HIGH TECHNOLOGY LETTERS | Vol. 29 No. 1 | Mar. 2023 | pp. 12-21

doi:10.3772/j.issn.1006-6748.2023.01.002

Research on mechanical wear life feature fusion prediction method based on temporal pattern attention mechanism①

JIANG Zhinong (江志农)[*,**], CHEN Yuyang[*,**], ZHANG Jinjie②[*,**], LI Zhaoyang[*,**], MAO Zhiwei[*,**], ZHI Haifeng[***], LIU Fengchun[***]

([*] Beijing Key Laboratory of Health Monitoring and Self-Healing of High-end Mechanical Equipment, Beijing University of Chemical Technology, Beijing 100029, P. R. China)

([**] Key Laboratory of Engine Health Monitoring-control and Networking, Ministry of Education, Beijing University of Chemical Technology, Beijing 100029, P. R. China)

([***] China North Engine Research Institute (Tianjin), Tianjin 300400, P. R. China)

Abstract

In order to solve the problem of low prediction accuracy when only vibration or oil signal is used to predict the remaining life of gear wear, a gear wear life feature fusion prediction method based on temporal pattern attention mechanism is proposed. Firstly, deep residual shrinkage network (DRSN) is used to extract the features of the original vibration time series signals with low signal-to-noise ratio, and the vibration features associated with gear wear evolution are obtained. Secondly, the extracted vibration features and the oil monitoring data that can intuitively reflect the wear process information are jointly input into the bi-directional long short-term memory neural network based on temporal pattern attention mechanism (TPA-BiLSTM), the complex nonlinear relationship between vibration features, oil features and gear wear process evolution is further explored to improve the prediction accuracy. The gear life cycle dynamic response and wear process signals are obtained based on the gear numerical simulation model, and the feasibility of the proposed method is verified. Finally, the proposed method is applied to the residual life prediction of gear on a test bench, and the comparison between different methods proved the validity of the proposed method.

Key words: prediction of gear remaining useful life, information fusion, numerical simulation, neural network, oil monitoring

0 Introduction

Gearbox is a power transmission device with a wide range of applications[1]. Due to long-term complex working conditions of heavy load and variable load and harsh working environment, excessive alternating stress on the gear tooth mating surface in the gearbox will cause rolling contact fatigue on the contact surface[2] and wear failure. In practical engineering application, once the gearbox teeth breaks, such as teeth and shaft breakage, it will lead to serious economic loss and even casualties. Therefore, real-time monitoring of gear operation status and early warning of wear faults are of great significance[3-4].

At present, most scholars have carried out gear wear state assessment and fault diagnosis based on oil monitoring parameters. Diagnosis methods can be roughly divided into three categories: methods based on linear regression, methods based on gray theory, and methods based on machine learning[5-8]. Sejkorova et al.[5] proposed a method combining partial least squares (PLS) and principal component regression (PCR) to accurately predict the viscosity of worn oil samples. Li et al.[6] applied the improved Euler algorithm to solve the grey theory model and obtained the change trend curve of Fe mass concentration in lubricating oil, which successfully predicted the upcoming wear failure of the transmission device. However, the prediction method based on linear regression and grey theory is suitable for trend prediction based on small sample information under offline oil monitoring, and the prediction accuracy is limited by the defect of sample size and method itself.

① Supported by the National Natural Science Foundation of China (No. 52101343), the Aeronautical Science Foundation (ASFC) (No. 201834S9002) and Chongqing Technology Innovation and Application Development Special General Project (No. cstc2020jscx-msxm0411).

② To whom correspondence should be addressed. E-mail: zjj87427@163.com.

Received on June 10, 2022

With the further development of online oil monitoring devices and machine learning methods, the traditional fault diagnosis methods based on oil parameters can no longer meet the practical application requirements, and the machine learning methods represented by neural networks are gradually applied to the wear state assessment of various equipment. David et al.[7] combined fuzzy inference system (FIS) and artificial neural network (ANN) to effectively predict the development trend of Fe particle concentration and soot particle concentration in lubricating oil. Bazi et al.[8] proposed a tool wear prediction model of integration of variational mode decomposition (VMD), convolutional neural network (CNN) and bi-directional long short-term memory neural network (BiLSTM), which was better than the traditional method. Although oil monitoring can directly reflect the evolution process of gear tooth surface wear, there are problems such as signal delay and difficulty in fault tracing, so vibration and oil monitoring technology are often used to monitor the real-time condition of gear box in practical engineering applications[9]. Furthermore, most of the researches focus on gearbox fault diagnosis methods, but there are few reports on the prediction of residual life of gear wear. Therefore, it is of practical significance to study the prediction method of gear wear residual life based on vibration and oil features.

A gear wear residual life prediction method based on DRSN-TPA-BiLSTM is proposed, which integrated vibration and oil features. Deep residual shrinkage network (DRSN) is used to extract the vibration signal of wear and tear associated features, the vibration and oil signals are fused, and the nonlinear complex relationship between vibration, oil features and gear wear evolution is further explored by using the bi-directional long short-term memory neural network based on temporal pattern attention mechanism (TPA-BiLSTM) to extract sensitive features. Based on the numerical simulation model of gear and the signal of experimental bench, the proposed method is tested and verified, which proves the feasibility and advantages of the proposed method.

1 Prediction model of vibration and oil feature fusion based on temporal pattern attention mechanism

The feature fusion prediction model based on temporal pattern attention (TPA) mechanism is mainly composed of feature fusion and life prediction. Due to slow and gradual gear wear process and strong background noise, vibration acceleration signals have weak and slow variation features, which makes it difficult to extract vibration features. As a result, vibration signals are firstly input into BatchNormalization layer and then into the residual shrinkage building unit with channel-wise thresholds (RSBU-CW) after a Conv layer. Then Maxpooling is carried out to complete feature extraction of vibration signals. RSBU is the core of DRSN[10]. RSBU-CW is a threshold sharing residual shrinkage unit, which introduces soft threshold into the network structure of ResNet as a nonlinear layer to improve the feature extraction effect of deep learning method on noisy data or complex data. Secondly, TPA is introduced into BiLSTM, and advanced features are extracted by integrating vibration and oil signals, so as to achieve gear wear life prediction. The structure of gear wear life prediction network based on temporal pattern attention mechanism proposed in this paper is shown in Fig. 1.

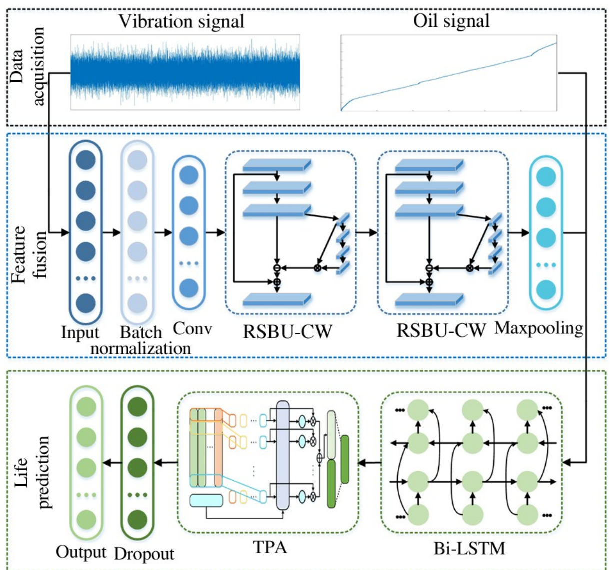

Fig. 1 Fusion prediction model based on temporal pattern attention mechanism

1.1 Soft threshold function

The traditional ReLU activation function sets all the signals less than 0 to 0, resulting in all the information in the negative signal being discarded, which is prone to the problem of incomplete feature extraction. However, the soft threshold function only sets the signal close to 0 to 0, which can retain the useful information in the negative signal, extract the signal feature to the maximum extent, and realize the denoising of vibration signal. RSBU-CW unit introduces the residual term based on CNN, adaptively learns the threshold through the subnetwork, and softs the characteristic threshold to achieve the denoising effect of the input signal. RSBU-CW structure is shown in Fig. 2, where,

C represents the number of channels, W represents the signal width, K represents the number of convolution kernels, and M represents the number of neurons at the full connection layer. Threshold softening is a classical method commonly used in signal denoising. The soft threshold function formula can be expressed as

$$y = \begin{cases} x - \tau & x > \tau \\ 0 & -\tau \leqslant x \leqslant \tau \\ x + \tau & x < -\tau \end{cases} \tag{1}$$

where, x represents input feature, y represents the output feature, τ indicates the threshold ($\tau > 0$).

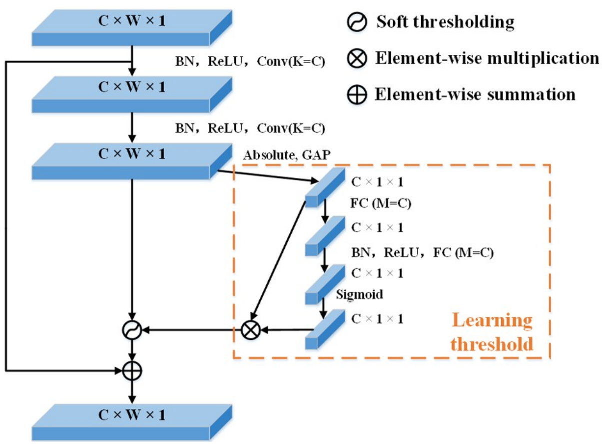

Fig. 2 Residual shrinkage building unit with channel-wise thresholds (RSBU-CW)

1.2 Temporal pattern attention mechanism

It is easy to cause low prediction accuracy when only using vibration signal to predict the remaining life of gear, and the number of ferromagnetic particles in lubricating oil can directly reflect the evolution process of gear wear. Therefore, the extracted vibration features and oil data are input into BiLSTM network based on TPA[11] to further explore the nonlinear and complex relationship between vibration, oil features and gear wear evolution process and obtain advanced fusion features. Traditional deep convolutional neural networks tend to convolve in one direction when extracting features, and can only extract feature information in time dimension, ignoring the nonlinear relationship between input features. After the combination of vibration and oil features, TPA convolves the different features of vibration and oil first, and then convolves the time series sequence to deeply extract the features of space-time and space domains. The constructed TPA network unit is shown in Fig. 3.

After the vibration features and oil information are transmitted through BiLSTM network, the hidden state matrix $\boldsymbol{H} = [\boldsymbol{h}_{t-w}, \boldsymbol{h}_{t-w+1}, \boldsymbol{h}_{t-1}]$ is obtained, where w is the length of the time series. Convolving m features from top to bottom along $\boldsymbol{H}$ matrix, the time characteristic matrix $\boldsymbol{H}^C$ is obtained as

$$\boldsymbol{H}_{i,j}^{C} = \sum\nolimits_{l=1}^{w} \boldsymbol{H}_{i,(t-w-1+l)} * \boldsymbol{C}_{j,\,T-w+l} \tag{2}$$

where, $\boldsymbol{C}_j$ represents the j-th filter; T represents the maximum characteristic length that needs attention, and its value is usually w; $*$ represents the convolution operation. The number of convolution kernels of the filter is k, and the filter is convolved along the row vectors of the hidden state matrix $\boldsymbol{H}$.

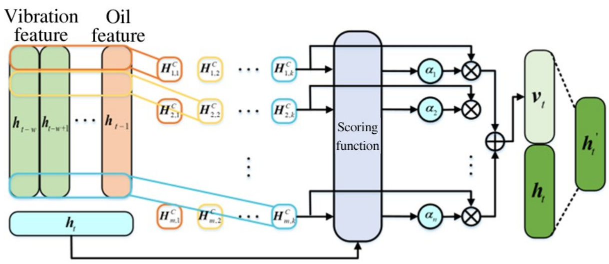

Fig. 3 Temporal pattern attention mechanism unit (TPA)

The correlation of feature information is calculated through the attention mechanism, and the scoring function is as follows.

$$f(\boldsymbol{H}_i^C, \boldsymbol{h}_t) = (\boldsymbol{H}_i^C)^T \boldsymbol{W}_a \boldsymbol{h}_t \tag{3}$$

$$\boldsymbol{\alpha}_i = \mathrm{Sigmoid}(f(\boldsymbol{H}_i^C, \boldsymbol{h}_i)) \tag{4}$$

where, $\boldsymbol{W}_a$ is the weight matrix and $\boldsymbol{\alpha}_i$ is the weight of attention. Define attention vector:

$$\boldsymbol{v}_t = \sum\nolimits_{i=1}^{n} \boldsymbol{\alpha}_i \boldsymbol{H}_i^C \tag{5}$$

where, n represents the total number of features of vibration and oil.

Predicted values are obtained by linear mapping:

$$\boldsymbol{y} = \boldsymbol{W}_{h'}(\boldsymbol{W}_h \boldsymbol{h}_t + \boldsymbol{W}_v \boldsymbol{v}_t) \tag{6}$$

where, $\boldsymbol{y}$ represents the predicted value, $\boldsymbol{W}_{h'}$, $\boldsymbol{W}_h$ and $\boldsymbol{W}_v$ represent the weight matrices of variables $\boldsymbol{h}'$, $\boldsymbol{h}$ and $\boldsymbol{v}$, respectively.

2 Numerical calculation model of gear considering dynamic wear

2.1 Calculation model of tooth surface wear of spur gear

In order to verify the feasibility of DRSN-TPA-BiLSTM based vibration and oil information fusion prediction model, the simulation vibration response and simulation wear accumulation of gear are obtained by constructing a numerical calculation model of gear considering dynamic wear.

Calculate tooth surface wear according to Archard wear formula[12]:

$$\frac{\mathrm{d}h}{\mathrm{d}s} = kp \tag{7}$$

where, h is the wear depth, s is the relative sliding distance, k is the wear coefficient, p is the Hertz contact pressure. The wear depth of any meshing point a can be obtained by integrating the relative sliding distance:

$$h_{w_a} = \int_0^{s_a} k_w p_a \mathrm{d}s_a \tag{8}$$

where, h_{w_a} is the wear depth of point a, k_w is the slip distance of point a, and p_a is the contact pressure at point a. According to the single point observation method[13], the accumulated wear depth of any meshing point a on the tooth surface after $(n+1)$ meshing cycles is as follows.

$$h_{w_a,(n+1)} = h_{w_a,(n)} + k_w p_{a,(n)} S_a \tag{9}$$

By integrating the wear depth of each meshing point along the tooth profile, the cumulative wear volume of a single tooth can be obtained as

$$V = \int_{x_1}^{x_N} h_{w_i} \mathrm{d}x_i \tag{10}$$

The contact mode of involute cylindrical spur gear is usually simplified as equivalent cylinder contact with time-varying radius, and the curvature radius and comprehensive curvature radius of the equivalent cylinder can be calculated as follows.

$$\rho_1 = r_1 \sin\alpha + y,\ \rho_2 = r_2 \sin\alpha - y \tag{11}$$

$$\frac{1}{\rho} = \frac{1}{\rho_1} + \frac{1}{\rho_2} \tag{12}$$

$$y = \omega_2 r'_2 (t - t_0) \tag{13}$$

where, ρ_1, ρ_2 are the radius of curvature of the driving and driven gears respectively; ρ is the radius of composite curvature; r_1, r_2 are the radius of the indexing circle of the driving and driven gears respectively; r'_2 is the radius of the base circle of the driven gears; ω_2 is the rotation speed of the driven gear; α is the pressure angle; y is the distance between the meshing point and the node of the driving and driven gears; t_0 is the time from the meshing point to the node. Generally, $t_0 = 0.5t_p$, t_p is a meshing cycle. The coiling speed on the tooth profile surface is expressed as follows[14].

$$U = \frac{U_1 + U_2}{2} \tag{14}$$

$$U_1 = \omega_1 \rho_1,\ U_2 = \omega_2 \rho_2 \tag{15}$$

According to Hertz contact theory, contact half-width and contact pressure at meshing point a are as follows.

$$a_H = \sqrt{\frac{4F_a \rho}{\pi b E^*}} \tag{16}$$

$$p_a = \frac{2F_a}{\pi b a_H^2} (a_H - y_i^2)^{\frac{1}{2}} \tag{17}$$

where, F_a is the normal meshing force; b is the tooth width; E^* is the equivalent elastic modulus, and its calculation formula is

$$\frac{1}{E^*} = \frac{1 - v_1^2}{E_1} + \frac{1 - v_2^2}{E_2} \tag{18}$$

where, E_1, E_2, v_1, v_2 are the elastic modulus and Poisson's ratio of driving gear and driven gear respectively.

2.2 A dynamic model of single-stage gear considering dynamic wear

In this paper, a single-stage gear dynamics model is constructed based on the lumped mass method, and the dynamics equation of gear train is obtained by Newton's laws of mechanics[15]:

$$\begin{aligned} I_1 \ddot{\theta}_1 &+ c_m (R_1 \dot{\theta}_1 - R_2 \dot{\theta}_2 - \dot{e}(t)) R_1 \\ &+ k(t) f(R_1 \theta_1 - R_2 \theta_2 - e(t)) R_1 = T_1 \end{aligned} \tag{19}$$

$$\begin{aligned} I_2 \ddot{\theta}_2 &- c_m (R_1 \dot{\theta}_1 - R_2 \dot{\theta}_2 - \dot{e}(t)) R_2 \\ &- k(t) f(R_1 \theta_1 - R_2 \theta_2 - e(t)) R_2 = -T_2 \end{aligned} \tag{20}$$

where, R_1, R_2 are the base circle radius of driving gear and driven gear respectively; I_1, I_2 are the moment of inertia; θ_1, θ_2 are torsional displacement; T_1, T_2 are the torques acting on the driving and driven gears respectively; c_m is meshing damping; $k(t)$ is time-varying mesh stiffness; $e(t)$ is the static transfer error. Taking the gear pair backlash as $2b$, $f(\cdot)$ represents the backlash function:

$$f(x) = \begin{cases} x - b & x > b \\ 0 & -b \leqslant x \leqslant b \\ x + b & x < -b \end{cases} \tag{21}$$

Suppose that the displacement of the master-slave driven wheel on the meshing line is x_1 and x_2, then $x_1 = R_1\theta_1$, $x_2 = R_2\theta_2$, Eqs(19) and (20) can be rewritten as

$$\begin{aligned} m_1 \ddot{x}_1 &+ c_m (\dot{x}_1 - \dot{x}_2 - \dot{e}(t)) \\ &+ k(t) f(x_1 - x_2 - e(t)) = F_1 \end{aligned} \tag{22}$$

$$\begin{aligned} m_2 \ddot{x}_2 &- c_m (\dot{x}_1 - \dot{x}_2 - \dot{e}(t)) \\ &- k(t) f(x_1 - x_2 - e(t)) = -F_2 \end{aligned} \tag{23}$$

where, $m_1 = Ip/R_1^2$, $m_2 = Ip/R_2^2$ are equivalent masses of the master and slave driven wheels respectively; $F_1 = T_1/R_1$, $F_2 = T_2/R_2$ respectively represent the meshing forces of the main driving wheel. Due to the existence of backlash, the dynamic transmission error of the gear transmission system is assumed to be $\delta(t) = x_1 - x_2 - e(t)$, it can be obtained by combining Eqs(22) and (23):

$$m\ddot{\delta} + c_m \dot{\delta} + k(t) f(\delta) = F(t) \tag{24}$$

where, $m = \dfrac{m_1 m_2}{m_1 + m_2}$ is the equivalent mass of the gear pair, $F(t) = \dfrac{m}{m_1} F_1 + \dfrac{m}{m_2} F_2 - m\ddot{e}(t)$.

Due to the large difference in the order of magnitude between the stiffness and damping terms, in order to avoid instability in the solution, the dynamics equation is dimensionlessly processed, and the sorted dynamics equation is as follows.

$$\ddot{x} + 2\xi\dot{x} + k(\tau)f(x) = f_m + f_h \tag{25}$$

where, $\frac{k_m}{m_c} = \omega_n^2$, $2\xi = \frac{c_m}{\omega_n m_c}$, $t = \omega_n \tau$, $x = \frac{\delta}{b}$, $\omega_e = \frac{\omega}{\omega_n}$, $f_m = \frac{F}{\omega_n^2 b}$, $f_h = \frac{\ddot{e}(\tau)}{\omega_n^2 b}$.

Considering tooth surface wear, the change of wear depth is introduced into the dynamic equation, then the static transfer error becomes $e(t) + h(t)$, and the clearance function[16] considering wear is

$$f(x) = \begin{cases} x - (b + h(t)) & x > b + h(t) \\ 0 & -b - h(t) \leqslant x \leqslant b + h(t) \\ x + (b + h(t)) & x < -b - h(t) \end{cases} \tag{26}$$

3 Gear wear life prediction based on simulation data

3.1 Analysis of numerical simulation results

According to the gear numerical calculation model considering dynamic wear in Section 2, the wear accumulative amount[17] and dynamic transmission error of the whole life cycle of the gear are simulated respectively. It is assumed that the gear is in all-weather operation with a life of 100 d. The simulation parameters of the gear system are shown in Table 1.

Table 1 Gear system simulation parameters

Parameter	Driving gear	Driven gear
Number of teeth	55	75
Pressure angle/°	20	20
Module/mm	2	2
Tooth width/mm	20	20
Roughness/μm	0.2	0.2
Addendum coefficient	1	1
Backlash in circular tooth/mm	0.1	0.1
Poisson's ratio	0.3	0.3

Fig. 4 shows the simulation results of single tooth profile wear depth under three different meshing cycles. As the number of meshing cycles increases, the wear depth increases, and the simulation results of tooth surface wear accumulative amount in the whole life cycle of the gear system are shown in Fig. 5. The dynamic equation is solved by using the Runge-Kutta method with variable step length. Fig. 6 shows the dynamic transmission error at a certain operating time, and Fig. 7 shows the dynamic transmission error comparison under six wear degrees from slightness to severeness. With the deepening of wear degree, the dynamic transmission error keeps increasing. Fig. 8 shows the simulation results of gear system's life-cycle dynamic response.

Considering that gear degradation is not obvious at the early stage of operation, piecewise linear function[18] is adopted as the residual life degradation curve. As can be seen from Fig. 7, significant degradation begins to occur after 25 d of operation, so the RUL curve can be expressed as

$$y_{RUL}^{i} = \begin{cases} 75 & x_i < k_i \\ 75 - (x_i - k_i) & x_i \geqslant k_i \end{cases} \tag{27}$$

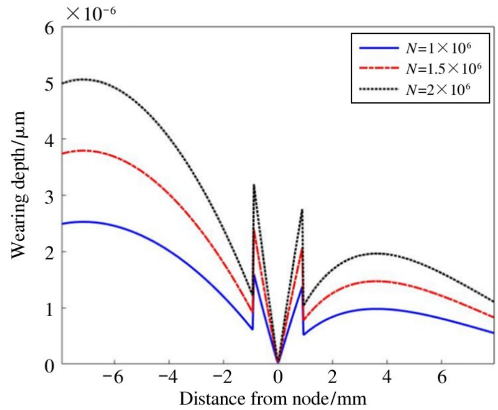

Fig. 4 Wearing depth of single tooth profile

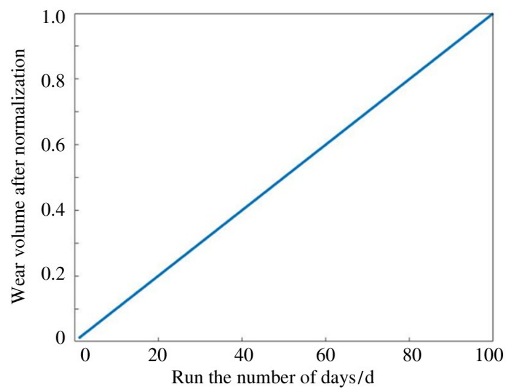

Fig. 5 Cumulative amount of tooth surface wear in the whole life cycle

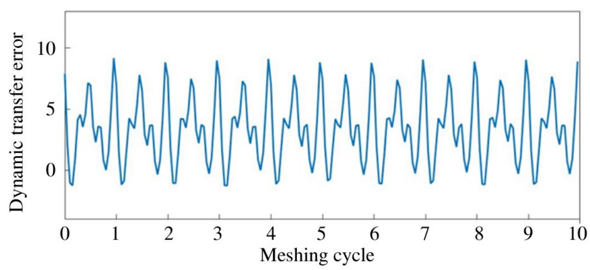

Fig. 6 Dynamic transmission error at a given running time

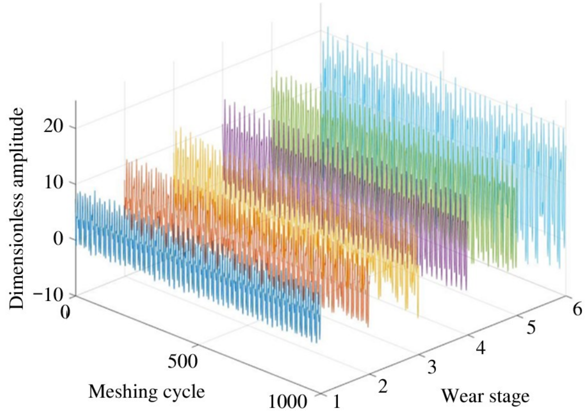

Fig. 7 Comparison of dynamic transmission error under different wear degrees

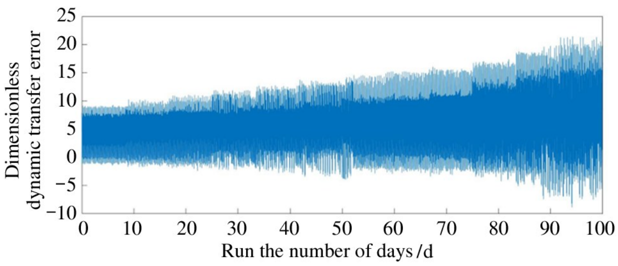

Fig. 8 Dynamic transmission error throughout life cycle

where, $k_i = c_{max} - 75$ is the inflection point of gear degradation, x_i is the current operating moment, and c_{max} is the number of days of operation when the gear fails.

3.2 Evaluation index of model performance

In order to evaluate the prediction effect of the proposed method and quantify the quality of the model, this paper adopts the evaluation indexes including root mean square error (RMSE), mean absolute percentage error (MAPE) and SCORE, whose calculation formulas are shown in Eqs(28) – (31).

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(f(x_t) - y_t)^2}/\bar{y} \times 100\% \tag{28}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n} | f(x_t) - y_t | /\bar{y} \times 100\% \tag{29}$$

$$S_t = \begin{cases} e^{-\ln(0.5)\times\left(\frac{(y_t - f(x_t))}{30}\right)} & y_t - f(x_t) \leqslant 0 \\ e^{+\ln(0.5)\times\left(\frac{(y_t - f(x_t))}{50}\right)} & y_t - f(x_t) > 0 \end{cases} \tag{30}$$

$$SCORE = \frac{1}{n}\sum_{t=1}^{n}(100 \times S_t) \tag{31}$$

where, $f(x_t)$ is the predicted value at moment t, y_t is the measure of the moment t, $\bar{y}$ is the mean of measured values, and n is the number of samples. Eq. (30) defines the penalty rule for early prediction and late prediction[19]. If the predicted RUL is smaller than the actual RUL, the penalty is small, because there is still time to replace the faulty parts before the fault occurs. If the predicted RUL is greater than the actual RUL, measures will be taken after the fault occurs, and a major system failure may occur. Therefore, the penalty in this case is large.

3.3 Performance evaluation of predictive models

In order to verify the feasibility of the proposed method, the dynamic transmission error of the whole life cycle is taken as the vibration signal, and the cumulative wear of the tooth surface of the whole life cycle is taken as the oil signal, and they are input into the proposed DRSN-TPA-BiLSTM fusion prediction model. The BatchNormalization layer is first used to normalize vibration signals, and then the superficial features are obtained through the convolutional layer with 16 neurons, 3 filters and 2 steps. Then, RSBU-CW is used to obtain the deep vibration features. The number of convolutional layer neurons in RSBU-CW module is 32, the number of filters is 3, and the step size is 2, and then the maximum pooling process is carried out. Finally, the deep vibration features and oil data are fused and input into TPA-BiLSTM network for prediction. After several experiments, the number of BiLSTM neurons of the proposed DRSN-TPA-BiLSTM model is finally set to 50. In order to prevent overfitting during training, the prediction results are input into the Dropout layer, and the optimal drop rate is 0.2.

In the process of model training, GridSearchCV is used to optimize the hyperparameters. The optimized hyperparameters include batch _ size {16, 20, 30}, epochs {10, 20, 30} and optimizer {Adam, Adadelta}. The optimization result is batch _ size = 30, epochs = 20, and optimizer = Adam. To verify the necessity of oil signal fusion, the fusion prediction results are compared with those of the model with only vibration signal input, as shown in Fig. 9. In order to verify the superiority of DRSN-TPA-BiLSTM model, it is compared with LSTM, BiLSTM, DRSN-BiLSTM and other methods, and the comparison results are shown in Fig. 10.

It can be seen from Fig. 9 that the prediction effect of residual life of gear by combining vibration and oil information is significantly better than that by using only vibration information, indicating that the prediction accuracy of residual life of gear can be greatly improved by adding oil information on the basis of vibration information. As can be seen from Fig. 10, the prediction results of DRSN-TPA-BiLSTM fusion prediction model proposed in this paper are closest to the real RUL curve, and the results of DRSN-BiLSTM model are relatively close to the real RUL curve, while the prediction effects of traditional LSTM and BiLSTM models are relatively poor. Compared with LSTM, BiLSTM and DRSN-BiLSTM models, the prediction effect of the proposed method in gear wear simulation data is the best.

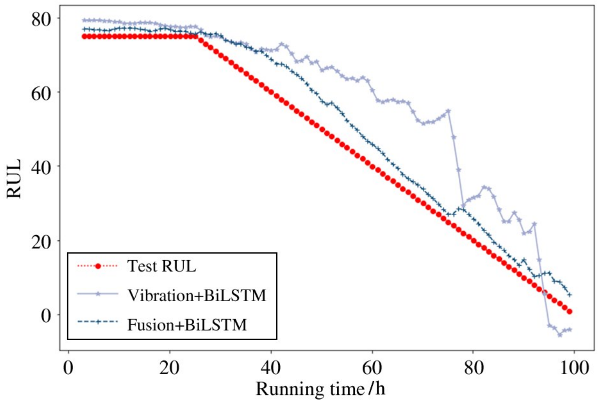

Fig. 9 Comparison of prediction effect before and after fusion

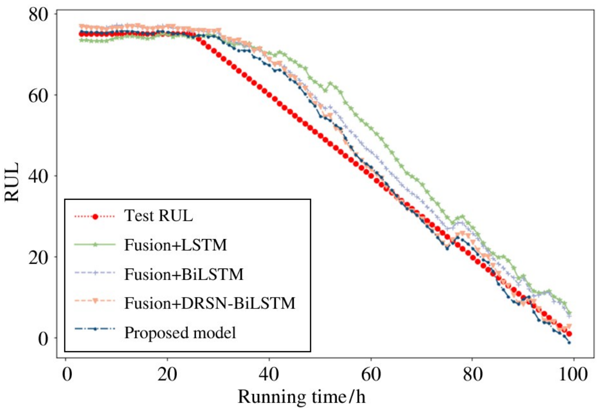

Fig. 10 Comparison of prediction effects of different models

According to the model performance evaluation indexes in subsection 3.3, the RMSE, MAPE error and SCORE of different prediction methods are calculated respectively. The RMSE and MAPE errors of vibration and oil information fusion are reduced by 9.171% and 7.939%, SCORE is improved by 8.705, respectively, compared with non-fusion methods. The prediction error of DRSN-TPA-BiLSTM model based on information fusion in this paper is reduced by 16.629% and 14.829%, SCORE is improved by 16.576, respectively, compared with non-fusion methods. The error comparison of each method is shown in Fig. 11. Obviously, this method has obvious superiority.

4 Experimental analysis of gear wear

4.1 Experimental equipment and data acquisition

After the above verification, the proposed method is further applied to the prediction of gear wear residual life on the test bench. The gearbox fault simulation test bench is shown in Fig. 12, and its main parameters are shown in Table 2. The vibration signal is obtained by installing acceleration sensors near each meshing gear pair of the gearbox, and the accumulation of ferromagnetic particles in lubricating oil is obtained by installing metal particle sensors at the lubricating oil circulation bypass of the gearbox, so as to realize real-time online monitoring and data acquisition of vibration and oil. The sampling frequency of the data acquisition system is 25 600 Hz, and the schematic diagram of the fault diagnosis system of multi-stage transmission gearbox is shown in Fig. 13. In order to accelerate the wear process, the first large spur gear at the input end of the gearbox is selected as the wear fault simulation gear. The tooth surface of this gear is not treated, and the other tooth surfaces are carburized.

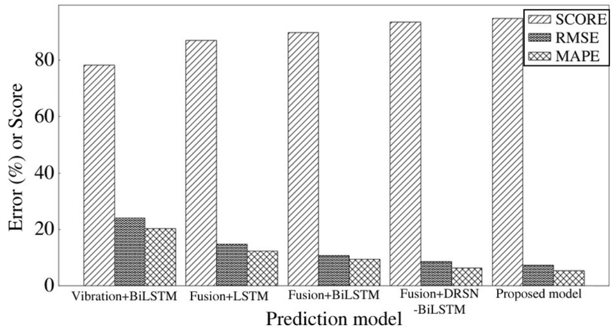

Fig. 11 Comparison of prediction errors of different methods

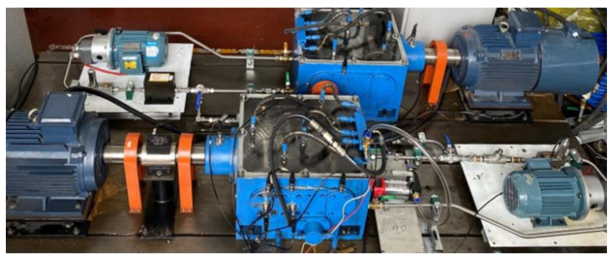

Fig. 12 Gearbox fault simulation test bench

Table 2 Main parameters of the test bench

Driving power /kW	Speed range/ (r/min)	Torque range/ (N · m)	Total transmission ratio	Driving stages
22	0 – 3000	0 – 150	1 : 1	4

In order to verify the effectiveness of the proposed method, a long-period gear wear experiment is carried out, which is run at 1200 r/min speed and 90 N · m load. After every 10 h of operation, the gear surface wear is checked out of the box until serious wear is visible to the naked eyes, and the experiment is stopped with a cumulative running time of 60 h. Tooth surface wear at different stages are shown in Fig. 14.

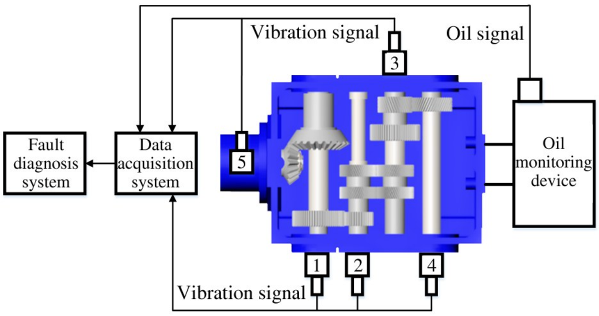

Fig. 13 Multi-stage transmission gearbox fault diagnosis system

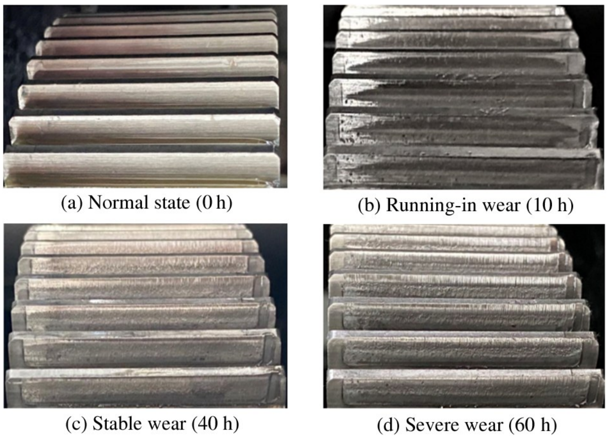

Fig. 14 Tooth surface wear at different stages

Fig. 15 and Fig. 16 show the vibration signal and oil signal of the whole life cycle collected by gear wear fault simulation experiment.

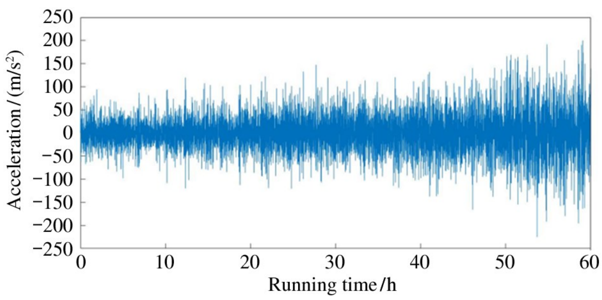

Fig. 15 Full life cycle vibration signal

Considering that the gear degradation is not obvious at the early stage of operation, piecewise linear function is adopted as the residual life degradation curve. It can be seen from Fig. 14 that significant degradation begins to occur after 10 h of operation, so the RUL curve can be expressed as

$$y_{\mathrm{RUL}}^{i}=\begin{cases}50 & x_i<k_i\\ 50-(x_i-k_i) & x_i\geqslant k_i\end{cases}\tag{26}$$

where, $k_i=c_{\max}-50$ is the inflection point of gear degradation, x_i is the current running moment, and $c_{\max}$ is the running time when the gear fails.

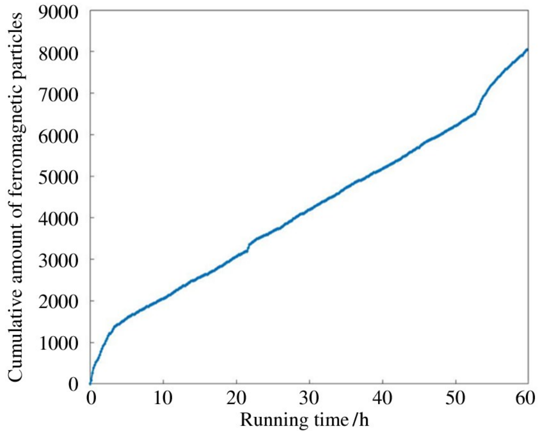

Fig. 16 Full life cycle oil signal

4.2 Gear wear life prediction based on DRSN-TPA-BiLSTM fusion of vibration and oil features

The vibration and oil signals obtained in the experiment are input into the vibration and oil information fusion prediction model based on DRSN-TPA-BiLSTM to predict the remaining service life of the gear. The parameter settings of the model are consistent with subsection 3.3.

In the process of model training, GridSearchCV is also used for hyperparameter optimization, the optimization result is batch _ size = 30, epochs = 30, and optimizer = Adam. In order to further verify the necessity of fusing the oil signal and the superiority of the proposed method, the fusion prediction results are compared with the non-fusion prediction results. At the same time, the proposed method is also compared with the existing prediction methods. The comparison results are shown in Fig. 17 and Fig. 18.

It can be seen from Fig. 17 that the prediction effect of gear remaining life by combining vibration and oil information is significantly better than that by using only vibration information, which is consistent with the

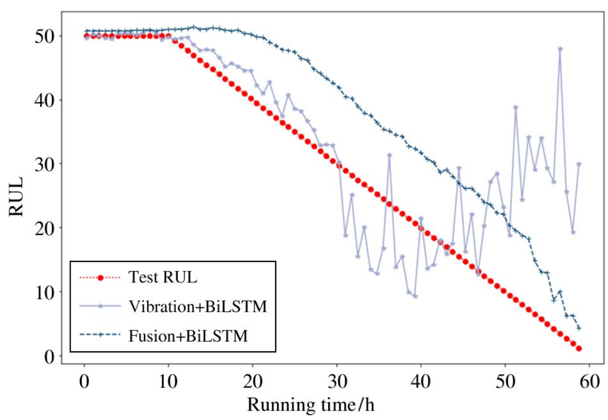

Fig. 17 Comparison of prediction effect before and after fusion

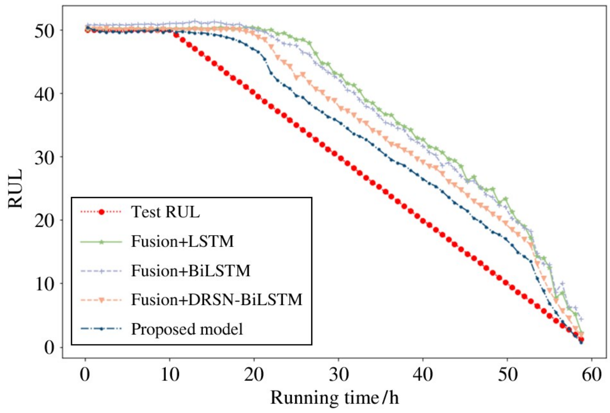

Fig. 18 Comparison of prediction effects of different models

verification result of simulation data. It can be seen from Fig. 18 that the prediction results of the DRSN-TPA-BiLSTM fusion prediction model proposed in this paper are closest to the real RUL curve. Compared with other models, the prediction effect of the proposed method is the best in the full-cycle experimental data of gear wear.

Similar to the simulation data, RMSE, MAPE error and SCORE of different prediction methods were calculated, and the RMSE and MAPE errors of vibration and oil information fusion are reduced by 10.272% and 2.790%, SCORE is improved by 0.196, respectively, compared with non-fusion methods. The prediction error of DRSN-TPA-BiLSTM model based on information fusion in this paper is reduced by 20.376% and 13.121%, SCORE is improved by 7.994, respectively, compared with non-fusion methods, as shown in Fig. 19.

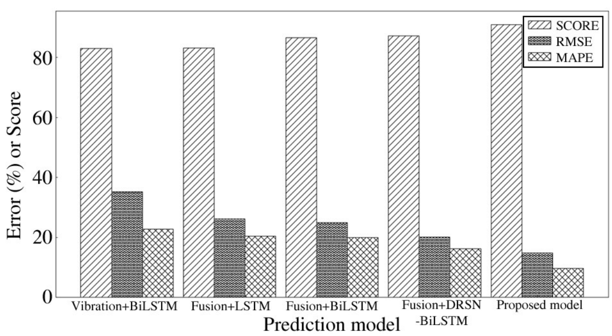

Fig. 19 Comparison of prediction errors of different methods

To further verify the superiority of the proposed method, the training duration of the proposed method is compared with that of the traditional method, and the comparison results are shown in Table 3.

As can be seen from Table 3, compared with the traditional LSTM prediction model, the training time of the prediction model based on the combination of DRSN and TPA-LSTM is greatly shortened. In conclusion, the vibration and oil information fusion model based on DRSN-TPA-BiLSTM proposed in this paper has the highest prediction accuracy and relatively short training time.

Table 3 Training duration comparison

Prediction model	LSTM	BiLSTM	DRSN-TPA-LSTM	DRSN-TPA-BiLSTM
Duration/s	1527	2335	460	706

5 Conclusions

A gear wear prediction method based on DRSN-TPA-BiLSTM is proposed, which integrates vibration and oil information, aiming to solve the problems such as difficulty in extracting vibration features of gear wear process, low prediction accuracy of residual life of gear only by collecting vibration or oil signals, and the deficient structure of traditional LSTM network. The main research conclusions are as follows.

(1) The prediction effect of residual life by combining vibration information with oil information is obviously better than that by using only vibration information.

(2) By solving the numerical calculation model of gear considering dynamic wear, it can be seen that the wear depth on the tooth profile increases with the increase of meshing cycles. With the deepening of wear, the dynamic transmission error increases. At the same time, the simulation data of gear life cycle can be obtained, which is of significance to the research of gear remaining life prediction.

(3) Adding RSBU-CW before the traditional BiLSTM model can improve the prediction accuracy of multivariate fusion feature data. Compared with LSTM, BiLSTM and DRSN-BiLSTM models, RMSE and MAPE of DRSN-TPA-BiLSTM model have the smallest error, the highest SCORE and its training time is shorter. The comprehensive performance of DRSN-TPA-BiLSTM model is the best.

References

[1] ZHANG S S. Research on gearbox wear fault diagnosis based on oil monitoring and grey theory [D]. Jinan: Shandong University, 2019: 1-10. (In Chinese)

[2] MORALES G E, RYCERZ P, KADIRIC A. Prediction of micropitting damage in gear teeth contacts considering the concurrent effects of surface fatigue and mild wear [J]. Wear, 2018, 398-399: 99-115.

[3] JAMES M, LILIANE P, PETER N, et al. A review on lubricant condition monitoring information analysis for maintenance decision support [J]. Mechanical Systems and Signal Processing, 2018, 118(MAR.1):108-132.

[4] SUN J Y, WANG L M, Li J F, et al. Online oil debris

monitoring of rotating machinery: a detailed review of more than three decades[J]. Mechanical Systems and Signal Processing, 2021, 149: 107341.
[5] SEJKOROVA M, KUCERA M, HURTOVA I, et al. Application of FTIR-ATR spectrometry in conjunction with multivariate regression methods for viscosity prediction of worn-out motor oils[J]. Applied Sciences, 2021, 11(9): 3842.
[6] LI H B, HUANG J M, RONG L. Application of grey prediction theory in ship machinery fault diagnosis [J]. Journal of Shanghai Maritime University, 2017, 38(3): 85-89. (In Chinese)
[7] DAVID V, JAKUB G, LIBOR Z. Potential for using the ANN-FIS meta-model approach to assess levels of particulate contamination in oil used in mechanical systems[J]. Tribology International, 2019, 135: 324-334.
[8] BAZI R, BENKEDJOUH T, HA H, et al. A hybrid CNN-BiLSTM approach-based variational mode decomposition for tool wear monitoring [J]. The International Journal of Advanced Manufacturing Technology, 2022, 119(1): 1-15.
[9] LU Y, ZHANG W, WANG W, et al. Application of oil and vibration monitoring in early fault warning of machine pump group bearings [J]. Lubrication Engineering, 2020,45(10): 136-140.
[10] ZHAO M H, ZHONG S S, FU X Y, et al. Deep residual shrinkage networks for fault diagnosis[J]. IEEE Transactions on Industrial Informatics, 2020, 16(7): 4681-4690.
[11] SUN-YAO S, FAN-KENG S, HUNG-YI L. Temporal pattern attention for multivariate time series forecasting[J]. Machine Learning, 2019, 108(8-9): 1421-1441.
[12] LIU F B. Simulation of wear process in spur gear[J]. Mechanical Science and Technology, 2004, 23(1): 55-56,59. (In Chinese)
[13] PRIIT P, SOREN A. Wear simulation with the Winkler surface model[J]. Wear, 1997, 207(1): 79-85.
[14] WANG H, ZHOU C, LEI Y, et al. An adhesive wear model for helical gears in line-contact mixed elastohydrodynamic lubrication[J]. Wear, 2019, 426-427: 896-909.
[15] WANG X S, WU S J, CHEN J, et al. Dynamic wear characteristics of spur gear transmission system considering dynamic load and dynamic wear coefficient [J]. Journal of Central South University (Science and Technology), 2014, 45(2): 408-413. (In Chinese)
[16] TAMMINANA V K, KAHRAMAN A, VIJAYAKAR S. A study of the relationship between the dynamic factors and the dynamic transmission error of spur gear pairs[J]. Journal of Mechanical Design, 2007,129(1): 75-84.
[17] CHEN W, LEI Y L, Li X Z, et al. Calculation of tooth surface adhesion wear of involute spur gear at low speed [J]. Journal of Jilin University (Engineering and Technology Edition), 2021,51(5): 1628-1634. (In Chinese)
[18] LIU H, LIU Z, JIA W, et al. A novel deep learning-based encoder-decoder model for remaining useful life prediction[C] // International Joint Conference on Neural Networks (IJCNN). Budapest: IEEE, 2019: 19657.
[19] SHUAI Z, RISTOVSKI K, FARAHAT A, et al. Long short-term memory network for remaining useful life estimation[C] // IEEE International Conference on Prognostics and Health Management (ICPHM). Dallas: IEEE, 2017: 88-95.

JIANG Zhinong, born in 1967. Now he is the director of Key Laboratory of Engine Health Monitoring and Networking, Ministry of Education. He received his M. S. degree from Dalian University of Technology and Ph. D degree from Beijing University of Chemical Technology. His research interests include the research and application of fault monitoring and diagnosis method for mechanical power equipment.