ORIGINAL ARTICLE

Modeling and analysis of tool wear prediction based on SVD and BiLSTM

Xiaoqiang Wu [1,2,3] · Jia Li [1] · Yongquan Jin [1] · Shuxian Zheng [1]

© Springer-Verlag London Ltd., part of Springer Nature 2020

Abstract
Wear is one of the main forms of tool failure during machining. The prediction of tool wear is of great significance for ensuring the high quality of the workpiece. In order to improve prediction accuracy of tool wear, a tool wear prediction model based on singular value decomposition (SVD) and bidirectional long short-term memory neural network (BiLSTM) is proposed. The cutting force signal is taken as the monitoring signal. Firstly, the raw cutting force signal is reconstructed by Hankle matrix, and the SVD of the reconstructed matrix is performed to extract the signal features. Then, SVD features of the current sampling period and the previous four sampling periods are taken as the input, and the tool wear prediction value at the current time is obtained based on the BiLSTM. The experimental results show that the proposed SVD-BiLSTM model can effectively predict the tool wear and obtain higher prediction accuracy than other comparison models.

1 Introduction

With the continuous advancement and development of manufacturing technology, the manufacturing field has higher and higher requirements for machining efficiency and quality of the workpiece. In machining process, tool wear is the main factor affecting the machining quality and efficiency of the workpiece. Researches show that machine downtime caused by tool wear and damage accounts for 20% of total downtime. Therefore, it is important to monitor the tool wear state in real time during machining process.

The online monitoring system for tool wear generally adopts indirect data-driven monitoring method. That is, signals like the cutting force in the cutting process [1–3], vibration [4–6], acoustic emission [7, 8], and current [9, 10] are captured in real time by sensors and data acquisition devices. Having preprocessed the data and extracted the features, the tool state recognition model based on machine learning is used to monitor the tool wear state. In recent years, the rapid development of sensor technology and computer storage performance has made it possible to collect sensor time series data on a large scale during the cutting process. Therefore, the data-driven method has been greatly developed and applied in the field of tool state monitoring [11].

In the research of data-driven tool wear state monitoring methods, the key to the research is the feature extraction and pattern recognition of the monitoring signals. At present, researchers have done a lot of work on these two aspects. Li et al. [12] extracted the features of acoustic emission signals by wavelet packet transform, and they also realized the recognition of tool wear state by the combination of wavelet packet transform and fuzzy clustering method. Based on the integration of principal component analysis and wavelet analysis, Wang [13] proposed a multi-scale principal component analysis method for online monitoring of tool wear during milling process. Babouri [14] put forward a spectral center method for analyzing vibration signals to identify three wear stages of plunge milling tool. Li [15] extracted the 14 time domain features of the cutting force signal by correlation analysis and proposed a V-SVR model for tool wear state recognition. Li Guihong [16] proposed a method for identifying wear state

✉ Xiaoqiang Wu
  wangzai8402@163.com

1 Tianjin Key Laboratory of Equipment Design and Manufacturing Technology, Tianjin University, Tianjin 300350, China

2 College of Mechanical Engineering, Inner Mongolia University for the Nationalities, Tongliao 028000, China

3 Key Laboratory of Mechanism and Equipment Design of Ministry of Education, Tianjin University, Tianjin 300350, China

of milling cutter based on EMD and SVM. In the above researches, the tool wear state is divided into initial wear, medium wear, and severe wear. Although the above methods can accurately identify which wear stage the tool is in, it cannot accurately predict the tool wear value. With the tool wear state, it is impossible to accurately identify the optimal time for tool change in the manufacturing process.

It is essential to achieve accurate tool wear prediction in real time. In this way, the tool can be timely replaced according to current tool wear value and preset tool wear threshold. As a result, the surface quality and dimensional accuracy of the workpiece could be ensured, less waste of workpiece raw materials there would be, downtime would be reduced. Moreover, it can avoid the waste of tool working life and the increase of tool cost caused by frequent tool change.

The wear of the tool is a long-lasting process. The wear state of the current moment is closely related to the wear state of the first few moments. The information contained in the data collected in the first few moments plays a guidance role in judging the wear state of the tool at the current moment. Therefore, in order to discover the rule of tool wear degradation and the time series dependency between tool wear states at various times, the tool wear prediction model should be able to learn the features of time series data.

With the development of deep learning in recent years, the ability of the recurrent neural network (RNN) to process time series data has become increasingly prominent and has been applied and developed in the field of mechanical fault diagnosis [17]. As a variant of RNN, long short-term memory (LSTM) can effectively extract the dynamic variation features and long-term time series dependency features of time series data. Zhao et al. [18] used LSTM for tool wear state monitoring. Malhotra P et al. [19] proposed a LSTM-based Encoder-Decoder scheme for reconstructing time series signals under normal conditions and detect the anomaly by calculating signals and reconstruction errors. Bruin et al. [20] used the LSTM network to diagnose railway track fault current in real time. In the above schemes and applications, since the raw signal input into the LSTM model is a low sampling rate, good results can be obtained when LSTM is directly used to process the raw signal.

However, the vibration signals commonly used in mechanical equipment fault diagnosis are mostly high sampling rate signals. As a result, the length of a signal that can effectively reflect the fault feature information is usually greater than 1000 or even longer. Since layers in the LSTM network are connected fully, when the raw signal containing a large amount of noise is directly processed by the LSTM, the model parameters are too large, making the model difficult to train and easily causing overfitting phenomenon. In addition, ordinary LSTM can only memorize the data change features before current time. In order to accurately reflect the features of the series data of the fault, the BiLSTM algorithm [21] is selected.

Based on above analysis, this paper proposes a tool wear prediction model based on SVD [22] and BiLSTM. Firstly, the SVD is used to extract the features of the raw signal. It can not only eliminate the influence of noise on the robustness and generalization ability of the raw signal but also shorten the input signal length of the BiLSTM algorithm. Besides, it can also reduce the complexity of the BiLSTM algorithm and improve the performance of the algorithm. Then, BiLSTM is used to extract the time series variation features of the current sampling period and the first few sampling periods to predict the tool wear value. According to the data of the open dataset in the 2010PHMdata challenge, the SVD-BiLSTM model is verified by comparative study with other models.

2 Raw signal feature extraction based on SVD

The signal processing method based on SVD has unique processing ability in nonlinear and non-stationary signal analysis [23], so it is very suitable for signal processing in the field of mechanical condition monitoring and fault diagnosis. Before extracting the singular value feature of signal, the raw signal needs to be preprocessed and the target matrix needs to be constructed. Since the SVD method based on Hankel matrix can well describe the features of the weak signal, the raw cutting force signal is used to construct the Hankel matrix.

Let $X = [x_1, x_2, x_3, ..., x_N]$ be the raw signal collected, and the reconstructed $m \times n$ Hankel matrix can be expressed as

$$H = \begin{bmatrix} x_1 & x_2 & ... & x_n \\ x_2 & x_3 & ... & x_{n+1} \\ ... & ... & ... & ... \\ x_m & x_{m+1} & ... & x_N \end{bmatrix} \quad (1)$$

where $N = m + n + 1$ and $1 < n < N$, $m \geq 2$, $n \geq 2$.

Having constructed the Hankel matrix, the SVD of the Hankel matrix is performed. According to the SVD theory, the decomposition of H is obtained as

$$H = U.S.V^T \quad (2)$$

where U is an $m \times m$ orthogonal matrix, V is an $n \times n$-orthogonal matrix, and S is a non-negative real number diagonal matrix, expressed as

$$S = \begin{bmatrix} \sigma_1 & 0 & ... & 0 \\ 0 & \sigma_2 & ... & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & ... & \sigma_n \end{bmatrix} \quad (3)$$

The diagonal element $\sigma_i (i = 1, 2, ..., n)$ is the singular value of the characteristic matrix. In this paper, the first k valid singular values are reserved as the feature set of the raw signal, and the feature set is expressed as $f = (\sigma_1, \sigma_2, ..., \sigma_k)$.

Fig. 1 LSTM unit diagram
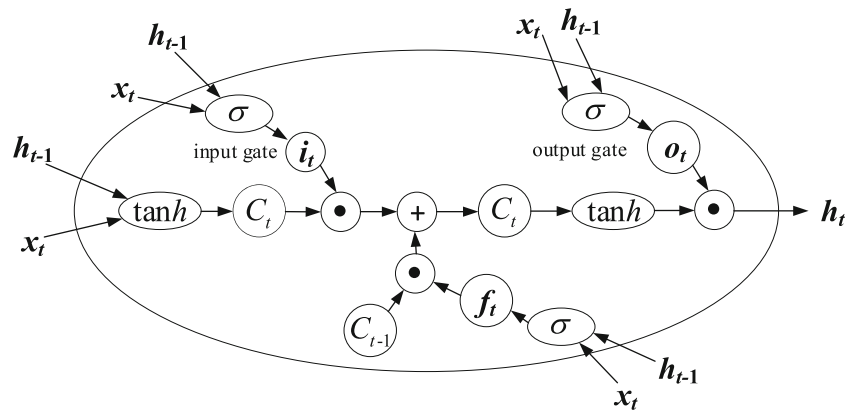
3 BiLSTM theory

As a special structure of RNN, LSTM not only avoids the phenomenon of "vanishing" gradient and "exploding" gradient but also has longer time memory than RNN and other models. The hidden layer of RNN has only one state h sensitive to short-term input. While in the hidden layer of LSTM model, a long-term state memory c is added. As shown in Fig. 1, the LSTM mainly controls the memory unit c through three gate structures, namely input gate, forget gate, and output gate. The main function of input gate is to control the inputs of the temporary status information into the memory unit c; the forget gate mainly controls which information is forgotten at the last moment; the output gate is responsible for controlling whether to output the current status information. The unit diagram of LSTM is shown in Fig. 1.
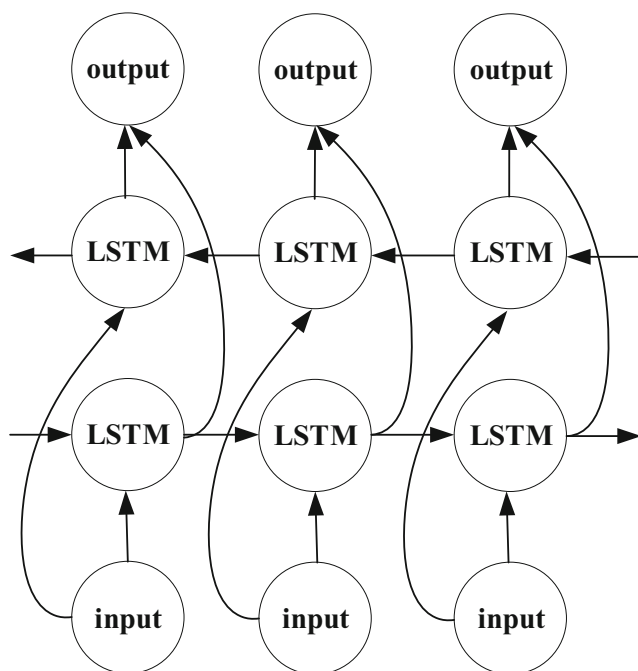

Fig. 2 BiLSTM network structure

The function of the forget gate is to output a number f_t between 0 and 1 by reading h_{t-1} and x_t. This number will determine the degree of retention of c_{t-1}, and the formula is

$$f_t = \sigma\left(W_f h_{t-1} + U_f x_t + b_f \right. \tag{4}$$

The role of the input gate is to determine which new information will be transmitted to the cell state in h_{t-1} and x_t. The specific data processing is divided into two parts. The first part determines which values are updated by sigmoid activation function. The second part outputs a_t by using tanh activation function, and then add it to the cell state. The formula is

$$i_t = \sigma\left(W_i h_{t-1} + U_i x_t + b_i \right. \tag{5}$$

$$a_t = tanh(W_a h_{t-1} + U_a x_t + b_a) \tag{6}$$

Then update the cell state:

$$c_t = f_t{}^\circ c_{t-1} + i_t{}^\circ a_t \tag{7}$$

where, "∘" denotes the Hadamard product. The output gate is also composed of two parts, the first part being h_{t-1}, x_t and the sigmoid activation function. The second part consists of c_t and the tanh activation function:

$$o_t = \sigma(W_o h_{t-1} + U_o x_t + b_o) \tag{8}$$

$$h_t = o_t{}^\circ tanh(c_t) \tag{9}$$

In all the above formulas, W, U, and b are linear correlation coefficient and bias coefficient, and σ is the sigmoid activation function.

The main characteristic of LSTM is that each new input contains certain features from the previous input. As a whole, the model has a long-term memory ability by continuously updating the cell state.

The BiLSTM consists of forward LSTM and backward LSTM that obtain front and rear sections features, respectively. Compared with LSTM, the state of BiLSTM current recurrent unit is affected by the pre and post data. With the BiLSTM, the whole information
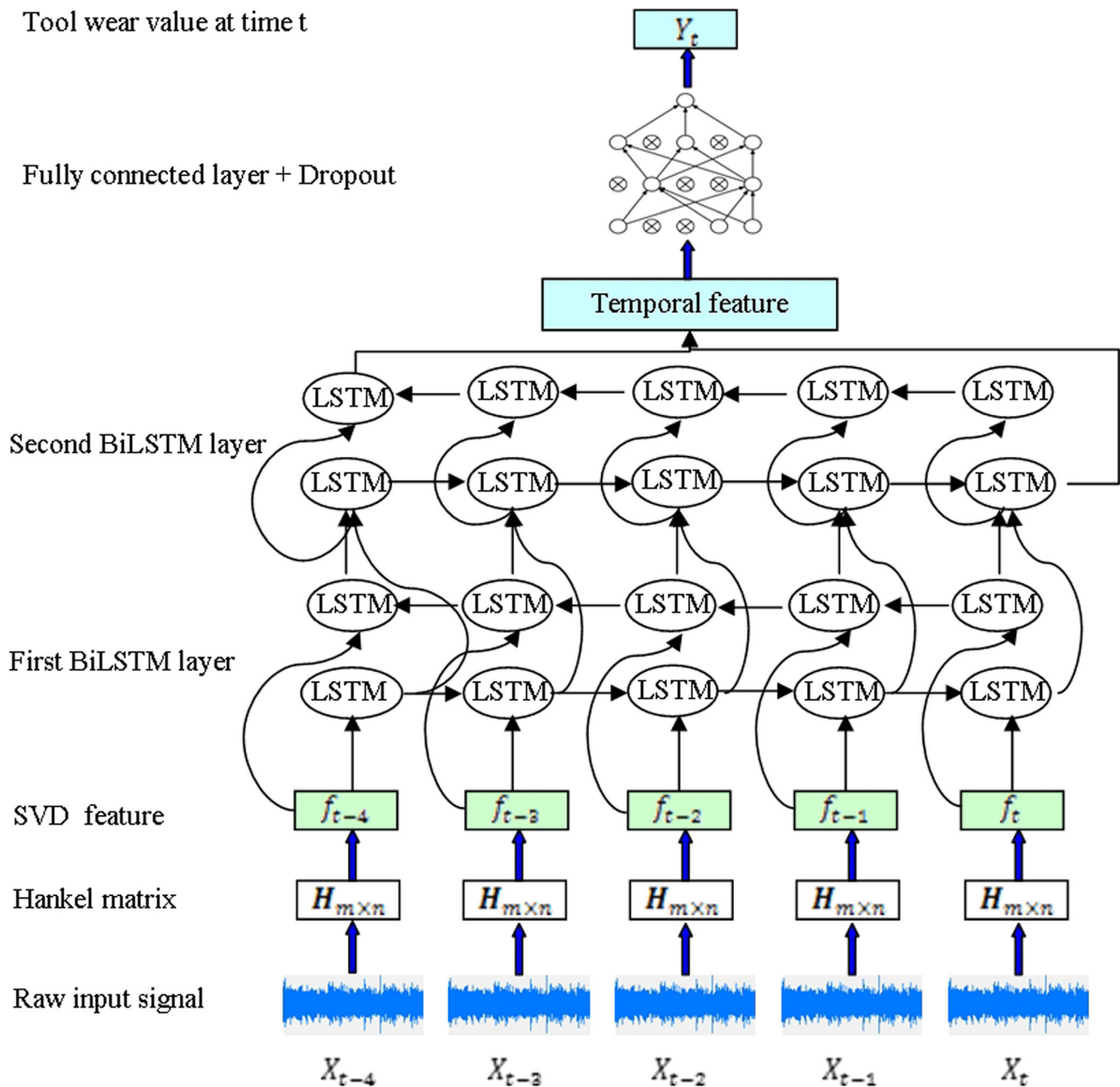
Fig. 3 Diagram of the SVD-BiLSTM tool wear prediction model

can be better grasped in processing time series data. Therefore, BiLSTM is used as the basis of the tool wear prediction model.
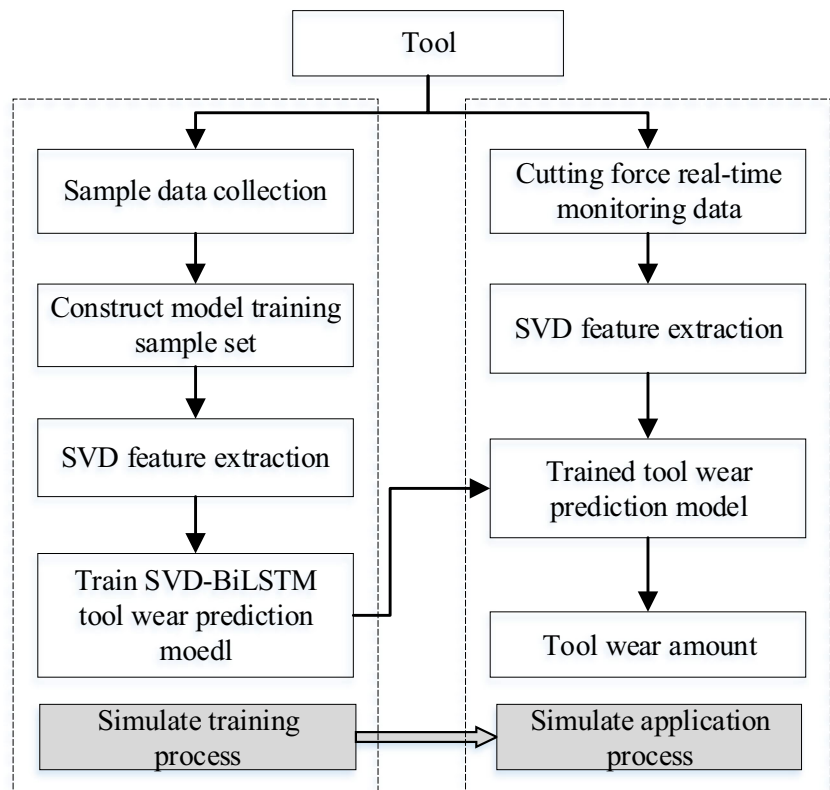
The tool wear prediction model takes the feature set of the raw cutting force signal as input, and the model structure is shown in Fig. 2. During the cutting process, a cutting force signal of time span t is acquired for each length of cutting, the time interval of which is taken as a signal acquisition period. The signal features of the cutting force acquired during the current sampling period reflect the current tool wear. The wear degradation of the tool is a continuous process. Since the current wear state is closely related to the wear state of the first few moments, the information

contained in the data collected in the first few moments is important for judging the wear state of the tool at the present time. Considering the characteristics of the BiLSTM theory, the BiLSTM can be used to learn the time series variation of tool wear during these adjacent sampling periods.

4 SVD-BiLSTM tool wear prediction model

The SVD-BiLSTM tool wear prediction model is shown in Fig. 3. Let the cutting force signal obtained by the current

Fig. 4 Flow chart of training and application process of the SVD-BiLSTM tool wear prediction model

sampling period be X_t, then the input of the tool wear prediction model is $X = (X_{t-4}, X_{t-3}, X_{t-2}, X_{t-1}, X_t)$, the output is the tool wear value Y_t corresponding to the current signal sampling period. The model consists of two parts, the singular value feature extraction process and the BiLSTM based tool wear prediction process. Firstly, the target matrix is reconstructed for X_{t-4}、 X_{t-3}、 X_{t-2}、 X_{t-1}, and X_t by Hankel matrix construction method, and then each target matrix is performed singular value decomposition respectively, and 15 singular values are extracted as features of each

signal. The extracted input signal features can be expressed as $F = (f_{t-4}, f_{t-3}, f_{t-2}, f_{t-1}, f_t)$, $F \in R^{5 \times k}$. The feature F is then input into two layers of BiLSTM for time series feature extraction. Finally, the extracted time series features are input into the fully connected layer with dropout and the linear regression output layer to complete the model construction.

The training and application process flow chart of the tool wear prediction model is shown in Fig. 4 which includes the following steps:

Fig. 5 Experiment device
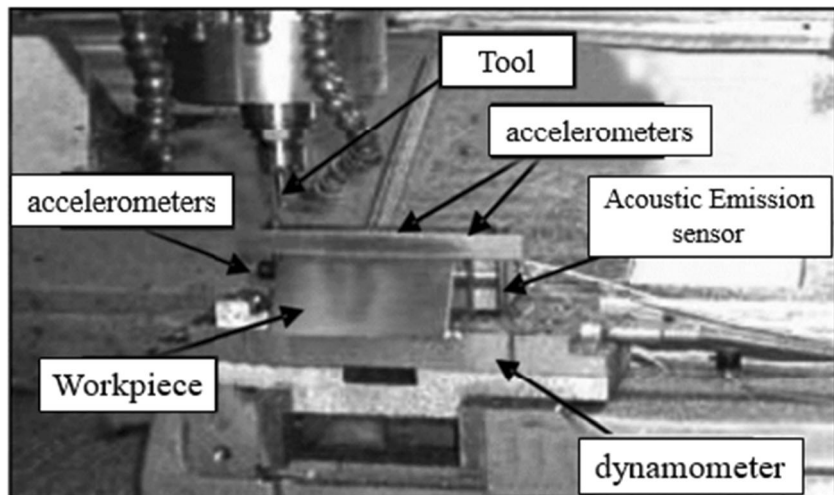
Table 1 Parameters of BiLSTM model

Type	Input signal dimension	Output signal dimension	Activation function	Dropout proportion
The first layer Bi-direction LSTM	((None, 5, 15)	(None, 5, 600)	Sigmiod	-
The second layer Bi-direction LSTM	(None, 5, 600)	(None, 300)	Sigmiod	-
dense_1 (dense)	(None, 300)	(None, 20)	Relu	0.5
dense_2 (dense)	(None, 20)	(None, 1)	Relu	-

(1) Acquiring the model training sample data: After cutting for a certain length of the bevel, the cutting force signal of time span t in the cutting process is taken as X_t, the tool is then removed, and the wear amount of the tool flank is measured by the microscope measuring device as Y_t. This constituted a sample data (X, Y_t), $X = (X_{t-4}, X_{t-3}, X_{t-2}, X_{t-1}, X_t)$. Loop through this process to obtain a lot of sample data.

(2) First, perform singular value decomposition on each X in the sample data to obtain the singular value feature set $F = (f_{t-4}, f_{t-3}, f_{t-2}, f_{t-1}, f_t)$, and construct a sample (F, Y_t) to train the BiLSTM tool wear prediction model.

(3) The sample set (F, Y_t) is divided into a model training sample and a model test sample. The first step is to select the appropriate number for BiLSTM network layers and neurons, initialize the network parameters, set the number of trainings, and then train the model BiLSTM. In each training, the cost function of the current training stage is obtained through forwarding propagation, and the network parameters are updated by error back propagation. The training would not stop until the set training times are reached. Then the test set is used to test the model prediction accuracy.

(4) After the model is trained, in the tool cutting process, the cutting force signal is collected in real-time. Having extracted SVD feature, the extracted SVD feature is input into the trained tool wear prediction model to obtain the tool wear value.

The mean absolute error (MAE), the mean absolute percentage error (MAPE), and the root mean square error (RMSE) are used as the evaluation indicators of the model, as shown in Eqs. (10), (11), and (12):

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_t - y_p| \tag{10}$$

$$MAPE = \frac{100}{n} \sum_{i=1}^{n} \left| \frac{y_t - y_p}{y_t} \right| \tag{11}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(y_t - y_p \right)^2} \tag{12}$$

where y_t is the test value of the wear, and y_p is the predicted value of the wear.

Fig. 6 Tool wear prediction result using SVD-BiLSTM
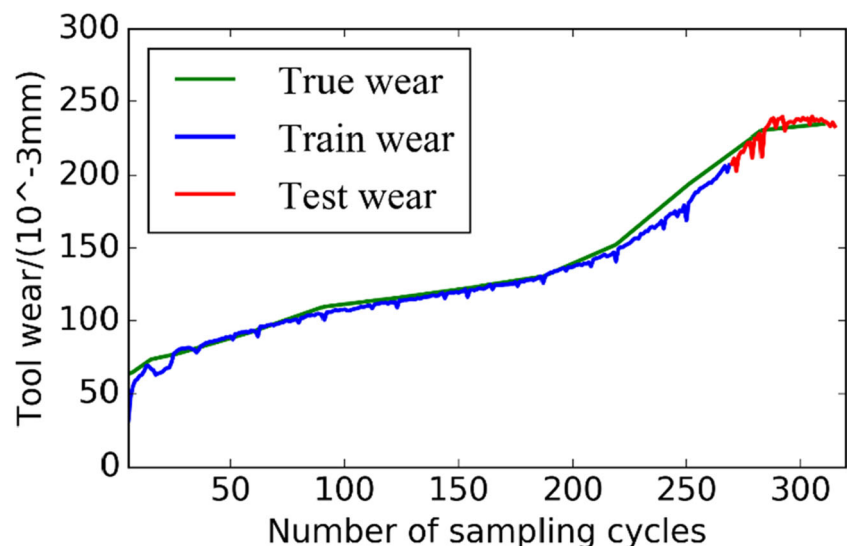
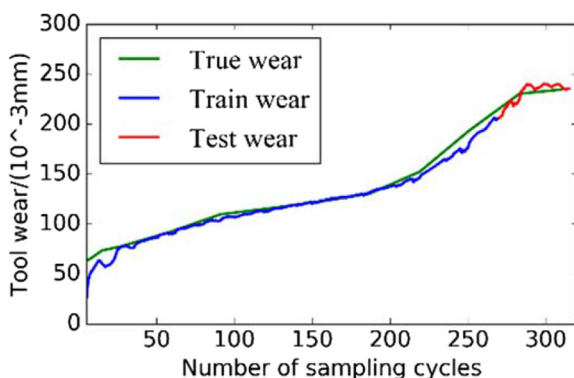**Table 2** Comparison of prediction results of BiLSTM algorithm based on different characteristics
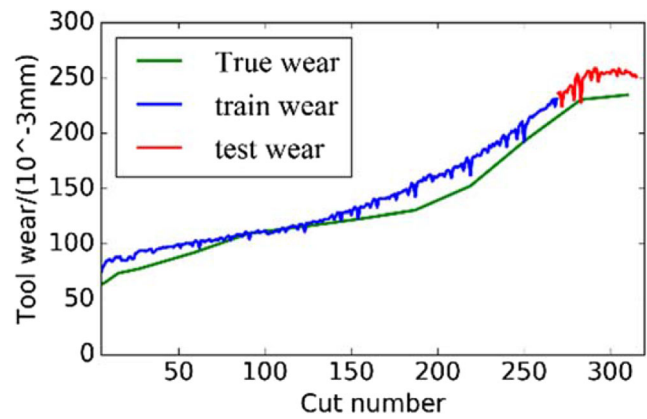
Feature type	MAE	MAPE	MSE
SVD feature	2.7642	2.334	3.9186
Time domain feature	9.0911	8.7446	3.9186

5 Experiment analysis and verification

5.1 Experiment data description

Data used in this paper is from the open dataset for milling tool wear prediction [24] in the 2010 PHM data challenge. The experiment was carried out on a high-speed CNC machine, and a ball-end milling cutter with three grooves was used to mill the workpiece of Inconel 718. The experiment device is shown in Fig. 5. During the experiment, the workpiece bevel was continuously milled to process the complete bevel. The machining parameters are as follows: the spindle speed was 10,400 r/min, the feed rate in x direction was 1555 mm/min; the cutting depth in y direction (radial) was 0.125 mm; the cutting depth in z direction (axial) was 0.2 mm. During the milling process, the Kistler quartz three-component dynamometer was used to measure the cutting force, three Kistler piezoelectric accelerometers were used to measure the machine vibration signal, and the Kistler Acoustic Emission (AE) sensor was used to measure the acoustic emission signal generated during the cutting process. Seven signal channels (force_x, force_y, force_z, vibration_x, vibration_y, vibration_z, AE_RMS) were acquired using an NI PCI1200 data acquisition device with a sampling frequency of 50 kHz. After a surface milling (one-shot) was performed, the flank wear of the three grooves was measured offline using a LEICA MZ12 microscope.

In this paper, the data of cutter No. 6 (C6) was used for the experiment, and cutter No. 6 was measured for 315 times. The cutting force signal in x direction with a length of 2048 was taken as an input signal of one sampling period, and this was
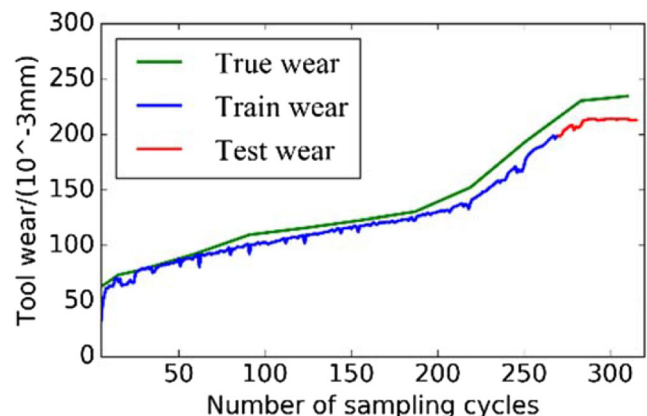

Fig. 8 Tool wear prediction result using SVD-RNN

an input signal of a timestep in the model. Then the average wear of the three grooves was calculated as the tool wear.

5.2 Model application and prediction results

The input data of the model was composed according to time step = 5, that is, the cutting force signal for each consecutive 5 sampling periods corresponding to the tool wear value of the 5th sampling period. Three hundred fifteen total measured data constituted 311 sets of sample data, of which the first 265 sets of samples were taken as the model training sample. The remaining 46 groups of samples were used as model test sample. Then the SVD-BiLSTM tool wear prediction model was established. When using the SVD, select $k = 15$. The main parameters of the BiLSTM model are shown in Table 1. In addition, during the model training, the epoch was 50, the batch size was 15, and the model optimizer was Adam [25]. The model runs on the i5-6500 CPU with a model training speed of 21 ms/batch.

The experiment results of the proposed tool wear prediction model based on SVD-BiLSTM are shown in Fig. 6. It can be seen from the figure that the predicted tool wear values can well follow the trend of real tool wear values with little errors. It is proved that the model can adaptively extract features

Fig. 7 Tool wear prediction result using SVD-LSTM

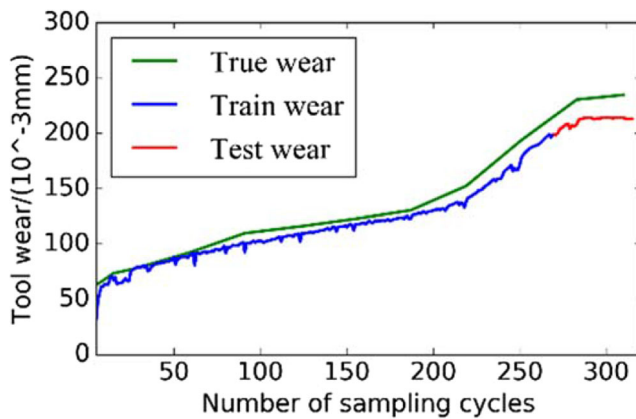**Fig. 9** Tool wear prediction result using SVD-GRU

Fig. 10 Tool wear prediction result using SVD-BiGRU

related to tool wear from the raw data and use it for tool wear prediction.

5.3 Model verification and discussion

In order to verify the performance of SVD features in tool wear prediction, the prediction effects of BiLSTM algorithm based on SVD features and time domain features were compared. In order to construct the time domain features of the same scale as the SVD features, 15 time domain features were extracted, namely mean, root mean square, RMS amplitude, absolute mean, skewness, kurtosis, variance, maximum, minimum, peak-to-peak, waveform index, peak index, pulse index, margin index, and kurtosis index. The prediction results of the two models are shown in Table 2.

To further verify the performance of the SVD-BiLSTM tool wear prediction model, LSTM, RNN, GRU, and BiGRU were used for comparative experiments. All recurrent neural network models had the same structure, so only the type of the recurrent neural network layer needed to be changed, while the network parameters remained unchanged, and all model inputs were SVD features.

The prediction results of each model are shown in Figs. 7, 8, 9, and 10 and Table 3. Each model can well reflect the trend of tool wear and can accurately predict the wear value of

Table 3 Comparison of prediction results of different recurrent neural network models

Type	Training set			Test set		
	MAE	MAPE	MSE	MAE	MAPE	MSE
BiLSTM	2.4948	2.4112	3.6782	*4.3044*	*1.8896*	*5.0873*
LSTM	2.5021	2.5266	4.2013	4.6459	2.0312	2.0312
RNN	14.5117	12.3313	16.9749	21.8238	9.6101	22.3491
FNN	*2.4186*	*2.3806*	*3.0905*	10.7307	4.6869	11.1283
BiGRU	5.7163	4.6654	6.7617	16.551	7.2364	16.9182

subsequent data. In each model, the effect of RNN is the worst, because there is only one hidden state in RNN. Although this state is constantly adjusted according to the input data and previous hidden states, it is more greatly affected by the previous hidden state. Therefore, when analyzing long-sequence data, its long-term memory ability is poor. The better-performing LSTM model has three kinds of gate structures, namely input gate, output gate, and forget gate, and cell state, so it owns good long-term memory ability, gaining great advantages in processing time series data. As a result, its final prediction results are significantly better than other structure models. The BiLSTM selected in this paper is the best model in all data and test sets. This is because the bidirectional LSTM can memorize the features of all data. Therefore, BiLSTM is better than LSTM in both the training set and the test set. In the error comparison of training set, GRU model performs better. Compared with the three gate structures of LSTM, GRU is simplified into two structures, the update gate and the reset gate. The optimization makes calculation simpler. So in the case of the same epoch, GRU can achieve better training results in the training set. However, GRU does not have cell state, and GRU's long-term memory ability is poor, which is directly reflected in GRU test set error value. The bidirectional GRU is even worse, because the poor long-term memory of each unit makes the unit weight more complicated and confusing.

6 Conclusion

The SVD-BiLSTM model proposed in this paper can be applied to the prediction of tool wear, and it can improve the prediction accuracy. Firstly, the Hankel matrix was used to reconstruct the raw cutting force signal, and then the SVD feature of the cutting force signal was extracted by singular value decomposition. From the comparative experiment between the tool wear prediction model based on time domain features and SVD features, it can be found that BiLSTM algorithm based on SVD features can obtain higher tool wear prediction accuracy than that based on time domain features. After extracting the features, the BiLSTM algorithm was used to further extract the time series features between adjacent sampling periods and predict the tool wear. From the comparison experiments of different recurrent neural network models, it can be found that the SVD-BiLSTM tool wear prediction model has the lowest prediction error, and its prediction performance is better than other recurrent neural networks.

References

1. Guo JC, Li AH (2019) Advances in monitoring technology of tool wear condition [J]. Tool Engineering 53(05):3–13
2. Patra K, Jha AK, Szalay T et al (2017) Artificial neural network based tool condition monitoring in micro mechanical peck drilling using thrust force signals [J]. Precis Eng 48:279–291
3. Li XL, Li HX, Guan XP (2004) Fuzzy estimation of feed-cutting force from current measurement - a case study on intelligent tool wear condition monitoring [J]. IEEE Trans Syst Man Cybern Part C Appl Rev 34(4):506–512
4. Li X, Dong S, Venuvinod PK (2000) Hybrid learning for tool wear monitoring [J]. Int J Adv Manuf Technol 16(5):303–307
5. Dimla DE (2002) The correlation of vibration signal features to cutting tool wear in a metal turning operation[J]. Int J Adv Manuf Technol 19(10):705–713
6. Wu D, Huang M (2014) Application of vibration signal monitoring in tool wear fault diagnosis [J]. Mechanical Engineering & Automation (02):121–122+125
7. Pai PS, Rao PKR (2002) Acoustic emission analysis for tool wear monitoring in face milling[J]. Int J Prod Res 40(5):1081–1093
8. Zhou JH, Pang CK, Zhong ZW et al (2011) Tool wear monitoring using acoustic emissions by dominant-feature identification[J]. IEEE Trans Instrum Meas 60(2):547–559
9. Li X, Tso SK (1999) Drill wear monitoring based on current signals [J]. Wear 231(2):172–178
10. Li X, Djordjevich A, Venuvinod PK (2000) Current-sensor-based feed cutting force intelligent estimation and tool wear condition monitoring[J]. IEEE Trans Ind Electron 47(3):697–702
11. Chen Y, Jin Y, Jiri G (2018) Predicting tool wear with multi-sensor data using deep belief networks[J]. Int J Adv Manuf Technol
12. Li XL, Yuan ZJ (1998) Tool wear monitoring with wavelet packet transform-fuzzy clustering method[J]. Wear 219(2):145-154
13. Wang G, Zhang Y, Liu C et al (2016) A new tool wear monitoring method based on multi-scale PCA [J]. J Intell Manuf:1–10
14. Babouri MK, Ouelaa N, Djamaa MC et al (2017) Prediction of tool wear in the turning process using the spectral center of gravity [J]. J Fail Anal Prev:1–9
15. Li N, Chen Y, Kong D et al (2017) Force-based tool condition monitoring for turning process using v-support vector regression [J]. Int J Adv Manuf Technol 91(1–4):351–361
16. Li G, Du X, Zhao LL et al (2019) Design of milling-tool wear monitoring system based on EEMD-SVM [J]. Automation & Instrumentation 06:30–32
17. Rui Z, Ruqiang Y, Zhenghua C et al (2019) Deep learning and its applications to machine health monitoring[J]. Mech Syst Signal Process 115:213–237
18. Zhao R, Wang J, Yan R et al (2016) Machine health monitoring with LSTM networks[C]// International Conference on Sensing Technology. IEEE
19. Malhotra P , Ramakrishnan A , Anand G , et al. LSTM-based encoder-decoder for multi-sensor anomaly detection[J]. 2016
20. Bruin TD, Verbert K, Babuška R (2017) Railway track circuit fault diagnosis using recurrent neural networks. IEEE Trans Neural Netw Learn Syst 28:523–533. https://doi.org/10.1109/TNNLS.2016.2551940
21. Graves A, Fernández S, Schmidhuber J, Bidirectional LSTM (2005) Networks for improved phoneme classification and recognition[M]. Artificial Neural Networks: Formal Models and Their Applications – ICANN 2005
22. Zhao XZ, Bang-Yan YE, Lin Y (2011) Amplitude modulation feature extraction of bearing vibration signal using singular value decomposition[J]. Transactions of Beijing Institute of Technology 31(5):572–577
23. Huang J N, Wang S H, Ma C. Fault diagnosis of rolling bearing based on SVD-EEMD and BP neural network [J]. J Beijing Inf Sci Technol University,2019,34(02):69–74
24. Li X, Lim B, Zhou J, Huang S, Phua S, Shaw K, Er M (September 2009) Fuzzy neural network modelling for tool wear estimation in dry milling operation. In Proceedings of the Annual Conference of the Prognostics and Health Management Society, San Diego, CA, USA, pp 27–30
25. Kingma D, Ba J (2014) Adam: a method for stochastic optimization[J]. Comput Sci