Tool wear condition monitoring in milling process based on data fusion enhanced long short-term memory network under different cutting conditions

Guoxiao Zheng [a], Weifang Sun [a], Hao Zhang [b], Yuqing Zhou [a,*], Chen Gao [c,*]

[a] College of Mechanical and Electrical Engineering, Wenzhou University, Wenzhou, China, 325035
[b] Shaoxing Customs, Shaoxing, China, 312099
[c] School of Mechatronics and Transportation, Jiaxing Nanyang Polytechnic Institute, Jiaxing, China, 314031

Highlights

- A data fusion- LSTM is proposed to estimate tool wear under different cutting conditions.

- NCA is used to select useful features fusioned by EMD VMD and FSST.

- Experimental results show the proposed method outperforms significantly SVR and RNN.

Abstract

Tool wear condition monitoring (TCM) is essential for milling process to ensure the machining quality, and the long short-term memory network (LSTM) is a good choice for predicting tool wear value. However, the robustness of LSTM- based method is poor when cutting condition changes. A novel method based on data fusion enhanced LSTM is proposed to estimate tool wear value under different cutting conditions. Firstly, vibration time series signal collected from milling process are transformed to feature space through empirical mode decomposition, variational mode decomposition and fourier synchro squeezed transform. And then few feature series are selected by neighborhood component analysis to reduce dimension of the signal features. Finally, these selected feature series are input to train the bidirectional LSTM network and estimate tool wear value. Applications of the proposed method to milling TCM experiments demonstrate it outperforms significantly SVR- based and RNN- based methods under different cutting conditions.

Keywords

tool wear condition monitoring, empirical mode decomposition, variational mode decomposition, fourier synchro squeezed transform, neighborhood component analysis, long short-term memory network.

1. Introduction

In the modern numerical control milling process, tool condition is one of the key factors affecting the machining quality of workpiece [19, 22]. Tool breakage is the main cause of abnormal shutdown and lead to time lost and capital destroyed [27]. It has reported that severe tool failure causes at least 20% of abnormal downtime [4, 32]. However, traditional tool condition monitoring (TCM) methods are based on the machining time or the number of workpiece machined resulting in the effective utilization rate of tool is only 50%-80%, which affect the processing efficiency and increase the machining cost significantly [15, 35]. It is predicted that an effective TCM method can increase the cutting efficiency by 10-50% and reduce the machining cost by 10-40% [23, 33]. Therefore, the development of effective online TCM method has received broadly positive reviews and is a research hotspot nowadays [10, 11].

Recently, many deep learning models have been employed in TCM applications [9, 14, 21]. For example, Cao et al [1] recognized tool wear condition by derived wavelet frames and Convolutional neural network (CNN) using vibration signals. Recent advanced technology that have greatly increased the number of TCM study, Huang et al [8] proposed a tool wear predicting method by deep CNN, in which multi- domain features are respectively extracted from cutting force and vibration. Lei et al [16] employed Extreme learning machine (ELM) to classify tool wear condition in milling processes, and used genetic algorithm and particle swarm optimization to optimize model parameters of ELM. Tim and Chris [26] proposed a disentangled- variation- autoencoder CNN method to estimate tool wear condition in a self-supervised way. Zhi et al [30] proposed a hybrid CNN and edge-labeling graph neural network (EGNN) method for limited tool wear image training samples, in which the CNN is employed to extract features of tool wear image and the EGNN is applied to distinguish the tool's category. However, these TCM methods have been generally applied for diagnosis (classification) rather than prognosis (regression), tool wear is a progressive and continuous cumulative process, regressive prediction of tool wear is more suitable than classification that the CNN- based methods are difficult to use [34]. Recurrent neural networks (RNN) could be solve the problem

of regression and increase the accuracy of the prognosis, but the error of backpropagation in RNN would increase sharply or decrease exponentially, which lead to the problem of long lag [5,18]. As a significant branch of RNN, Long short-term memory (LSTM) network is proposed to overcome the above problem. Due to the special unit structure with learning long-term dependencies, LSTM can deal with the long-distance dependence problem in time sequence data [6]. Therefore, LSTM is potential to obtain good performance for TCM [31]. Tao et al [24] designed a TCM method based on LSTM and hidden Markov model (HMM) to estimate the tool wear value and predict it's remaining useful life. Zhao et al [29] proposed a convolutional Bi-directional LSTM network, in which CNN extracted local feature of original signal and Bi-directional LSTM encoded temporal information and predict tool wear value. However, it is found that the regression accuracies of LSTM- based TCM method are poor when the cutting conditions of testing samples are different with that of training samples in our experiment. That is, the cutting condition could affect significantly the performance of LSTM- based TCM method. Therefore, this paper try to alleviated the influence of cutting condition to LSTM model through a data fusion way.

In this paper, a data fusion enhanced LSTM- based TCM method is established to estimate tool wear value under variable cutting conditions. The paper is organized as follows: Section 2 introduces the proposed data fusion enhanced LSTM method, Section 3 describes the experimental setup, data analysis and experimental results. Finally, conclusion is in Section 4.

2. Proposed method

2.1. Framework of the proposed method

The proposed TCM method framework based on data fusion enhanced LSTM is illustrated in Figure 1. Firstly, vibration time series signal collected from milling process are transformed to feature space through Empirical mode decomposition (EMD), Variational mode decomposition (VMD) and Fourier synchro squeezed transform (FSST), and then few feature series are selected by neighborhood component analysis (NCA) to reduce dimension of the signal features. Finally, these new feature series selected by NCA are input into bidirectional LSTM network to train the regression model.
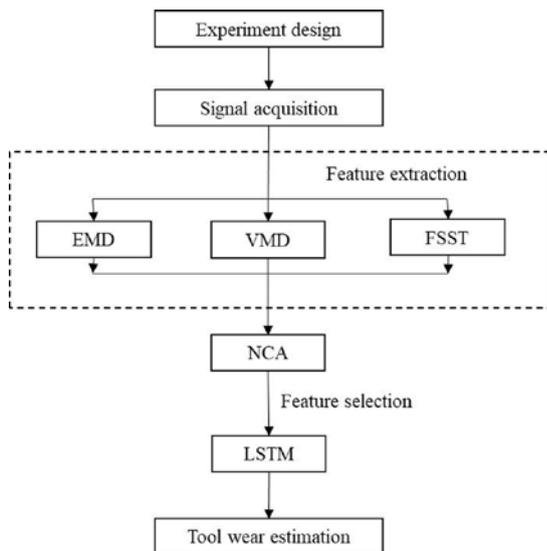

Fig. 1. Framework of the data fusion enhanced LSTM- based TCM method

2.3. Data preprocessing

For extracting more features of time series under limited samples, the collected signals are divided into multiple segments using a sliding window method. In addition, these segmented data are normalized by batch normalization method [17] as follows:

$$\mu_B = \frac{1}{m}\sum_{i=1}^{m} x_i \tag{1}$$

$$\sigma_B^2 = \frac{1}{m}\sum_{i=1}^{m}\left(x_i - \mu_B\right)^2 \tag{2}$$

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}} \tag{3}$$

where x_i and y_i denote the input and output value after batch normalization respectively, m denotes the number of inputs in minibatch, μ_B and σ_B denote the mean of input and the average variance of the input respectively, $\hat{x}_i$ is the normalized x_i.

2.3. Feature extraction

2.3.1. Empirical mode decomposition

EMD is a nonlinear time-frequency decomposition algorithm that decompose the signal into several intrinsic mode functions (IMFs) and a residual [7], shown in Equation (4). In EMD, all decomposed IMFs contain the local feature information in different time scales of the original signal. Finally, each IMF contains approximately a single frequency component, and the instantaneous frequency of the original signal can be obtained after the weighted average of the instantaneous frequency of each IMF:

$$X(t) = \sum_{i=1}^{N} \text{IMF}_i(t) + r_N(t) \tag{4}$$

EMD decomposes the signal according to the time scale features of the original data, without pre-setting any basis function, which is the most significant advantage compared with other time-frequency decomposition methods, such as wavelet transform. Due to the complexity and uncertainty of milling process, it is very difficult to find a basis function suitable for milling signal, EMD could be employed for feature extraction in milling TCM.

2.3.2. Variational mode decomposition

VMD is an adaptive time-frequency signal decomposition algorithm, its framework is the solution of variational problems [3]. VMD considers the signal is composed of sub signals with different frequencies dominant, and transforms the decomposition of signal into the solution of constrained variational model [13,28]. In this process, the central frequency and bandwidth of each IMF are updated alternately and iteratively. Finally, the signal band is decomposed adaptively and obtain the preset K narrowband IMFs in equation (5).

$$x(t) = \sum_{k=1}^{K} u_k(t) \tag{5}$$

In VMD, each IMF u_k is a bandwidth limited frequency modulation and amplitude modulation signal shown in equation (6):

$$u_k(t) = A_k(t)\cos\left(\phi_k(t)\right) \tag{6}$$

VMD has perfect mathematical theory support, its essence is an adaptive optimal Wiener filter group, which can get high signal-to-noise ratio IMFs.

2.3.3. Fourier synchro squeezed transform

Fourier synchro squeezed transform (FSST) is based on the short-time Fourier transform (SFT) implemented in the spectrogram function [12,25]. The FSST function determines the SFT of a function, f using a spectral window, g, and computing in equation (7):

$$V_g f(t,\eta) = \int_{-\infty}^{\infty} f(x)g(x-t)e^{-j2\pi\eta(x-t)}dx \qquad (7)$$

Unlike the conventional definition, this definition has an extra factor of $e^{j2\pi\eta t}$. The transform values are then "squeezed" so that they concentrate around curves of instantaneous frequency in the time-frequency plane.

2.3.4. Neighborhood component analysis

Neighborhood component analysis (NCA) is a distance metric method in metric learning and dimension reduction fields [2]. NCA is based on K-Nearest Neighborhood (KNN) including feature parameters and response label [20]. NCA selects randomly neighbors, obtains the transformation matrix in Mahalanobis distance by optimizing the results of the leave-one-out cross validation (LOOCV) method, and finds the feature parameter set maximizing the average LOO classification / regression accuracy to achieve the purpose of feature selection.

2.4. Long short-term memory network

An LSTM network is a type of RNN that can learn long-term dependencies between time steps of sequence data [6,29]. The framework of LSTM is shown in Figure 2.
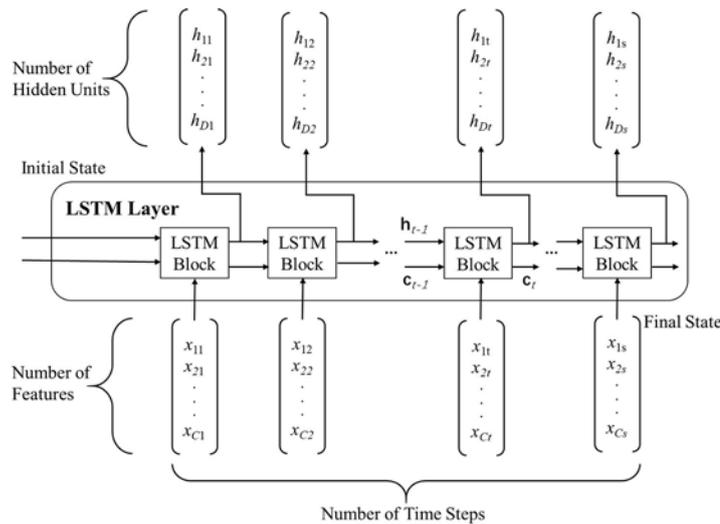

Fig. 2. LSTM architectures

Let $X_t = \{X_{1t}\ X_{2t}\ X_{Ct}\}$ is a time series with C features, $\mathbf{h}_t$ and $\mathbf{c}_t$ are the hidden state and cell state at time t, respectively. At time t, the state of the network $(\mathbf{c}_t\ \mathbf{h}_t)$ is calculated by X_t and $(\mathbf{c}_{t-1}\ \mathbf{h}_{t-1})$ by Equation (8) and (9):

$$\mathbf{c}_t = f_t \odot \mathbf{c}_{t-1} + i_t \odot g_t \qquad (8)$$

$$\mathbf{h}_t = o_t \odot \sigma_c(\mathbf{c}_t) \qquad (9)$$

The definition and expression of it ft gt ot are as shown in Table 1.

3. Experimental observation and research

3.1. Experimental setup

The experimental device for milling TCM is shown in Figure 3. In the milling TCM experiment, a CNC milling machine (DMTG VDL850A, China) is used to finish milling process, and a piece of #45 steel (30 cm ×10 cm × 8 cm) is used as the workpiece material. What's more, the milling vibration signals of spindle X and Y directions are acquired by two accelerometers with a signal acquisition device (ECON Dynamic Signal Analyzer, shown in Figure 3(b)). In addition, the signal sampling frequency in the experiment is 12KHz.

Fourteen uncoated three-insert tungsten steel end milling cutters with diameter of 10 mm are employed to mill the workpiece under different cutting conditions, listed in Table 2. For each tool, the workpiece is milled surface 10 times, and the tool wear value is measured after milling each surface using a tool microscope (GP-300C Figure 3(c)). The length of rake face wear (KB) is employed as the tool wear criterion in the experiment, and the max value $KB = max\ (KB_1\ KB_2\ KB_3)$ of three inserts is adopted as the final tool wear value. Figure 4 illustrates the tool wear conditions after milling the workpiece surface 1-st, 5-th and 10-th times.
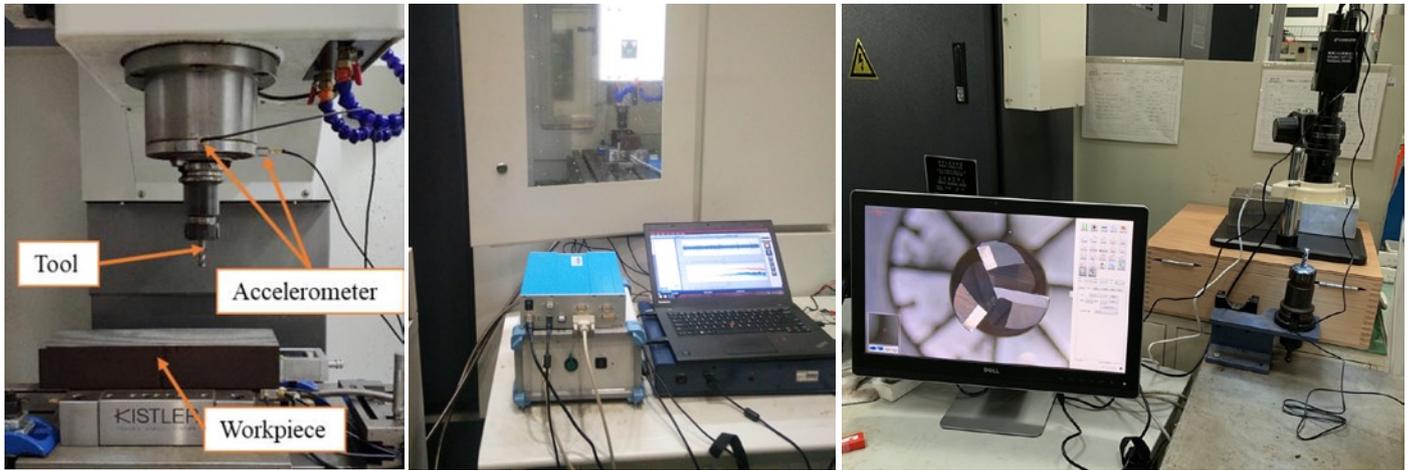
In the 14 milling TCM experiments, the training, verification and testing sets are generated randomly shown in Table 2, 7 sets of samples for training, 3 sets of samples for verification, and 4 sets of samples for testing.

3.2. Results and analysis

3.2.1. Samples and metrics

Acceleration signals of Spindle X and Y direction are used in the network, 272 training set, 120 validations set, and 80 test set are made up from spindle sensor signals. In all analyzed samples, there is no same cutting condition combination in the three dataset. Besides, in the signal pre-processing, the original signal of each sample is divided into 10 parts by slide window method, in which the window size is 2000 points, and the sliding distance is 1000 points.

To evaluate the performance of the proposed method, three indexes are employed, including the mean absolute error (MAE), root mean squared error (RMSE), and R-squared (R^2).

Table 1. Definition and expression of the LSTM layer gate

Component	Purpose	Formula
Input gate (i)	Control level of cell state update	$i_t = \sigma_g\left(W_i\mathbf{x}_t + R_i\mathbf{h}_{t-1} + b_i\right)$
Forget gate (f)	Control level of cell state reset (forget)	$f_t = \sigma_g\left(W_f\mathbf{x}_t + R_f\mathbf{h}_{t-1} + b_f\right)$
Cell candidate (g)	Add information to cell state	$g_t = \sigma_c\left(W_g\mathbf{x}_t + R_g\mathbf{h}_{t-1} + b_g\right)$
Output gate (o)	Control level of cell state added to hidden state	$o_t = \sigma_g\left(W_o\mathbf{x}_t + R_o\mathbf{h}_{t-1} + b_o\right)$

where W_t and R_t are the input weights and recurrent weight in the t-th layer, and b_k is the bias of each component.
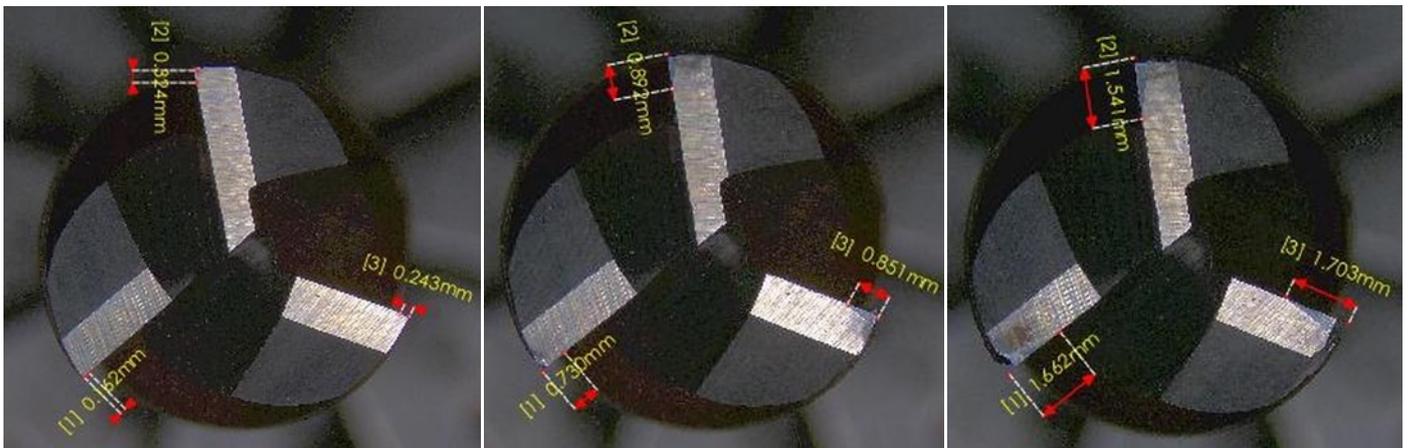
| (a) Experimental platform | (b) Data acquisition system | (c) Tool microscope |

Fig. 3. The experimental setup [16]

| (a) 1-st milling | (b) 5-th milling | (c) 10-th milling |

Fig. 4. Tool images indicative of different tool-wear values [33]

Table 2. Experimental cutting parameters

Case No	Spindle speed (rpm)	Axial cut deep (mm)	Feed speed (mm/min)	Dataset type
1	2300	0.4	400	Training
2	2300	0.5	450	Validation
3	2300	0.6	500	Testing
4	2400	0.4	450	Training
5	2400	0.5	500	Testing
6	2400	0.6	400	Validation
7	2500	0.4	500	Training
8	2500	0.5	400	Testing
9	2500	0.6	450	Training
10	2300	0.4	500	Testing
11	2300	0.6	400	Training
12	2500	0.6	500	Validation
13	2500	0.6	400	Training
14	2500	0.4	400	Training

3.2.2. Algorithm settings

For each cutting process in the experiment, there are two mutually perpendicular milling vibration signals which are collected from the equipment and a part of collected signal has 12000 points as shown in Figure 5, in which the corresponding cutting parameters is the Case 1 in Table 2: spindle speed is 2300 rpm, axial cutting depth is 4 mm, and feed rate is 400 mm/min.

Since the real monitoring signal is often nonlinear and non-stationary, it is suitable to use the EMD, VMD and FSST methods to obtain the features of vibration signals for tool wear. In order to obtain signal
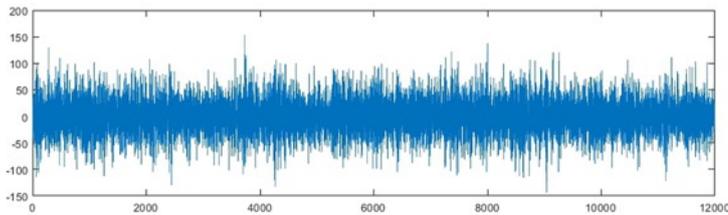
Fig. 5. Original vibration signal

Table 3. Features selection

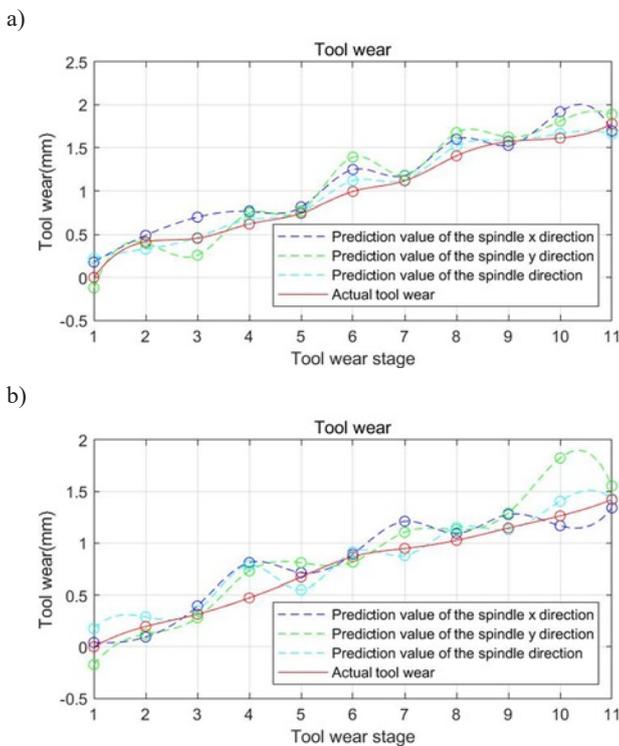| Characteristic	Fourier synchro squeezed transform(Hz)
Real part characteristic frequency	0;938;1875;2813;375;4688;5625;6563; 750;8438;9375;103130;1125;12188;13125 14063;4500(Hz)
Imaginary part characteristic frequency	938;1875;2813;375;4688;5625;6563; 750;8438;9375;10313;1125;12188;4500(Hz)

a)

b)

Fig. 6. Prediction results of tool wear: a) the 5-th tool, b) the 8-th tool

feature and more information from the vibration signal to predict the tool wear value, the original signal is transformed by EMD, VMD and FSST to expand the dimensionality. Furthermore, to remove irrelevant features and reduce the number of features, sensitive features that correlate well with tool wear are selected out through NCA.

In this work, the first 6 IMFs and residualsare taken in EMD, the first 5 IMFs and residualsare taken in VMD, and 60 IMFs is decomposed in FSST. In addition, it is necessary to take the real and imaginary parts of the IMF as the feature matrix of the vibration signal, and use NCA to take the effective characteristic matrix. The results are listed in the Table 3

By calculating, it was found that the feature matrix has first 6 numbers of IMFs and residuals of EMD, 5 numbers of IMFs and residuals of VMD, 17 real parts and 14 imaginary parts of FSST. Totally 45 feature matrixes as input signals. The two single-channel experimental data of the sensor are superimposed and fused into a new sample. Meanwhile, all data from experiments need batch normalization.

In this model, it is a way to use eleven layers as neural network architectures in our experiments: especially bidirectional LSTM layer, which has two hidden LSTM layers (forwards and backwards) as shown in Table 4.

Due to the limitations of experimental equipment conditions and cost, 14 sets of experiments were executed, 7 sets of samples under different working conditions were selected for training, 3 sets of samples were selected for verification, and 4 sets of samples were selected for testing.

For all architectures, complete error gradient was calculated and the weights are trained by using gradient descent with momentum. In all experiments, the same training parameters were kept: randomly assigned initial weights, keeping the training algorithm and parameters constant, allowing us to focus on the impact of changing the architecture.

3.2.3. Experimental results

The LSTM model established by the training set and verification set is applied to predict the testing set, including 4 tools with different cutting conditions. In Figure 6, the blue, green, and cerulean dotted lines denote the prediction results using the proposed method with the spindle vibration signal of X-direction, Y-direction, and dual-direction (composition of X and Y directions). It is noted that the cutting parameters of the 5-th and 8-th tools are different. For Figure 6(a), the spindle speed is 2400 rpm, the axial cutting depth is 0.6 mm, and the feed rate is 500 mm/min. For Figure 6(b), the spindle speed is 2500 rpm, the axial cutting depth is 0.5 mm, and the feed rate is 400 mm/min. It can be seen that the trend of the overall predicted value is similar the actual wear value, and the error at some stages is less than 0.1 or even close to the wear value.

To test the regression performance, the proposed method is compared with RNN and support vector machine (SVR). As a result, the MSE, RMSE and R^2 of three methods are presented in Table 5.

It can be seen from Table 5 that the proposed LSTM- based method is highly effective in improving the regression accuracy, the prediction accuracy of the proposed method is much higher than that of RNN and SVR according to the values of three evaluation indexes, except for the X- direction signal of the 3-rd and 5-th tools. In addition, the prediction accuracies with the dual- direction signal outperform that of signal- direction except for the 3-rd tool, while the results of three indexes are slightly worse than that of two other methods in the 3-rd tool.

Table 4. Network architectures

Serial number	Name	Type	Serial number	Name	Type
1	sequenceinput	Sequence input	7	dropout_2	20% dropout
2	biLSTM_1	BILSTM：300 hidden units	8	fc_1	1 fully connected layer
3	relu_1	ReLU	9	dropout_3	20% dropout
4	dropout_1	20% dropout	10	fc_2	1 fully connected layer
5	biLSTM_2	BILSTM：300 hidden units	11	Regression output	mean-squared-error: Response
6	relu_2	ReLU			

Table 5. *Prediction results of LSTM and RNN and SVR*

Vibration signal	Tool	MAE			RMSE			R−Squared		
		LSTM	SVR	RNN	LSTM	SVR	RNN	LSTM	SVR	RNN
Spindle X- direction	3	0.2150	0.5091	**0.1222**	0.2572	0.6160	**0.1490**	0.7996	−0.1498	**0.9328**
	5	0.1508	0.5518	**0.1127**	0.1741	0.6443	**0.1396**	0.9091	−0.2456	**0.9415**
	8	**0.1153**	0.4822	0.2507	**0.1483**	0.5824	0.2669	**0.8971**	−0.5871	0.6667
	10	**0.1762**	0.4686	0.3688	**0.2063**	0.6246	0.4857	**0.8039**	−0.7978	−0.0868
Spindle Y- direction	3	0.2829	0.3989	**0.2572**	**0.3327**	0.4891	0.4811	**0.6647**	0.2829	0.3989
	5	**0.1437**	0.6037	0.2458	**0.1813**	0.7576	0.4740	**0.9014**	0.1437	0.6037
	8	**0.1661**	0.5130	0.6122	**0.2159**	0.7089	0.7402	**0.7819**	0.1661	0.5130
	10	**0.1681**	0.5639	0.6996	**0.1985**	0.6385	0.7295	**0.8184**	0.1681	0.5639
Spindle dual- direction	3	**0.2413**	0.4406	1.0053	**0.2944**	0.5219	1.1609	**0.7373**	0.1749	−0.3083
	5	**0.0738**	0.5407	1.0043	**0.0974**	0.6101	1.2031	**0.9715**	−0.1169	−0.3344
	8	**0.1031**	0.3976	0.9307	**0.1332**	0.5097	1.0576	**0.9169**	−0.2158	−0.4234
	10	**0.1404**	0.5002	0.778	**0.1691**	0.5853	0.9284	**0.8683**	−0.5786	−0.2971

4. Conclusion

This paper proposed a novel method based on data fusion enhanced LSTM to estimate tool wear value under different cutting conditions. Firstly, the original vibration signals are decomposed and transformed to obtain high-dimensional feature series set through EMD, VMD and FSST, and then NCA is employed to select useful features and reduce the feature dimension, in order to reduce operational burden and improve the accuracy of regression. Finally, these selected feature series are input into bidirectional LSTM network to estimate tool wear value. Hence, applications of the proposed method to milling TCM experiments demonstrate it outperforms significantly SVR- based and RNN- based methods under different cutting conditions.

References

1. Cao XC, Chen BQ, Yao B, He WP. Combining translation-invariant wavelet frames and convolutional neural network for intelligent tool wear state identification. Computers in Industry 2019; 106: 71-84, https://doi.org/10.1016/j.compind.2018.12.018.
2. Dong W, Tan X. Bayesian Neighborhood Component Analysis. IEEE Transactions on Neural Networks & Learning Systems 2018;29(7): 3140-3151, https://doi.org/10.1109/TNNLS.2017.2712823.
3. Dragomiretskiy K, Zosso D. Variational mode decomposition. IEEE Transactions on Signal Processing 2014; 62(3): 531-544, https://doi.org/10.1109/TSP.2013.2288675.
4. He K, Gao M, Zhao Z. Soft ComputingT echniques for Surface Roughness Prediction in Hard Turning: A LiteratureReview. IEEE Access 2019; 7: 89556-89569, https://doi.org/10.1109/ACCESS.2019.2926509.
5. Hochreiter S, Schmidhuber J. Long short-term memory. Neural Computation 1997; 9(8): 1735-80, https://doi.org/10.1162/neco.1997.9.8.1735.
6. Hua YS, Mou LC, Zhu XX. Recurrently exploring class-wise attention in a hybrid convolutional and bidirectional LSTM network for multi-label aerial image classification. Isprs Journal of Photogrammetry And Remote Sensing 2019; 149: 188-199, https://doi.org/10.1016/j.isprsjprs.2019.01.015.
7. Huang NE, Shen Z, Long SR, Wu MC, Shih HH,Zheng Q. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings Mathematical Physical & Engineering Sciences 1998, 454(1971): 903-995, https://doi.org/10.1098/rspa.1998.0193.
8. Huang Z, Zhu J, Lei J, Li X, Tian F. Tool wear predicting based on multi-domain feature fusion by deep convolutional neural network in milling operations. Journal of Intelligent Manufacturing 2020;31(4): 953-966, https://doi.org/10.1007/s10845-019-01488-7.
9. Jasiulewicz-Kaczmarek M, Antosz K, Żywica P,Mazurkiewicz D, Sun B, Ren Y. Framework of machine criticality assessment with criteria interactions. Eksploatacja i Niezawodnosc - Maintenance andReliability 2021; 23(2): 207-220, https://doi.org/10.17531/ein.2021.2.1.
10. Kozlowski E, Mazurkiewicz D, Zabinski T,Prucnal S, Sep J. Assessment model of cutting tool condition for real-time supervision system. Eksploatacja i Niezawodnosc - Maintenance and Reliability2019; 21(4): 679-685, https://doi.org/10.17531/ein.2019.4.18.
11. Kozłowski E, Mazurkiewicz D, Żabiński T,Prucnal S, Sęp J. Machining sensor data management for operation-level predictive model. Expert Systems with Applications 2020; 159: 1-22, https://doi.org/10.1016/j.eswa.2020.113600.
12. Kumar A, Gandhi CP, Zhou YQ, Kumar R, XiangJW. Improved CNN for the diagnosis of engine defects of 2-wheeler vehicle using wavelet synchro-squeezed transform (WSST). Knowledge based System 2020; 208,106453, https://doi.org/10.1016/j.knosys.2020.106453.
13. Kumar A, Gandhi CP, Zhou YQ, Kumar R, XiangJW. Variational mode decomposition based symmetric single valued neutrosophiccross entropy measure for the identification of bearing defects in acentrifugal pump. Applied Acoustics 2020; 165, 107294, https://doi.org/10.1016/j.apacoust.2020.107294.
14. Kumar A, Kumar R. Adaptive artificial intelligence for automatic identification of defect in the angular contact bearing. Neural Computing & Applications 2018; 29: 277-287, https://doi.org/10.1007/s00521-017-3123-4.
15. Lei Z, Zhou YQ, Sun BT, Sun WF. An intrinsic time scale decomposition-based kernel extreme learning machine method to detect tool wear conditions in the milling process. International Journal of Advanced Manufacturing Technology 2020; 106(3-4): 1203-1212, https://doi.org/10.1007/s00170-019-04689-9.
16. Lei Z, Zhu QS, Zhou YQ, Sun BT, Sun WF, PanXM. A GAPSO-Enhanced Extreme Learning Machine Method for Tool Wear Estimationin Milling Processes Based on Vibration Signals. International Journal ofPrecision Engineering and Manufacturing- Green Technology 2021;

8: 745-759, https://doi.org/10.1007/s40684-021-00353-4.

17. Liu M, Wu W, Gu Z, Yu Z, Qi F, Li Y. Deep learning based on Batch Normalization for P300 signal detection. Neurocomputing 2018; 275: 288-297, https://doi.org/10.1016/j.neucom.2017.08.039.

18. Ma X, Tao Z, Wang Y, Yu H, Wang Y. Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. Transportation Research Part C- Emerging Technologies2015; 54: 187-197, https://doi.org/10.1016/j.trc.2015.03.014.

19. Musavi SH, Davoodi B, Eskandari B. Evaluation of surface roughness and optimization of cutting parameters in turning of AA2024 alloy under different cooling-lubrication conditions using RSM method. Journal of Central South University 2020, 27(6): 1714- 1728, https://doi.org/10.1007/s11771-020-4402-2.

20. Raghu S, Sriraam N. Classification of focaland non-focal EEG signals using neighborhood component analysis and machine learning algorithms. Expert Systems with Applications 2018; 113: 18-32, https://doi.org/10.1016/j.eswa.2018.06.031.

21. Rosienkiewicz M. Artificial intelligence-based hybrid forecasting models for manufacturing systems. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2021; 23 (2): 263-277, https://doi.org/10.17531/ein.2021.2.6.

22. Sedlacek M, Podgornik B, Vizintin J. Influence of surface preparation on roughness parameters friction and wear. Wear 2009; 266(3-4): 482-487, https://doi.org/10.1016/j.wear.2008.04.017.

23. Siddhpura A, Paurobally R. A review of flankwear prediction methods for tool condition monitoring in a turning process. International Journal of Advanced Manufacturing Technology 2013; 65(1-4):371-393, https://doi.org/10.1007/s00170-012-4177-1.

24. Tao Z, An Q, Liu G, Chen M. A novel methodfor tool condition monitoring based on long short-term memory and hidden Markov model hybrid framework in high-speed milling Ti-6Al-4V. International Journalof Advanced Manufacturing Technology 2019; 105(7-8): 3165-3182, https://doi.org/10.1007/s00170-019-04464-w.

25. Tary JB, Herrera RH, Baan MV. Analysis of time-varying signals using continuous wavelet and synchrosqueezed transforms. Philosophical Transactions of the Royal Society A 2018; 376(2126), 20170254, https://doi.org/10.1098/rsta.2017.0254.

26. Tim VH, Chris KM. Self-supervised learning for tool wear monitoring with a disentangled-variational- autoencoder. International Journal of Hydromechatronics 2021; 4(1): 69-98, https://doi.org/10.1504/IJHM.2021.114174.

27. Wang B, Liu Z. Influences of tool structure tool material and tool wear on machined surface integrity during turning and milling of titanium and nickel alloys: a review. International Journal of Advanced Manufacturing Technology 2018; 98(5-8): 1925-1975, https://doi.org/10.1007/s00170-018-2314-1.

28. Zhang X, Zhao J. Compound fault detection in gearbox based on time synchronous resample and adaptive variational mode decomposition. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020;22(1): 161-169, http://dxdoiorg/1017531/ein2020119. https://doi.org/10.17531/ein.2020.1.19

29. Zhao R, Yan R, Wang J, Mao K. Learning to Monitor Machine Health with Convolutional Bi-Directional LSTM Networks. Sensors 2017; 17(2), 273, https://doi.org/10.3390/s17020273.

30. Zhi GF, He DD, Sun WF, Zhou YQ, Pan XM, Gao C. An edge-labeling graph neural network method for tool wear condition monitoring using wear image with small samples. Measurement Science and Technology 2021; 32(6), 064006, https://doi.org/10.1088/1361-6501/abe0d9.

31. Zhou JT, Zhao X, Gao J. Tool remaining usefullife prediction method based on LSTM under variable working conditions. International Journal of Advanced Manufacturing Technology 2019; 104(9-12): 4715-4726, https://doi.org/10.1007/s00170-019-04349-y.

32. Zhou YQ, Sun BT, Sun WF. A tool condition monitoring method based on two-layer angle kernel extreme learning machine and binary differential evolution for milling. Measurement 2020; 166: 108186, https://doi.org/10.1016/j.measurement.2020.108186.

33. Zhou YQ, Sun BT, Sun WF, Lei Z. Tool wear condition monitoring based on a two-layer angle kernel extreme learning machine using sound sensor for milling process. Journal of Intelligent Manufacturing 2020, https://doi.org/10.1007/s10845-020-01663-1.

34. Zhou YQ, Xue W. Review of tool condition monitoring methods in milling processes. International Journal of Advanced Manufacturing Technology 2018; 96(5-8): 2509-2523, https://doi.org/10.1007/s00170-018-1768-5.

35. Zhu QS, Zhou YQ, Sun BT, He DD, Sun WF. A tool wear condition monitoring approach for end milling based on numerical simulation. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2021; 23(2):371-380, https://doi.org/10.17531/ein.2021.2.17.