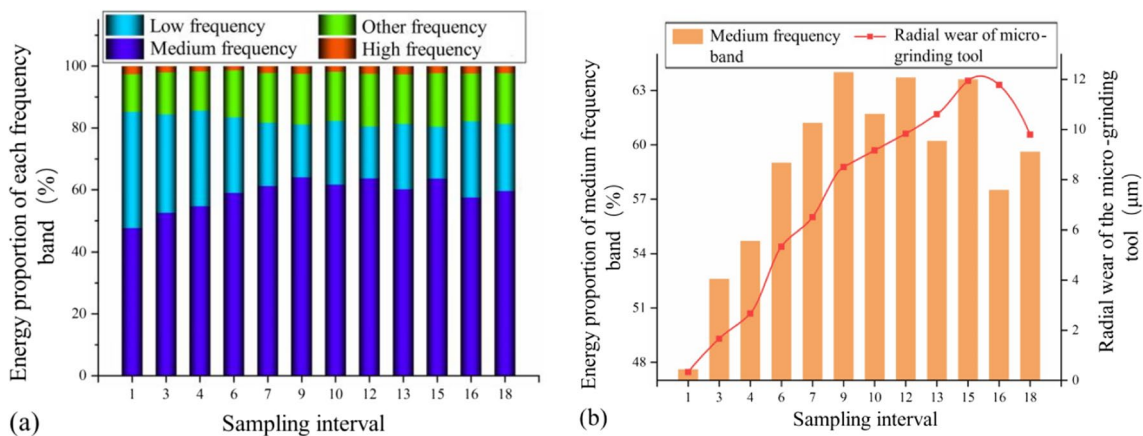

Fig. 3 Signal energy analysis in the sampling interval. **a** The energy proportion diagram for each frequency band. **b** The proportion of medium frequency band energy with the corresponding radial wear of the micro-grinding tool
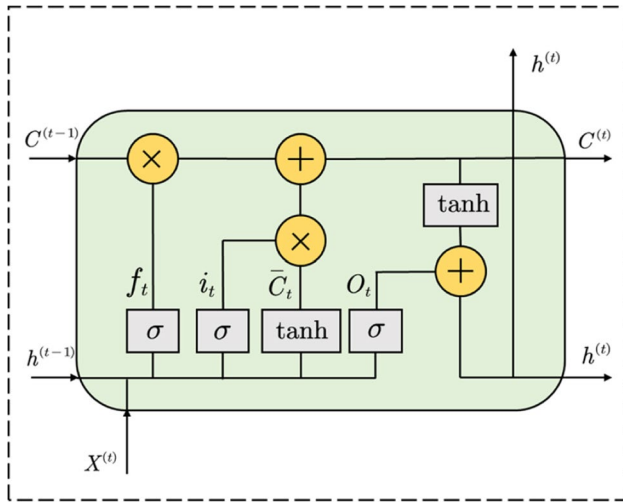
Fig. 4 Schematic diagram of LSTM unit structure

output gate. The input of the unit body includes the input of the current time and the output of the previous time unit body and the output of the unit body serves as the input of the next unit body. The specific calculation process is as shown in Eq. (3).

$$
\begin{cases}
f_t = \sigma\left(W_f\left(h_{t-1}, X^t\right) + b_f\right) \\
i_t = \sigma\left(W_i\left(h_{t-1}, X^t\right) + b_i\right) \\
\overline{C}_t = \tanh\left(W_C\left(h_{t-1}, X^{(t)}\right) + b_C\right) \\
C^{(t)} = f_t \odot C^{(t-1)} + i_t \odot \overline{C}_t \\
O_t = \sigma\left(W_O\left(h_{t-1}, X^{(t)}\right) + b_O\right) \\
h^{(t)} = O_t \odot \tanh\left(C^{(t)}\right)
\end{cases}
\tag{3}
$$

where the state of the unit $C^{(t)}$ is obtained by updating the forgetting gate f_t, input gate i, output gate O_t, the previous hidden layer state $h^{(t-1)}$, and the previous unit state $C^{(t-1)}$ at the moment of t. The hidden layer state in the unit $h^{(t)}$ is updated based on the signal data input at the moment of t and the unit state $C^{(t)}$. The parameters W_f, W_i, W_C, W_O and b_f, b_i, b_C, b_O are obtained through model training and shared at all times. The n represents the number of hidden neurons. The T represents the time step, which is the length of real-time sequential data. The "$\odot$" represents the accumulation by element. The σ represents sigmoid activation function and the *tanh* represents the tanh activation function.

In response to the problem of insufficient early feature memory ability and difficulty in deep mining the temporal features of the entire network in LSTM when learning temporal features, BiLSTM is proposed to enhance the representation ability of data features and improve the accuracy of wear prediction. BiLSTM consists of two independent LSTM layers, which input temporal feature data in forward and reverse directions for feature extraction. The two
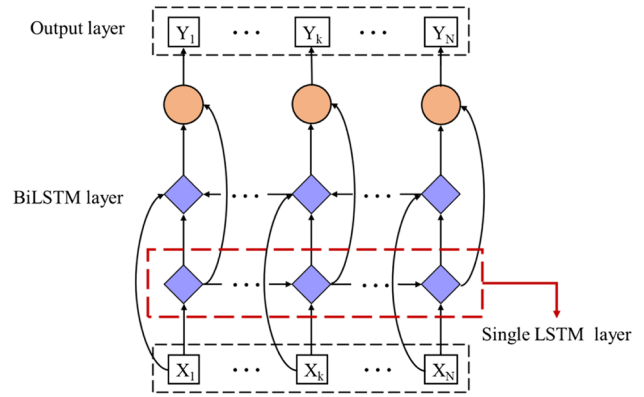

Fig. 5 Schematic diagram of BiLSTM unit structure

output feature vectors are concatenated as the final output to achieve feature representation. The parameters of the two LSTM layers are independent of each other and the specific network structure is as shown in Fig. 5. In the prediction model, $N = 90$, which represents the sample size. X_k represents that the input in group k is a 4-dimensional feature vector composed of processing variables including micro-groove depth (μm) and micro-groove length (mm), AE signals including the proportion of medium frequency band energy (%) and visual images including the cross-sectional area of the micro-grinding head (μm²). Y_k represents that the output in group k is the loss of cross-sectional area of the micro-grinding head (μm²).

In order to reduce the interference of singular sample data on training, accelerate the training and convergence speed of the model, and improve the accuracy of the model, normalization process is required to limit the data to the range of [0,1]. The prediction model uses Adam algorithm to minimize the loss function, as shown in Eq. (4).

$$
E_{loss} = \frac{1}{n}\sum_{k-1}^{n}\left(y_k^{pre} - y_k\right)^2
\tag{4}
$$

In Eq. (4), y_k^{pre} represents the predicted wear value, y_k represents the actual wear value, n represents the data volume of training samples, and E_{loss} represents the loss of function. Evaluate the performance of the model by calculating the root mean squared error (RMSE) and determination coefficient R^2, as shown in Eqs. (5) and (6). The hyperparameters are selected according to the RMSE after training and the final optimal hyperparameters are shown in Table 2.

$$
RMSE = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(y_k^{pre} - y_k)^2}
\tag{5}
$$

Table 2 The final optimal hyperparameters of BiLSTM

Hyperparameters	Value
NumHiddenUnits	110
MaxEpochs	2750
InputSize	4
FullyConnectedLayer	1
LearnRate	0.01

$$R^2 = \sum_{k-1}^{n} \left(y_k^{pre} - \bar{y}\right)^2 / \sum_{k-1}^{n} \left(yk - y_k^{pre}\right)^2 \qquad (6)$$

2.3 Comparative model

The prediction performance of the micro-grinding wear prediction model is related to the fitness between the depth of the model architecture and the sample data volume [33]. To prove the accuracy of the BiLSTM model by comparing the prediction accuracy of different models with different depth of model architecture, two derived models of GA-BiLSTM and LSTM are constructed. To increase the ability of feature signal extraction and representation, GA-BiLSTM introduces more algorithm structures to deepen the model architecture. To adapt to feature extraction and representation of smaller sample data, LSTM applies shallower network structures to simplify the model architecture. The prediction results are compared with the BiLSTM model.

In data-driven tool wear prediction methods, traditional machine learning methods often show better results in small sample data prediction than deep learning methods [34]. In traditional machine learning algorithms, BP has good adaptive feature extraction ability and nonlinear mapping ability [35]. Therefore, two wear prediction models based on BP and GA-BP were constructed for prediction and comparison. The BP structure is as shown in Fig. 6. The ω_{ij} is the weight value of the connection between the input layer unit and the

hidden layer unit. The ν_j is the weight value of the connection between the hidden layer unit and the output layer unit. The specific calculation process is as shown in Eq. (7).

$$\begin{cases} H_j(k) = f\left(\sum_{i=1}^{m} \left(\omega_{ij}\mathcal{X}_i(k) - a_j\right)\right), j = 1, 2, \cdots h \\ Y(k) = f\left(\sum_{j-1}^{h} \left(v_j h_j(k) - b\right)\right) \\ \varepsilon(k) = Y(k)(1 - Y(k))\left(Y(k) - T(k)\right) \\ \delta_j(k) = Y(k)\left(1 - Y(k)\right)\varepsilon(k)v_j \end{cases} \qquad (7)$$

where the a_j is the threshold between input layer and hidden layer units and the b is the threshold between hidden layer and output layer units. The $x_i(k)$ represents the input of the model. The $Y(k)$ represents the predicted output value. The $T(k)$ represents the expected output value.
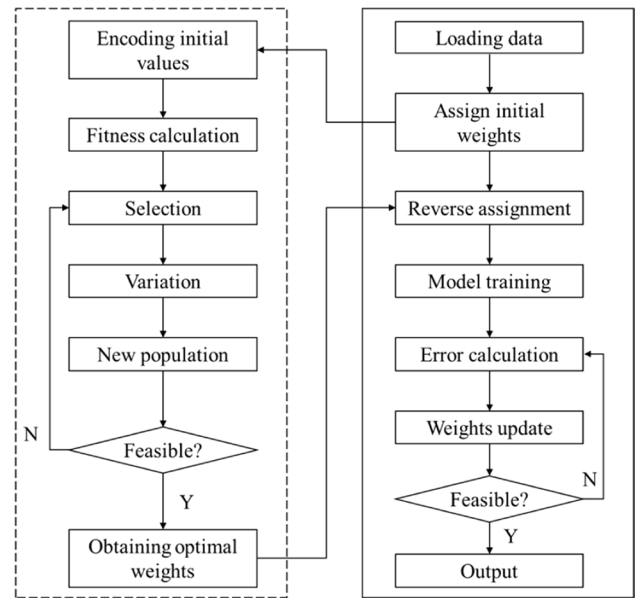

Fig. 7 Schematic flowchart of GA-BP structure

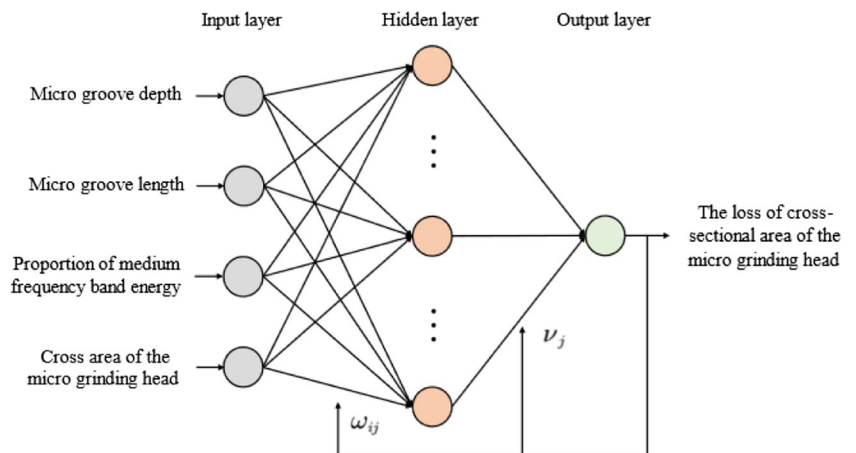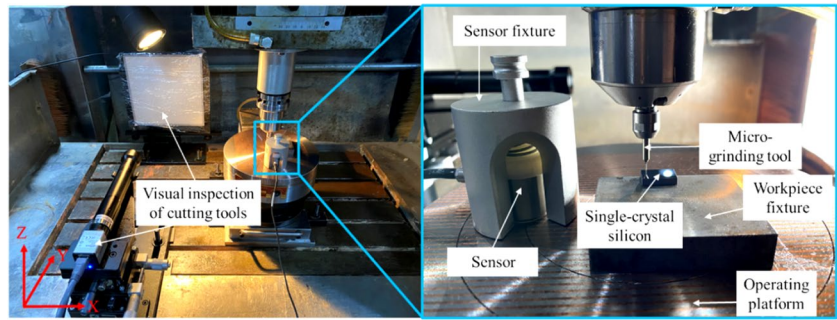
Fig. 6 Schematic diagram of BP structure

Fig. 8 The micro-grinding platform

The schematic flowchart of GA-BP structure is as shown in Fig. 7. After assigning the initial weight value of the connection to the model, genetic algorithm is used for fitness calculation, selection, and variation. When the iteratively updated weight value reaches feasible, the optimal weight value is reversely assigned to the model to achieve model optimization.

3 Experiment verification

3.1 Design of micro-grinding experiment

The micro-grinding experiment is conducted by the electroplated diamond micro-grinding tool to verify the feasibility and accuracy of the models proposed. The micro-grinding platform used in the experiment is modified from the MK2945C CNC coordinate grinder, equipped with a micro-grinding tool online detection system and AE detection system, as shown in Fig. 8. The micro-grinding tool online detection system includes camera with a resolution of 20 million pixels named Dahua A3B00MG000 CMOS, fixed magnification fixed focal distance telephoto lens named WWK40-100-111 and LED backlight board with model named CF-200-W. The AE detection system includes high sensitivity resonant AE sensor with model named R6α, preamplifier model named 2/4/6 with an amplification gain of 40 dB, and PCI-2 AE acquisition system.

Reasonable processing parameters can effectively reduce the risk of the processing process, which is conducive to

Table 3 The parameters in experiment

Parameters	Value
Tool hanging length (l, mm)	15
Spindle speed (v, r/min)	60,000
Feed rate (f, mm/min)	1
Micro-groove length (L, mm)	8
Micro-groove depth (d, μm)	50, 80, 100, 150, 200

detect feature signals and study their characterization relationships. The parameters used in the micro-grinding experiment are shown in Table 3. A total of six micro-grooves are processed with three sampling intervals set at the beginning, middle, and end of each micro groove, as shown in Fig. 9. The lengths corresponding to the head of the micro-groove are, respectively, 2 mm, 4 mm, and 6 mm. The experimental group with the same micro-groove depth corresponds to 18 sampling intervals, which is 18 sample data groups and five experimental groups contain a total of 90 sampling intervals, which is 90 sample data groups. The first 80% of the 90 data groups are divided as training sets to train the model and the last 20% are divided as test sets to test the predictive performance of the model.

3.2 Results of prediction models

The test data is loaded into the BiLSTM, GA-BiLSTM, LSTM, BP, and GA-BP models and the test results of prediction models are shown in Fig. 10a–e, respectively. In
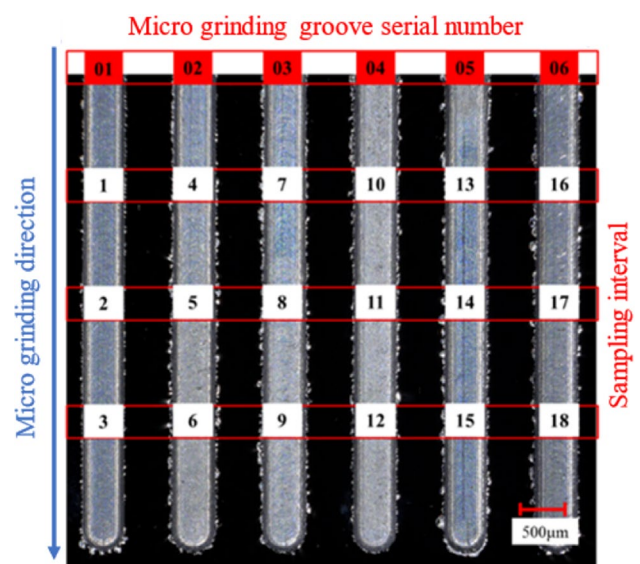

Fig. 9 The schematic diagram of sampling intervals for AE signals
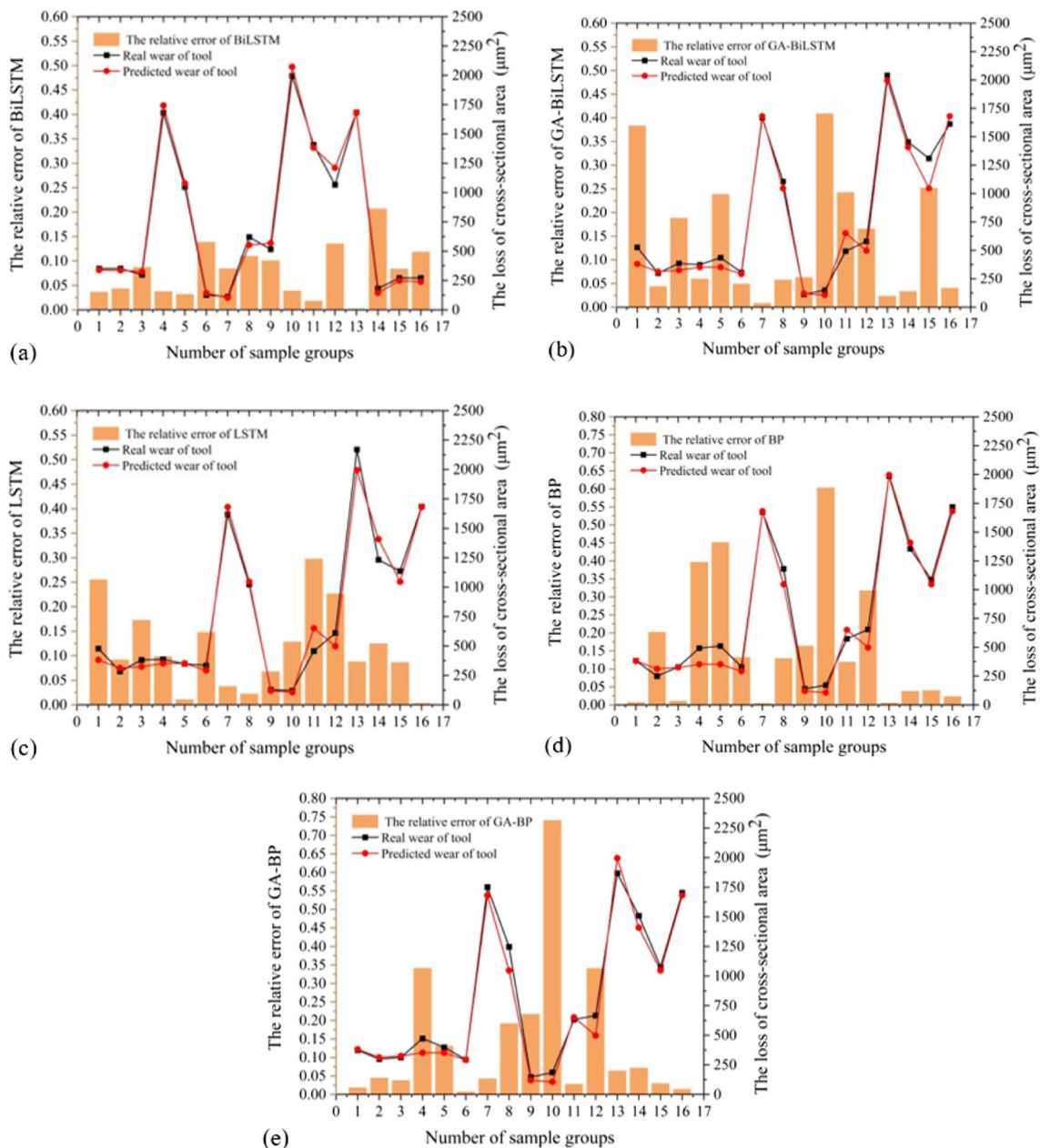
Fig. 10 The test results of prediction models. **a** BiLSTM; **b** GA-BiLSTM; **c** LSTM; **d** BP; **e** GA-BP

Fig. 10a, the predicted value of BiLSTM model is basically consistent with the real wear value. Most data groups show high prediction accuracy. Nearly half of the data prediction relative error is within 5% and only relative error of group 14 is about 20%. By averaging the relative errors of all groups, the relative error predicted by the model is 7.92%, indicating the prediction accuracy of BiLSTM model is 92.08%. In Fig. 10b, the prediction trend of GA-BiLSTM prediction model is in good agreement with the actual change trend. However, prediction results in some data groups have significant deviations, such as the relative errors of the 1st, 5th,

10th, 11th, and 15th groups exceeding 20%. By averaging the relative errors of all groups, the relative error predicted by the model is 12.8%, which means the prediction accuracy of the model is 87.2%. In Fig. 10c, the relative error predicted of LSTM model is 11.42% by averaging the relative errors of all groups, which means the prediction accuracy of the model is 88.58%.

In Fig. 10d, the prediction trend of BP prediction model is in good agreement with the actual change trend. The predicted values of the model and the actual wear value are in good agreement when the actual wear value is large, but there

is a significant error when the actual wear value is low. For example, the deviation between the predicted value and the actual value of the 4th, 5th, 10th, and 12th groups exceeds 30%, and even the relative prediction error of the 10th group is as high as 60.25%. By averaging the relative errors of all sample data, the relative error predicted by the model is 15.63%, which means the prediction accuracy of the model is 84.37%. In Fig. 10e, the prediction trend of the GA-BP neural network micro-grinding tool wear prediction model is in good agreement with the actual change trend. However, some data prediction results have significant deviations, such as the relative errors of the 4th, 10th, and 12th groups exceeding 30%, and the prediction result of the 10th group of data is even as high as 74.03%, which is consistent with the sample data group corresponding to the large error of the BP neural network. The relative error of the model prediction is 14.2%, which means the prediction accuracy of the model is 85.8%.

3.3 Comparison and analysis

In order to more intuitively compare the differences in predictive performance between models, the comparison figures of relative error for prediction models (BiLSTM, GA-BiLSTM, LSTM, BP, and GA-BP models) are plotted to compare the prediction accuracy of different models with different depths of model architecture. In order to compare differences of the prediction accuracy among five models more intuitively, the prediction results were plotted every four groups as a rank and a total of four relative error comparison figures (Fig. 11a–d) were drawn.

BiLSTM model has a lower relative error in the predicted values corresponding to most samples and the relative error fluctuation is smaller than the other two models. GA BiLSTM and LSTM models have relatively high relative errors in 1st, 5th, 10th, 11th, 12th, and 15th groups, while the BiLSTM model has relatively high relative error in 14th group. The relatively large prediction error in the 1st group of samples is due to the special memory structure of the model (Fig. 11a). As the output of the current unit will affect the output of the next unit, while the 1st group of samples is not affected by this, which means that they do not have memory ability and cannot mine temporal features of the data, there is a certain error. The large prediction errors of the remaining samples are caused by the randomness of the deep learning model. In addition, it can be seen that the prediction errors of BP neural network and GA-BP neural network fluctuate greatly, showing extremely high prediction accuracy in some samples and significant mismeasurement errors in some samples, especially the relative error in 10th group (Fig. 11c) is close to 80%.

Although, the BiLSTM model do not achieve extremely high prediction accuracy in individual samples, the overall relative error fluctuation is small and stable and the average prediction accuracy is higher than other four models. In conclusion, the BiLSTM model has the highest fitness with the small sample multi-source heterogeneous data and the results prove the accuracy, stability, and superiority of the wear prediction model of micro-grinding tool based on the BiLSTM.

Fig. 11 The comparison figures of relative error for prediction models. **a** 1st–4th groups; **b** 5th–8th groups; **c** 9th–12th groups; **d** 13th–16th groups
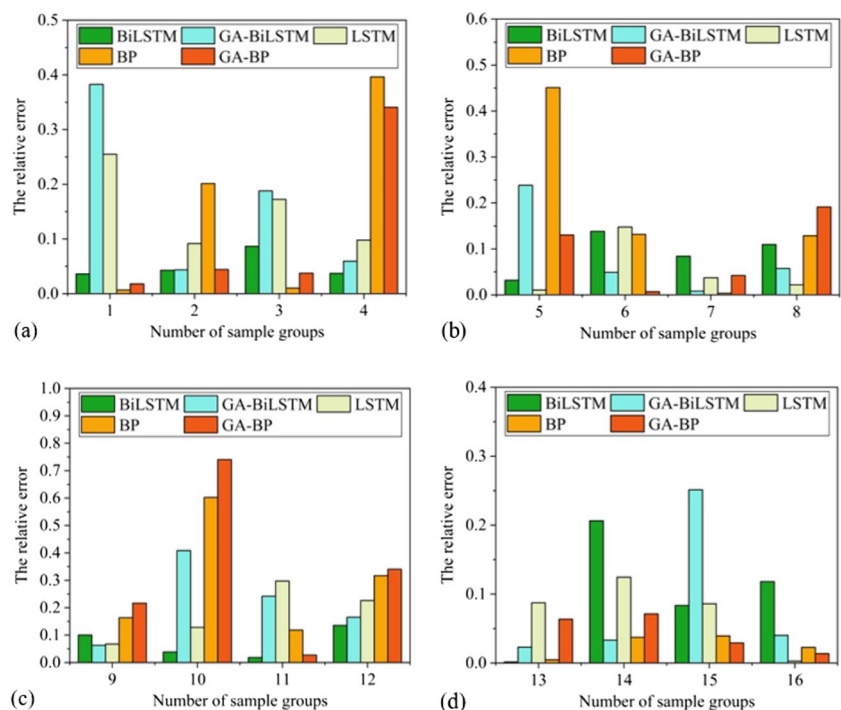
4 Conclusions

In this work, the BiLSTM prediction model was constructed to integrate small-sample and multi-source heterogeneous data. The micro-grinding tool wear is characterized and predicted.

The calculated and experimental results show that the prediction accuracy of BiLSTM model is 92.08%, which proves the feasibility of the prediction model. In comparison among the GA-BiLSTM, LSTM, and BiLSTM models, the BiLSTM model has the highest prediction accuracy, which indicates that the model has the strongest fitness of architectural depth of model and sample data. In comparison among the BP, GA-BP, and BiLSTM models, the BiLSTM model has the highest prediction accuracy and the smallest relative error fluctuation, which proves the stability and superiority of the BiLSTM model compared with the machine learning methods.

The BiLSTM model proposed in this paper has important scientific significance for the wear characterization and prediction of micro-grinding tools. On the one hand, it is able to guide the replacement strategy so that the processing quality will be ensured and the waste will be reduced. On the other hand, it can provide more solutions for sustainable manufacturing and provide theoretical basis for independent decision-making in precision intelligent manufacturing.

Code availability The data generated during this study are available from the corresponding author on reasonable request.

Declarations

Ethical approval The work does not contain libelous or unlawful statements and does not infringe on the rights of others, or contains material or instructions that might cause harm or injury.

Consent to participate All authors consent to participate.

Consent for publication All authors agree to transfer copyright of this article to the publisher.

Conflict of interest The authors declare no competing interests.

References

1. Li W, Ren YH, Li CF, Li ZP, Li MJ (2020) Investigation of machining and wear performance of various diamond micro-grinding tools. Int J Adv Manuf Technol 106(3-4):921–935. https://doi.org/10.1007/s00170-019-04610-4

2. An Q, Tao Z, Xu X, El Mansori M, Chen M (2020) A data-driven model for milling tool remaining useful life prediction with convolutional and stacked LSTM network. Measurement 154:107461. https://doi.org/10.1016/j.measurement.2019.107461

3. Ren Y, Li C, Li W, Li M, Liu H (2019) Study on micro-grinding quality in micro-grinding tool for single crystal silicon. J Manuf Process 42:246–256. https://doi.org/10.1016/j.jmapro.2019.04.030

4. Mohanraj T, Shankar S, Rajasekar R, Sakthivel NR, Pramanik A (2020) Tool condition monitoring techniques in milling process - a review. J Mater Res Technol-Jmr&T 9(1):1032–1042. https://doi.org/10.1016/j.jmrt.2019.10.031

5. Wu SH, Li Y, Li WG, Zhao XZ, Luo CL, Yu QL, Lin SJ (2023) A hybrid network capturing multisource feature correlations for tool remaining useful life prediction. Intal J Adv Manuf Technol 125(5-6):2815–2831. https://doi.org/10.1007/s00170-023-10837-z

6. Yang X, Fang Z, Yang Y, Mba D, Li X (2019) A novel multi-information fusion grey model and its application in wear trend prediction of wind turbines. Appl Math Modell 71:543–557. https://doi.org/10.1016/j.apm.2019.02.043

7. Gai XY, Cheng YN, Guan R, Jin YB, Lu MD (2022) Tool wear state recognition based on WOA-SVM with statistical feature fusion of multi-signal singularity. Int J Adv Manuf Technol 123(7-8):2209–2225. https://doi.org/10.1007/s00170-022-10342-9

8. Banda T, Farid AA, Li C, Jauw VL, Lim CS (2022) Application of machine vision for tool condition monitoring and tool performance optimization-a review. Int J Adv Manuf Technol 121(11-12):7057–7086. https://doi.org/10.1007/s00170-022-09696-x

9. Zhang XH, Yu HY, Li CC, Yu ZJ, Xu JK, Li YQ, Yu HD (2023) Study on in-situ tool wear detection during micro end milling based on machine vision. Micromachines 14(1). https://doi.org/10.3390/mi14010100

10. Qin Y, Liu X, Yue C, Zhao M, Wei X, Wang L (2023) Tool wear identification and prediction method based on stack sparse self-coding network. J Manuf Syst 68:72–84. https://doi.org/10.1016/j.jmsy.2023.02.006

11. Zhuo RJ, Deng ZH, Chen B, Liu GY, Bi SH (2021) Overview on development of acoustic emission monitoring technology in sawing. Int J Adv Manuf Technol 116(5-6):1411–1427. https://doi.org/10.1007/s00170-021-07559-5

12. Wang S, Zhao QL, Wu T (2022) An investigation of monitoring the damage mechanism in ultra-precision grinding of monocrystalline silicon based on AE signals processing. J Manuf Process 81:945–961. https://doi.org/10.1016/j.jmapro.2022.07.055

13. Yan BL, Zhu LD, Dun YC (2021) Tool wear monitoring of TC4 titanium alloy milling process based on multi-channel signal and time-dependent properties by using deep learning. J Manuf Syst 61:495–508. https://doi.org/10.1016/j.jmsy.2021.09.017

14. Hou W, Guo H, Luo L, Jin M (2022) Tool wear prediction based on domain adversarial adaptation and channel attention multiscale convolutional long short-term memory network. J Manuf Process 84:1339–1361. https://doi.org/10.1016/j.jmapro.2022.11.017

15. Wang J, Li Y, Zhao R, Gao RX (2020) Physics guided neural network for machining tool wear prediction. J Manuf Syst 57:298–310. https://doi.org/10.1016/j.jmsy.2020.09.005

16. Baydoun S, Fartas M, Fouvry S (2023) Comparison between physical and machine learning modeling to predict fretting wear volume. Tribol Int 177:107936. https://doi.org/10.1016/j.triboint.2022.107936

17. Li Y, Wang J, Huang Z, Gao RX (2022) Physics-informed meta learning for machining tool wear prediction. J Manuf Syst 62:17–27. https://doi.org/10.1016/j.jmsy.2021.10.013

18. Qiang BY, Shi KN, Liu N, Ren JX, Shi YY (2023) Integrating physics-informed recurrent Gaussian process regression into instance transfer for predicting tool wear in milling process. J Manuf Syst 68:42–55. https://doi.org/10.1016/j.jmsy.2023.02.019

19. Huang W, Zhang X, Wu C, Cao S, Zhou Q (2022) Tool wear prediction in ultrasonic vibration-assisted drilling of CFRP: a hybrid data-driven physics model-based framework. Tribol Int 174:107755. https://doi.org/10.1016/j.triboint.2022.107755

20. Zhou Y, Xue W (2018) Review of tool condition monitoring methods in milling processes. Int J Adv Manuf Technol 96(5):2509–2523. https://doi.org/10.1007/s00170-018-1768-5

21. Cheng M, Jiao L, Yan P, Jiang H, Wang R, Qiu T, Wang X (2022) Intelligent tool wear monitoring and multi-step prediction based on deep learning model. J Manuf Syst 62:286–300. https://doi.org/10.1016/j.jmsy.2021.12.002

22. Wei WH, Cong R, Li YT, Abraham AD, Yang CY, Chen ZT (2022) Prediction of tool wear based on GA-BP neural network. Proc Inst Mech Eng Part B-J Eng Manuf 236(12):1564–1573. https://doi.org/10.1177/09544054221078144

23. Chen SH, Lin YY (2023) Using cutting temperature and chip characteristics with neural network BP and LSTM method to predicting tool life. Int J Adv Manuf Technol 127(1-2):881–897. https://doi.org/10.1007/s00170-023-11570-3

24. Siami-Namini S, Tavakoli N, Namin AS (2019) The performance of LSTM and BiLSTM in forecasting time series. Paper presented at the 2019 IEEE International Conference On Big Data (Big Data)

25. Wu XQ, Li J, Jin YQ, Zheng SX (2020) Modeling and analysis of tool wear prediction based on SVD and BiLSTM. Int J Adv Manuf Technol 106(9-10):4391–4399. https://doi.org/10.1007/s00170-019-04916-3

26. Ma JY, Luo DC, Liao XP, Zhang ZK, Huang Y, Lu J (2021) Tool wear mechanism and prediction in milling TC18 titanium alloy using deep learning. Measurement 173. https://doi.org/10.1016/j.measurement.2020.108554

27. Pfeifer T, Wiegers L (2000) Reliable tool wear monitoring by optimized image and illumination control in machine vision. Measurement 28(3):209–218. https://doi.org/10.1016/S0263-2241(00)00014-2

28. Li XL (2002) A brief review: acoustic emission method for tool wear monitoring during turning. Int J Mach Tools Manuf 42(2):157–165. https://doi.org/10.1016/S0890-6955(01)00108-0

29. Liao TW, Ting C-F, Qu J, Blau PJ (2007) A wavelet-based methodology for grinding wheel condition monitoring. Int J Mach Tools Manuf 47(3-4):580–592. https://doi.org/10.1016/j.ijmachtools.2006.05.008

30. Zhang XX, Ren YH, Yang WC, Li W, Yu KN (2022) Study on acoustic emission signal perception of unsteady characteristic in micro-grinding of monocrystalline silicon. J Mech Eng 58(15):121–133

31. Marani M, Zeinali M, Songmene V, Mechefske CK (2021) Tool wear prediction in high-speed turning of a steel alloy using long short-term memory modelling. Measurement 177:109329. https://doi.org/10.1016/j.measurement.2021.109329

32. Goodfellow IBY, Courville A (2016) Deep learning. MIT Press, Cambridge,Massachusetts

33. Harari O, Bingham D, Dean A, Higdon D (2018) Computer experiments: prediction accuracy, sample size and model complexity revisited. Stat Sinica 28(2):899–919. https://doi.org/10.5705/ss.202016.0217

34. Shi J, Zhang YY, Sun YH, Cao WF, Zhou LT (2022) Tool life prediction of dicing saw based on PSO-BP neural network. Int J Adv Manuf Technol 123(11-12):4399–4412. https://doi.org/10.1007/s00170-022-10466-y

35. Meng X, Zhang J, Xiao G, Chen Z, Yi M, Xu C (2021) Tool wear prediction in milling based on a GSA-BP model with a multisensor fusion method. Int J Adv Manuf Technol 114(11-12):3793–3802. https://doi.org/10.1007/s00170-021-07152-w