Measurement 167 (2021) 108277

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

A BiGRU method for remaining useful life prediction of machinery

Daoming She [a,b], Minping Jia [a,*]

[a] School of Mechanical Engineering, Southeast University, Nanjing 211189, PR China
[b] Center for Advanced Life Cycle Engineering (CALCE), University of Maryland, College Park 20742, USA

ARTICLE INFO

Keywords:
Prediction uncertainty
RUL
Bootstrap
Deep learning

ABSTRACT

Remaining useful life (RUL) prediction, allowing for mechanical prediction maintenance, reduces the unplanned expensive maintenance greatly. Deep learning methods have provided better point estimation for RUL prediction due to their powerful feature extraction capability. Because of the measurement noise and model parameters, the prediction results usually vary greatly. In order to express the uncertainty of prediction, it is necessary to calculate not only the determined RUL prediction value, but also the confidence interval (CI) of RUL. In this paper, a bidirectional gated recurrent unit (BiGRU) RUL prediction method based on bootstrap method is proposed. The confidence interval (CI) of RUL can be obtained through bootstrap method. The validity of the proposed method is demonstrated by ABLT-1A bearing data. Obtaining the uncertainty in the RUL prediction has great significance for the actual production and manufacturing.

1. Introduction

In 2013, Germany put forward the concept of "industry 4.0″ on Hannover Messe, and presented that the future manufacturing industry would be highly digital, networked, and self-organized based on the (Cyber-Physicalsystem, CPS) [1]. Under this background, it is of great value to conduct the RUL prediction and health management under industrial big data.

Fault prediction and health management include the following four stages: fault detection, fault diagnosis, RUL prediction and health management [2]. When a fault is found, the machine is usually stopped immediately to avoid serious economic losses, and the failure usually occurs at an inconvenient time [3,4]. Therefore, accurate RUL prediction for effective mechanical predictive maintenance is considered to reduce costly unplanned mechanical maintenance.

RUL prediction methods mainly include physical degradation model based approaches and data-driven RUL prediction approaches. The model-based approaches attempt to establish a mathematical or physical model describing the process of mechanical degradation, and update the model parameters using collected data [5]. Common models include Markov process model [6,7], winner process model [8.9], gauss mixture model [10]. Wang [11] put forward a mechanical degradation state prediction method based on particle filter. But it is so difficult to use a concrete model to describe complex mechanical systems, which affects the prediction effectiveness of model-based methods. In this case, without understanding the physical model and failure mechanism, data-driven methods emerge as the times require.

Data-driven methods do not consider the internal failure mechanism of machinery, but directly excavate the implicit machinery health information and its evolution law from the condition monitoring data (acceleration, temperature, oil residue quality, etc.), so as to realize the RUL prediction [12,13]. Qian et al. [14] realized bearing damage tracking and RUL prediction using multi-time scale method. Ahmad [15] described a reliable rolling bearing health prediction technology, which predicted the health status of bearing by new health indicator (HI), and estimated the RUL by dynamic regression model. In literature [16], a new measurement method of signal-to-noise ratio (SNR) was proposed. Based on the new HIs, a multi-data fusion model RUL prediction method was given.

Deep learning approach is a popular direction in machine learning. Deep learning networks have a strong ability to extract potential fault features and degraded information automatically. Deep learning methods such as deep belief network (DBN) [17–19], convolutional neural networks (CNN) [20], long short-term memory (LSTM) [21] have emerged in the RUL prediction. Zhao et al. [22] extracted the degradation characteristics from a large number of vibration data of centrifugal pump bearing, and proposed a deep feature optimization fusion method [23]. Wu et al. [24] introduced a vanilla LSTM network to maximize the use of long-term memory ability under complex operation and working conditions to achieve higher prediction accuracy. Zhao

* Corresponding author.
E-mail address: mpjia@seu.edu.cn (M. Jia).

https://doi.org/10.1016/j.measurement.2020.108277
Received 10 August 2019; Received in revised form 3 June 2020; Accepted 19 July 2020
Available online 25 July 2020

et al. [25] presented local feature-based GRU networks that used for condition monitoring. Zheng [26] shown a LSTM method for RUL prediction. Elsheikh [27] introduced a new LSTM architecture to predict RUL. Chen et al. [28] put forward a recurrent neural network based on encoder-decoder framework with attention to predict HI values, which are designed closely related with the RUL values. Li et al. [29] considered the size optimization accuracy of vibration characteristics and the iteration speed of industrial model.

The deep learning methods provide effective point estimation for RUL prediction. However, when the collected data contains a lot of noise and the model parameters are complex, only point prediction is of little practical significance. In multi-fault mode, multi-working condition and multi-noise environment, the uncertainty of measurement and prediction model should be taken into account, so that the prediction results are much closer to reality. The main uncertainty calculation methods were Bayesian [30,31], bootstrap [32], mean–variance estimation (MVE) [33], lower upper bound estimation (LUBE) [34] et al. Zhang et al. [35] described an improved LSTM to predict RUL of lithium batteries, and calculated the CI of the RUL with Monte Carlo simulation method. Yarin et al. [36] calculated the uncertainty of the deep learning model using dropout method. Nick et al. [37] discussed the uncertainty of weights in neural networks. Balaji et al. [38] introduced a deep ensemble method to obtain prediction uncertainty. This method was easy to implement and parallelize, and required fewer hyper parameters. Bootstrap method is a statistical inference method of simulated sampling based on raw data. It can be used to study the distribution features of a certain statistic of a group of data. Bootstrap method is simple to operate and does not depend on the distribution of observed data. It has provided reliable estimation of uncertainty of neural network [39]. Moreover, there is little research on uncertain calculation in recurrent neural networks.

To sum up, aiming at the problem that the uncertainty of life prediction based on deep learning method is difficult to quantify, a BiGRU method for predicting the RUL of machinery based on bootstrap method is proposed. The CI of RUL can be obtained through bootstrap method. The validity of the presented method is demonstrated by ABLT-1A bearing experimental data.

The outline of the paper is organized in the following aspects. Section 2 will present brief introduction of GRU model, along with the proposed BiGRU model and then the detailed procedure of the method. Section 3 will provide the validation of the presented method by the experimental data from ABLT-1A platform. The conclusions and future work are provided.

2. The proposed method

2.1. The BiGRU model

GRU model is a simplified version of LSTM model. GRU combines input gate and forgetting gate in LSTM into update gate by using the difference between gate structure and LSTM. It contains only two gate structures, reset gate and update gate. The influence of the output hidden layer at the previous time on the current hidden layer is controlled by the update gate. The larger the update gate is, the greater the influence of the hidden layer output at the previous moment on the current hidden layer is. The degree of neglect of the hidden layer information at the previous moment is decided by the reset gate, and the smaller the reset gate value is, the more information is ignored. The reset gate mainly determines how the previous information is combined with the current input information, and the updated door mainly determines how much previous information is retained [40,41]. The gate structure of GRU is depicted in Fig. 1.

In Fig. 1, x denotes the input data, and h represents the output of GRU unit. r is the reset gate, and z is the update gate. r and z decide how to get the new hidden state h_t from the previous hidden state h_{t-1} calculation. The update gate controls both the current input x_t and the previous memory h_{t-1}, and outputs a numerical value z_t between 0 and 1. The calculation formula is as follows,

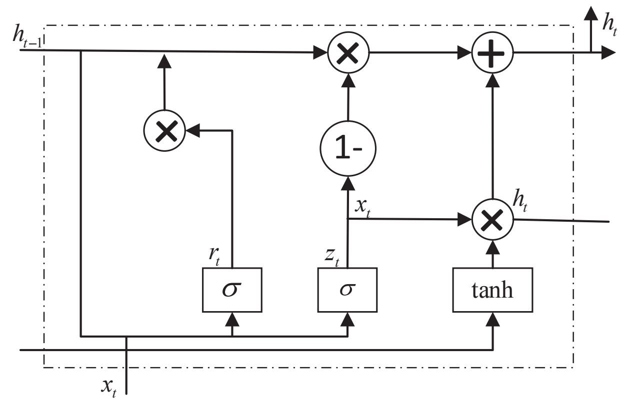

Fig. 1. Basic GRU structure.

$$z_t = \sigma(W_z[h_{t-1}, x_t] + b_x) \tag{1}$$

where z_t determines to what extent h_{t-1} should be passed to the next state, which is available from formula (1). In formula (1), σ is sigmoid activation function, W_z is the update gate weight and b_z is the bias. Reset gate controls the importance of h_{t-1} to result h_t. If the previous memory h_{t-1} is totally unrelated to new memory, the reset gate can play a role in removing the influence of the previous memory,

$$r_t = \sigma(W_r[h_{t-1}, x_t] + b_r) \tag{2}$$

Generating new memory information $\tilde{h}_t$ based on update gate,

$$\tilde{h}_t \ = \ \tanh(W_h[r_t h_{t-1}, x_t] + b_h) \tag{3}$$

The output at the current moment is h_t,

$$h_t = (1 - z_t)h_{t-1} + z_t\tilde{h}_t \tag{4}$$

The base unit of the BiGRU model consists of a forward-propagated GRU unit and a backward-propagated GRU unit. In a one-way neural network structure, the state always propagates from the front to the back. However, in the RUL prediction, if the output at this moment can be related to the state of the previous and the state of the latter moment closely. The current hidden layer state of BiGRU is determined by the current input x_t, the forward hidden state $\overrightarrow{h}_{t-1}$ at *t-1* and the output $\overleftarrow{h}_{t-1}$ of the reverse hidden layer state,

$$\overrightarrow{h}_t \ = \ G(x_t, \overrightarrow{h}_{t-1}) \tag{5}$$

$$\overleftarrow{h}_t \ = \ G(x_t, \overleftarrow{h}_{t-1}) \tag{6}$$

$$h_t = w_t\overrightarrow{h}_t + v_t\overleftarrow{h}_t + b_t \tag{7}$$

The $G(\cdot)$function represents the nonlinear transformation of the input, and encodes the degradation indicator into the corresponding GRU hidden state. w_t and v_t represent the weights of forward hidden layer state $\overrightarrow{h}_{t-1}$ and reverse hidden state output $\overleftarrow{h}_{t-1}$ corresponding to bidirectional GRU at time t, respectively, and b_t denotes the bias corresponding to hidden state at time t.

This paper constructs BiGRU layer and fully connected regression layers to predict the degradation state of machinery based on Tensorflow. The deep neural network (DNN) consists of five layers: one input layer, four hidden layers and one output layer. D_m is the output of recurrent neural network. The hidden layer and output layer can be calculated by the following formula:

$$Y_{m+1} = f(W_m D_m + b_m) \tag{8}$$

where W_m is the weight connecting layer m and layer $m + 1$, and b_m is the bias. $f(\cdot)$ is the activation function. The activation function is not utilized in the last layer. Fig. 2 shows the BiGRU model proposed in this paper. BiGRU model can make full use of the relevant knowledge of past and future degradation state of mechanical equipment and transfer between layers effectively, thereby improving the accuracy of RUL prediction.

2.2. Bootstrap method

Because the sensor data is full of noise, it is difficult to directly use the sensor data to reflect the degradation state of the machinery. Firstly, HI describing the degradation process of mechanical equipment is constructed. To include degradation information of mechanical equipment, sliding window technology [42] is adopted in this paper. Assume that the existing HI describing the whole life degradation process of mechanical equipment has N values, namely $(z(t_1), z(t_2), \ldots, z(t_N))$, $n = 1, 2, \ldots N$. Let L represent the length of the sliding window. $(z(t_i), z(t_{i+1}), \ldots, z(t_{i+L-1}))$ is a degenerate state of window length and its corresponding output is $z(t_{i+L})$. N degradation state data can generate N-L samples, and the predicted system state can be expressed by the following function,

$$z(t_{i+L}) = \phi(z(t_i), z(t_{i+1}), \ldots, z(t_{i+L-1})) \tag{9}$$

The CI of results is to quantify the uncertainty of point prediction. Bootstrap method is a statistical inference method of simulated sampling based on raw data. It can be used to study the distribution features of a certain statistic of a group of collected data, especially for those problems such as interval estimation of parameters and hypothesis test which are difficult to derive by conventional methods. Bootstrap method theoretically reduces the deviation between the predicted value and the regression mean, and provides a new scheme for quantifying the prediction uncertainty.

As shown in the flowchart of Fig. 3, the main ideology of this method is to resample K times from the original training data by an alternative method. Bootstrap model $\phi_k (k = 1, 2, \ldots, K)$ is trained with resampled data $S_k(t_1 : t_i)(k = 1, 2, \ldots, K)$ every time. Finally, the integration of multiple models will produce the mean and variance of RUL prediction, which is formulated as follows,

$$Z_k(l_i + t_i) = \phi_k(S_k(t_1 : t_i)) \tag{10}$$

where, $Z_k(l_i + t_i)$ represents the state prediction value of the system obtained by bootstrap method.

The RUL l_i of time t_i is defined as follows

$$l_i = inf\{l_i : Z_k(l_i + t_i) \geqslant \tau | z(t_1 : t_i)\} \tag{11}$$

where, τ denotes a pre-set failure threshold. $z_{0:i}$ is the estimated system state value from t_0 to t_i, and $Z_k(t_i + l_i)$ represents the estimated system state value at $t_i + l_i$. The CI of the RUL can be obtained by calculating the percentage of the RUL l_i at t_i.

The concrete steps of the approach proposed in this paper are described below:

1. HI quantifies the degradation process of mechanical equipment by extracting effective feature information from measured data. Firstly, HI is constructed to quantify the degradation process of mechanical equipment. The training and testing data sets are constructed by sliding window technology.

2. BiGRU neural network is constructed, and integrated learning network is constructed by bootstrap method to obtain the uncertainty expression of prediction results.

The training data and test data is acquired in step 1. The BiGRU is optimized through error back-propagation. The stochastic gradient descent optimizer is used to train the network, so as to reduce the training error of the network, and finally reduce the prediction error and uncertainty.

3. Obtaining the mean and CI of RUL: The test set is input to the trained network, and the output corresponds to the expected degradation state. At last, the mean and CI of RUL can be obtained.

3. Case study

The whole life vibration signal of rolling bearing came from the accelerating bearing life test rig ABLT-1A in School of Mechanical Engineering, Southeast University, China. The test bench consisted of test head, loading system, transmission system, lubrication system, electrical control system and data acquisition system, as depicted in Fig. 4 (a). Four bearings were installed simultaneously for accelerated bearing life test.

To monitor each bearing, the test bed was adjusted to 4 accelerometers to pick up vibration signals from three rigid body shells respectively. The data was acquired by National Instruments NI 9234. The data acquisition system was displayed in the Fig. 4 (b). The accelerating bearing life test conditions is depicted in Table 1.The installation diagram is shown in Fig. 4 (c). The rated dynamic load of the bearing 6308 was 42.3 kN, and the actual weight was 35 kg, i.e., the rated dynamic load on each bearing was 17.5 kN. The loading condition of radial load is shown in Table 2. After full loading, the bearing ran for 16 h, and the final test machine stopped because the root mean square (RMS) of vibration reached the pre-set threshold. The RMS of vibration with full load was 11.0, and the shutdown threshold was set to 45.0. After cutting the bearing by wire cut electrical discharge machining, it could be seen that the outer ring of the bearing at position 3 had obvious spalling

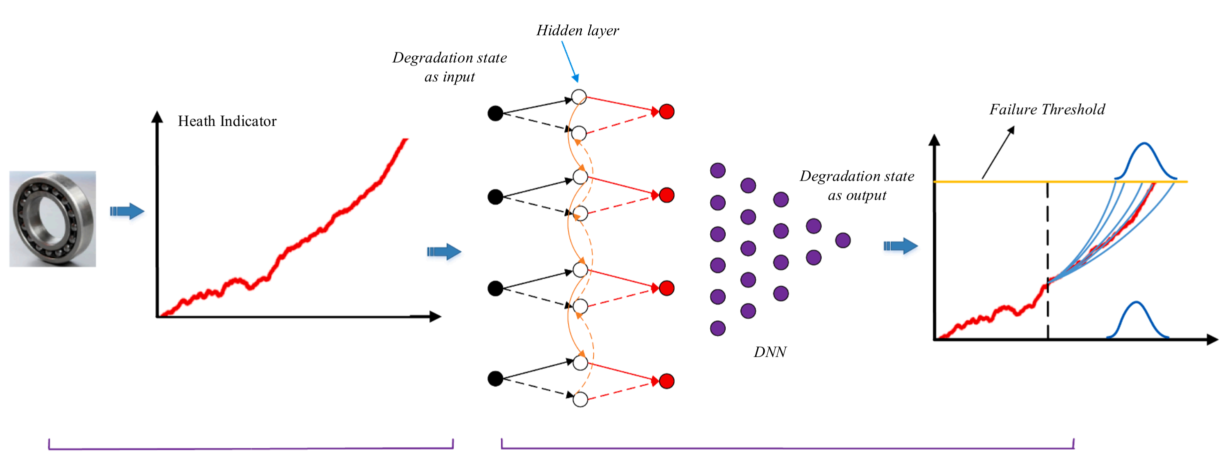

Fig. 2. The proposed BiGRU model.

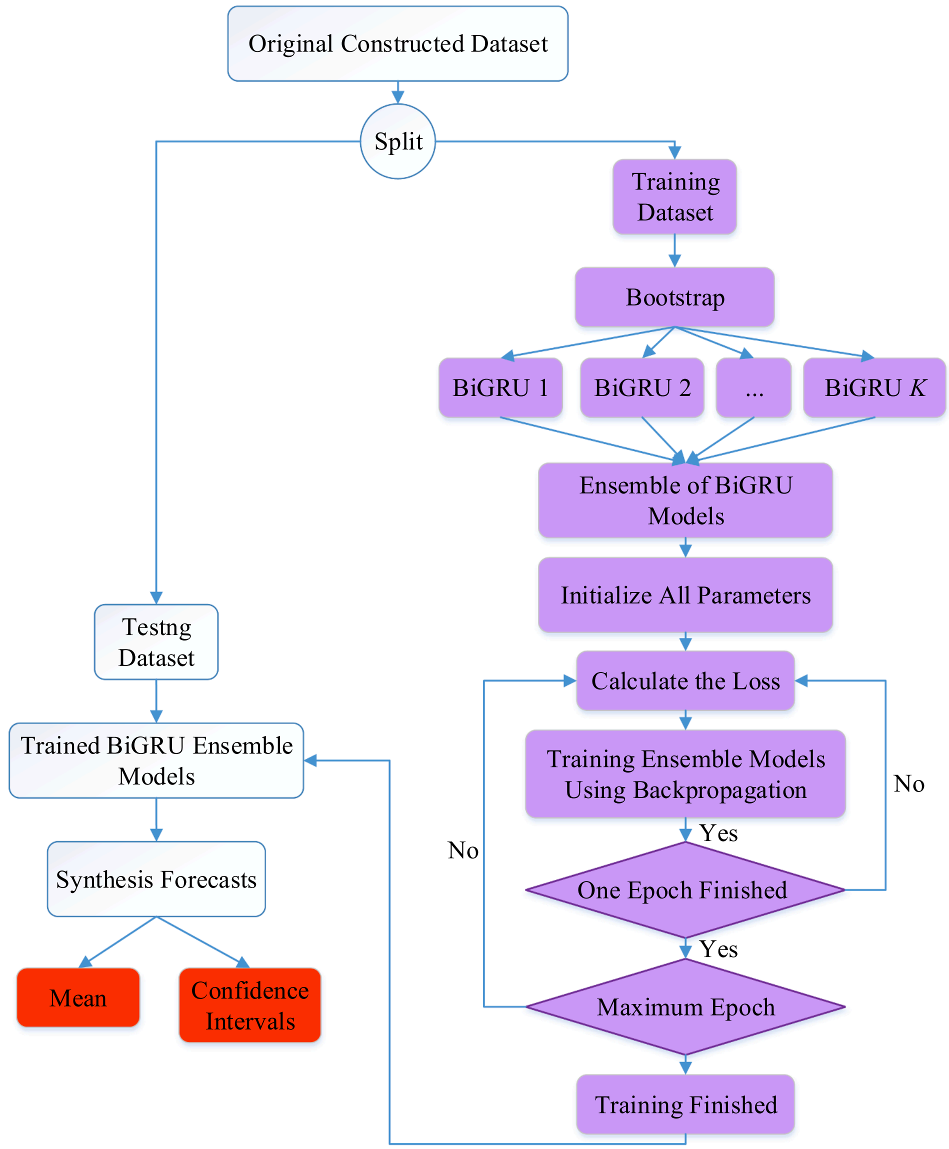

Fig. 3. Flowchart of bootstrap method proposed in this paper.

phenomenon, as shown in Fig. 4 (d).

3.1. HI construction

The configuration of the computer used in the experiment was as follows: Intel Core i7-8550U 1.80 GHz, 8 GB, Nvidia GeForce MX150, Windows 10 (64-bit), Python 3.5 (64-bit). The BiGRU model was implemented by tensorflow-gpu 1.7.0 package. Fig. 5 displays the time domain waveform of bearing. The original features extracted from the vibration signal refer to the literature [43] and [44]. Feature subsets are composed of features with large trendability. Finally, the HI describing the bearing degradation is constructed with the minimum quantization error method [45], as shown in Fig. 6.

3.2. RUL prediction

The neuron number of BiGRU network was set to 300 and the sliding window length L was 40. The BiGRU was trained by using the stochastic gradient descent optimizer. The total training step was 200 steps. The training and test data sets were constructed by using the constructed HIs and sliding window technology. The test results were the predicted degradation state of rolling bearings. The CI of the predicted results was obtained by Bootstrap method. And the mean and CI of RUL could be obtained.

To verify the effectiveness of the method in this paper, four compared RUL prediction methods were taken in this paper. The three methods for comparison were the same HI as the method in this paper. In method 1, model-based prediction method was compared. It was assumed that the degradation process was based on exponential model, and then particle filter method was conducted to estimate the degradation state. In method 2, it was a single-layer GRU network. In method 3, a single-layer LSTM network was used. In method 4, the three-layer fully connected neural network prediction was adopted. Mean squared error (MSE) and mean absolute percentage error (MAPE) were used to evaluate the RUL prediction results. The calculation formulas are as follows

$$RMSE = \sqrt{\frac{1}{N_p}\sum_{i=1}^{N_p}(HI_{Act} - HI_{Pre})^2} \tag{12}$$

$$MAPE = \frac{1}{N_p}\sum_{i=1}^{N_p}\left|\frac{HI_{Act} - HI_{Pre}}{HI_{Act}}\right| \tag{13}$$

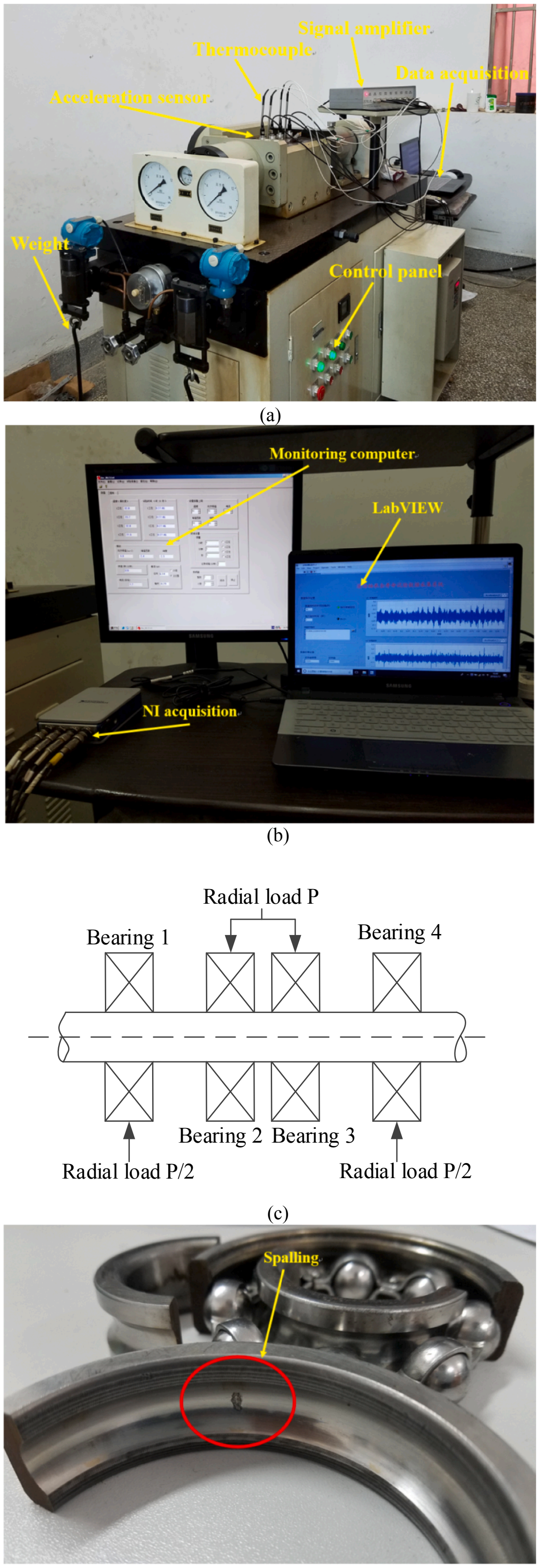

Fig. 4. The accelerating life test of bearing 6308. (a) ABLT-1A test platform. (b) Data acquisition. (c) The installation diagram of the 6308 bearing. (d) The failed bearing.

Table 1
Accelerating bearing life test conditions.

Parameter	Value
Bearing model	6308 single row deep groove ball bearing
Test speed/(r·min)	3000
Number of tested bearing	4
Sampling frequency /kHz	25.6
Save interval/s	30
Data length	25,600
Radial load on each bearing /kN	17.5

Table 2
Radial load.

Time	Weight (kg)	Radial load (kN)
2019-03-21 T 16:05--2019-03-21 T 19:41	0 → 35	0 → 17.5
2019-03-21 T 19:41-2019-03-22 T 11:19	35	17.5

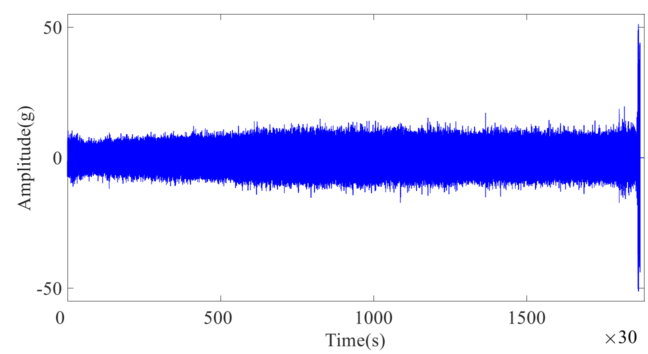

Fig.5. The vibration signal in time-domain of bearing 6308.

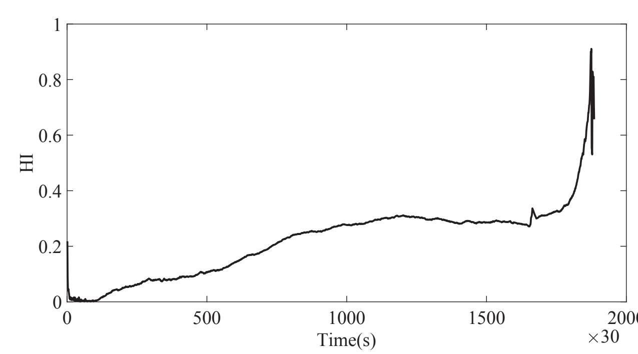

Fig. 6. HI of bearing 6308.

In the formula, HI_{Act} is the actual bearing degradation state. HI_{Pre} is the predicted bearing degradation state and N_p describes the number of prediction points. The results of five methods are shown in Fig. 7.

To evaluate the uncertainty of RUL prediction, median value and 95% CI of RUL were calculated. The prediction points of the five methods are all at 750 min. From Fig. 7 and Table 3, it can be seen that the median and 95% CI of the proposed method are 175.5 min and [167.0, 181.0] min, respectively, which covers the actual RUL value 168.0 min. In Fig. 7 (b), the part before the prediction point is the process of particle trajectory tracking. Fig. 7 (c), (d) and (e) show the prediction results of GRU, LSTM and NN methods, respectively. The CI of the GRU, LSTM and NN methods can not cover the true RUL. The MSE of the proposed method is 30.3%, 21.4%, 34.9%, 28.4% smaller than that of the other four methods, respectively. The MAPE of the presented method is 93.9%, 21.4%, 41.1%, 45.1% smaller than that of the other four methods, respectively. The prediction results of LSTM and NN

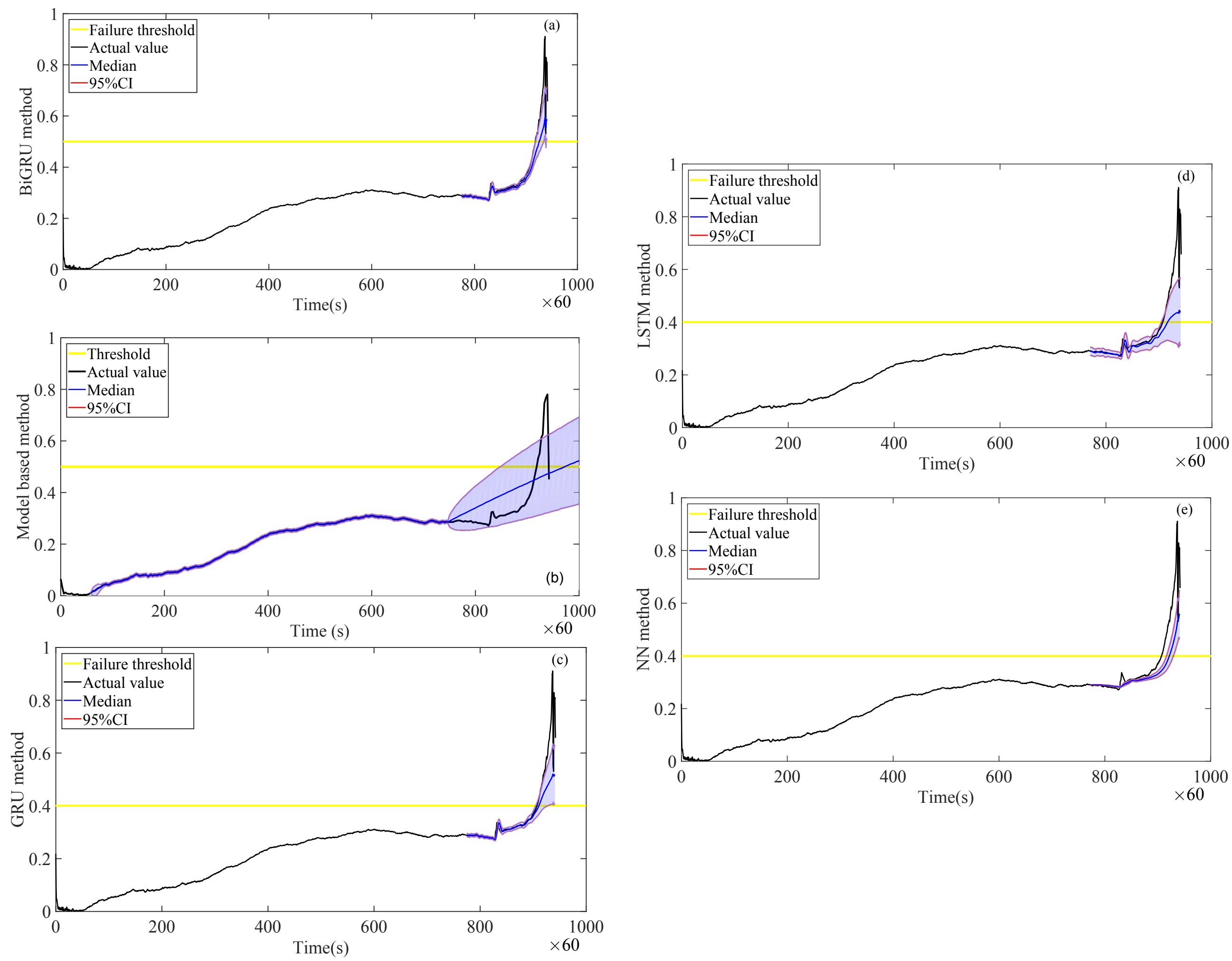

Fig. 7. RUL prediction by five methods.

Table 3
Comparison of five methods.

	RUL	Median	95%CI	Error	Error percentage	MSE	MAPE
BiGRU	168.0	175.5	[167.0, 181.0]	7.5	4.5%	0.0614	0.0609
Model based	168.0	222.0	[129.0, 357.0]	54.0	32.1%	0.0881	0.8817
GRU	157.0	165.3	[157.5, 171.0]	8.3	5.3%	0.0781	0.0724
LSTM	157.0	167.0	[160.0, 188.5]	10.0	6.4%	0.0943	0.1034
NN	157.0	171.0	[164.0, 175.3]	14.0	8.9%	0.0857	0.1110

methods can not reach the failure threshold of 0.5, so the failure threshold is set as 0.4. BiGRU model can make full use of the knowledge of past and future degradation state of mechanical equipment and transfer between layers effectively, thereby improving the accuracy of RUL prediction.

4. Conclusion

Aiming at the problem that the uncertainty of RUL prediction using deep learning methods is difficult to quantify, a BiGRU approach of RUL prediction based on bootstrap method is presented in this paper. The CI of RUL is obtained through bootstrap method. The value of our approach is demonstrated by the accelerating bearing life test data. The results indicate that the MSE of the proposed method is 30.3%, 21.4%, 34.9%, 28.4% smaller than that of the other four methods, respectively. The MAPE of the presented method is 93.9%, 21.4%, 41.1%, 45.1% smaller than that of the four compared methods, respectively. BiGRU model can make full use of the knowledge of past and future degradation state of mechanical equipment and transfer between layers effectively, thereby improving the accuracy of RUL prediction. Bootstrap method can obtain the CI of the RUL, and then express the uncertainty of the deep learning model effectively. Obtaining the CI representing the uncertainty in the deep learning model provide guidance value for the engineering and actual manufacturing.

For future work, the authors will further study the quantitative problem of the prediction uncertainty and the RUL prediction under multi factors and variable working conditions.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (Grant No. 51675098), and the Postgraduate Research & Practice Innovation Program of Jiangsu Province (Grant No. KYCX18_0066). The authors also would like to thank the China Scholarship Council.

References

[1] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, R.X. Gao, Deep learning and its applications to machine health monitoring, Mech. Syst. Signal Pr. 115 (2019) 213–237.
[2] D. She, M. Jia, Wear indicator construction of rolling bearings based on multi-channel deep convolutional neural network with exponentially decaying learning rate, Measurement 135 (2019) 368–375.
[3] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, J. Lin, Machinery health prognostics: A systematic review from data acquisition to RUL prediction, Mech. Syst. Signal Pr. 104 (2018) 799–834.
[4] M. Demetgul, K. Yildiz, S. Taskin, I.N. Tansel, O. Yazicioglu, Fault diagnosis on material handling system using feature selection and data mining techniques, Measurement 55 (2014) 15–24.
[5] G. Liang, N. Li, J. Feng, Y. Lei, L. Jing, A recurrent neural network based health indicator for remaining useful life prediction of bearings, Neurocomputing 240 (2017) 98–109.
[6] H. Dui, S. Si, M.J. Zuo, S. Sun, Semi-Markov Process-Based Integrated Importance Measure for Multi-State Systems, IEEE T Reliab. 64 (2015) 754–765.
[7] A. Veeramany, M.D. Pandey, Reliability analysis of nuclear piping system using semi-Markov process model, Ann. Nucl. Energy 38 (2011) 1133–1139.
[8] Z. Huang, Z. Xu, W. Wang, Y. Sun, Remaining Useful Life Prediction for a Nonlinear Heterogeneous Wiener Process Model With an Adaptive Drift, IEEE T Reliab. 64 (2015) 687–700.
[9] T. Li, H. Pei, Z. Pang, X. Si, J. Zheng, A Sequential Bayesian Updated Wiener Process Model for Remaining Useful Life Prediction, IEEE Access 8 (2020) 5471–5480.
[10] X.S. Si, W. Wangbde, D.H. Zhouc, Remaining useful life estimation – A review on the statistical data driven approaches, Eur. J. Oper. Res. 213 (2011) 1–14.
[11] J. Wang, R.X. Gao, Z. Yuan, Z. Fan, L. Zhang, A joint particle filter and expectation maximization approach to machine condition prognosis, J. Intell. Manuf. (2017) 1–17.
[12] D. Wang, K. Tsui, Q. Miao, Prognostics and Health Management: A Review of Vibration Based Bearing and Gear Health Indicators, IEEE Access. 6 (2018) 665–676.
[13] Z. Liu, Y. Cheng, P. Wang, Y. Yu, Y. Long, A method for remaining useful life prediction of crystal oscillators using the Bayesian approach and extreme learning machine under uncertainty, Neurocomputing 305 (2018) 27–38.
[14] Y. Qian, R. Yan, R.X. Gao, A multi-time scale approach to remaining useful life prediction in rolling bearing, Mech. Syst. Signal Process. 83 (2016) 549–567.
[15] W. Ahmad, S.A. Khan, M.M.M. Islam, J.M. Kim, A reliable technique for remaining useful life estimation of rolling element bearings using dynamic regression models ☆, Reliab. Eng. Syst. Safe., (2018) S0951832017310608.
[16] K. Liu, A. Chehade, C. Song, Optimize the Signal Quality of the Composite Health Index Via Data Fusion for Degradation Modeling and Prognostic Analysis, IEEE Trans. Autom. Sci. Eng. 14 (2017) 1504–1514.
[17] X. Zhang, L. Xiao, J. Kang, Degradation prediction model based on a neural network with dynamic windows, Sensors-Basel 15 (2015) 6996.
[18] J. Hu, L. Zhang, W. Tian, Z. Shuai, DBN based failure prognosis method considering the response of protective layers for the complex industrial systems, Eng. Fail. Anal. 79 (2017) 504–519.
[19] Z. Lin, W. Xue, A deep feature optimization fusion method for extracting bearing degradation features, IEEE Access, PP (2018) 19640–19653.
[20] G.S. Babu, P. Zhao, X.L. Li, Deep Convolutional Neural Network Based Regression Approach for Estimation of Remaining Useful Life, in: International Conference on Database Systems for Advanced Applications, Dallas, Texas, USA, 2016, pp. 214–228.
[21] M. Rao, Q. Li, D. Wei, M.J. Zuo, A deep bi-directional long short-term memory model for automatic rotating speed extraction from raw vibration signals, Measurement 158 (2020), 107719.
[22] L. Zhao, X. Wang, A deep feature optimization fusion method for extracting bearing degradation features, IEEE Access, PP (2018) 19640–19653.
[23] J. Zhang, P. Wang, R. Yan, R.X. Gao, Long short-term memory for machine remaining life prediction, J. Manuf. Syst. 48 (2018) 78–86.
[24] Y. Wu, Y. Mei, S. Dong, L. Li, Y. Liu, Remaining Useful Life Estimation of Engineered Systems using vanilla LSTM Neural Networks, Neurocomputing, (2017) S0925231217309505.
[25] R. Zhao, D. Wang, R. Yan, K. Mao, F. Shen, J. Wang, Machine Health Monitoring Using Local Feature-Based Gated Recurrent Unit Networks, IEEE T Ind. Electron. 65 (2018) 1539–1548.
[26] Z. Shuai, K. Ristovski, A. Farahat, C. Gupta, Long Short-Term Memory Network for Remaining Useful Life estimation, in: 2017 IEEE International Conference on Prognostics and Health Management (ICPHM), Ottawa, Canada, 2017, pp. 88–95.
[27] A. Elsheikh, S. Yacout, M. Ouali, Bidirectional handshaking LSTM for remaining useful life prediction, Neurocomputing 323 (2019) 148–156.
[28] Y. Chen, G. Peng, Z. Zhu, S. Li, A novel deep learning method based on attention mechanism for bearing remaining useful life prediction, Appl. Soft Comput. (2019), 105919.
[29] Y. Li, X. Meng, Z. Zhang, G. Song, A Remaining Useful Life Prediction Method Considering the Dimension Optimization and the Iterative Speed, IEEE Access 7 (2019) 180383–180394.
[30] P.L. Mcdermott, C.K. Wikle, Bayesian Recurrent Neural Network Models for Forecasting and Quantifying Uncertainty in Spatial-Temporal Data, (2018).
[31] A. Kendall, Y. Gal, R. Cipolla, Multi-Task Learning Using Uncertainty to Weigh Losses for Scene Geometry and Semantics, in: Proceedings of the IEEE conference on computer vision and pattern recognition, Salt Lake City, UT, USA, 2018, pp. 7482–7491.
[32] K. Li, R. Wang, H. Lei, T. Zhang, Y. Liu, X. Zheng, Interval prediction of solar power using an Improved Bootstrap method, Sol. Energy 159 (2018) 97–112.
[33] A. Khosravi, S. Nahavandi, D. Creighton, A.F. Atiya, Comprehensive review of neural network-based prediction intervals and new advances, IEEE Trans. Neural Netw. 22 (2011) 1341–1356.
[34] K. Abbas, N. Saeid, C. Doug, A.F. Atiya, Lower upper bound estimation method for construction of neural network-based prediction intervals, IEEE Trans. Neural Netw. 22 (2011) 337–346.
[35] Y. Zhang, X. Rui, H. He, M. Pecht, Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries, IEEE T Veh. Technol. 67 (7) (2018) 5695–5705.
[36] Y. Gal, Z. Ghahramani, Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning, in: International Conference on Machine Learning, New York City, NY, USA, 2016, pp. 1050–1059.
[37] N. Pawlowski, M. Rajchl, B. Glocker, Implicit Weight Uncertainty in Neural Networks, arXiv preprint arXiv:1711.01297, 2017.
[38] B. Lakshminarayanan, A. Pritzel, C. Blundell, Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles, in: Advances in Neural Information Processing Systems, Long Beach, CA, USA, 2017, pp. 6402–6413.
[39] Y. Liao, L. Zhang, C. Liu, Uncertainty Prediction of Remaining Useful Life Using Long Short-Term Memory Network Based on Bootstrap Method, in: 2018 IEEE International Conference on Prognostics and Health Management (ICPHM), Seattle, Washington, United States, 2018, pp. 1–8.
[40] M. Ravanelli, P. Brakel, M. Omologo, Y. Bengio, Light Gated Recurrent Units for Speech Recognition, IEEE Trans. Emerg. Top. Comput. Intell. 2 (2018) 92–102.
[41] Y. Yuan, C. Tian, X. Lu, Auxiliary Loss Multimodal GRU Model in Audio-visual Speech Recognition, IEEE Access, PP (2018) 5573–5583.
[42] J. Deutsch, D. He, Using Deep Learning-Based Approach to Predict Remaining Useful Life of Rotating Components, IEEE Trans. Syst., Man, Cybernet.: Syst. 48 (2018) 11–20.
[43] Y. Lei, N. Li, S. Gontarz, L. Jing, J. Dybala, A Model-Based Method for Remaining Useful Life Prediction of Machinery, IEEE T Reliab. 65 (2016) 1314–1326.
[44] K. Javed, R. Gouriveau, N. Zerhouni, P. Nectoux, A feature extraction procedure based on trigonometric functions and cumulative descriptors to enhance prognostics modeling, in: 2013 IEEE Conference on Prognostics and Health Management (PHM), Gaithersburg, Maryland, USA, 2013, pp. 1–7.
[45] S. Daoming, J. Minping, Z. Wan, Deep auto-encoder network method for health assessment of rolling bearings, J. Southeast Univ. (Nat. Sci. Ed.) (2018) 801–806.