COMPUTATIONAL RESEARCH PROGRESS IN APPLIED SCIENCE & ENGINEERING (CRPASE)

CRPASE: TRANSACTIONS OF MECHANICAL ENGINEERING

Journal homepage: http://www.crpase.com

CRPASE: Transactions of Mechanical Engineering 8 (4) Article ID: 2817, 1–9, December 2022 ISSN 2423-4591

Research Article

A Hybrid CNN-BiLSTM and Wiener Process-based Prediction Approach of Remaining Useful life for Rolling Bearings

Jia-Lun Wan [1], Yan Yang [2], Junyu Guo [3*], Le Dai [3]

[1] *School of Computer Science, Southwest Petroleum University, Chengdu, Sichuan, 610500, P.R. China*
[2] *School of Sciences, Southwest Petroleum University, Chengdu, Sichuan, 610500, P. R. China*
[3] *School of Mechatronic Engineering, Southwest Petroleum University, Chengdu, Sichuan, 610500, P.R. China*

Keywords

RUL,
CNN,
BiLSTM,
Wiener process,
Bayesian method.

Abstract

Predicting the remaining useful life (RUL) of rolling bearings can provide guidance and reference for effective maintenance of rolling bearings in advance to ensure the regular operation of the machine. Therefore, maintaining bearings' secure and reliable work is of great significance. Toward this end, this paper presents a RUL prediction model based on a convolutional neural network (CNN) and bidirectional long short-term memory (BiLSTM) hybrid neural network, combined with the Wiener process. This method has two components: extraction of vibration signal features based on the CNN-BiLSTM model and RUL prediction of bearings using the Wiener process. Since the technique of constructing feature engineering after dimensionality reduction of the time-frequency features of the bearing may lose important signal information, thus, this paper tries to use the vibration acceleration signal of the bearing as the input feature and then use CNN and BiLSTM to build the bearing degradation model. Health index construction by the advantages of CNN for feature extraction and BiLSTM for processing sequence data. Considering the uncertainty of the bearing degradation process, finally, the Wiener process is used to deduce the probability density function (PDF) for predicting RUL to predict the RUL of the constructed health index model. The PHM 2012 bearing datasets confirm the validity and superiority of the presented method in this study.

1. Introduction

Rolling bearings are one of the most usual and indispensable parts of machinery and equipment. The degree of safety and reliability of machinery and equipment will be directly influenced by the bearings' health status [1,2]. For example, in the case of electric motors, about 40% to 50% of motor failures are caused by rolling bearings damage [3]. Mechanical failures reduce the productivity of enterprises and even endanger the personal safety of employees. As an introductory section of prognostics and health management (PHM), RUL has been widely concerned and applied to decrease the failure probability and ensure the safety equipment service. Rolling bearings RUL aims to analyze and predict failure time using information such as bearings monitoring signals and working conditions environment. Based on RUL, it is possible to determine the optimal timing of maintenance of the equipment and reduce the cost of use [4].

Research on bearings RUL is currently divided into two main approaches: mechanistic modeling and data-driven. The mechanical modeling approach is mainly based on the

* Corresponding Author: Junyu Guo
E-mail address: Junyguo@163.com, ORCID: https://orcid.org/0000-0001-9462-9501

Received: 11 September 2022; Revised: 20 October 2022; Accepted: 25 October 2022
https://doi.org/10.52547/crpase.8.4.2817

Academic Editor: **He Li**

Please cite this article as: J. L. Wan, Y. Yang, J. Guo, L. Dai, A hybrid CNN-BiLSTM and Wiener Process-based Prediction Approach of Remaining Useful life for Rolling Bearings, Computational Research Progress in Applied Science & Engineering, CRPASE: Transactions of Mechanical Engineering 8 (2022) 1–9, Article ID: 2817.

physical structure of the bearing to model the degradation of the equipment to predict the remaining life. The mechanical modeling approach needs a deep understanding of the bearing's structure, while the mechanism of rolling bearings has a certain complexity. Therefore, the approach requires a lot of expertise and assumptions, which leads to complex and costly modeling [5-7]. The data-driven approach focuses on predicting the remaining useful life of a bearing by fitting data on key characteristics that reflect the bearing's state of degradation using intelligent algorithms [8-10]. With the development of artificial intelligence technology, data-driven methods have become the prevailing approach for RUL prediction in recent years. Further, data-driven methods can be classified into statistical theory methods and methods based on artificial intelligence [11]. Typical statistical theory methods include methods based on Wiener process [12,13], methods based on Gamma process [14,15], methods based on Markov process [16,17], etc. The main methods based on artificial intelligence are methods based on support vector machines (SVM) and methods based on neural networks (NN), etc. Hasanipanah et al. [18] proposed a particle swarm optimization (PSO) and support vector regression (SVR) hybrid model algorithm for Air-overpressure (AOp) prediction. Liu et al. [19] predicted the RUL of rolling bearings by continuous type hidden Markov model (HMM) and PSO-SVM. Zhang et al. [20] used a full CNN to train a RUL prediction model. Han et al. [7] used BiLSTM to build a rolling bearing degradation model for bearing RUL prediction. Zhu et al. [21] proposed a deep feature learning method to predict bearings RUL, which through time-frequency representation (TFR) and multiscale convolutional neural network (MSCNN), compared with traditional CNN, MSCNN can learn more significant features, thereby improving the accuracy of prediction. Zheng et al. [22] stacked multilayer LSTMs to build RUL prediction models, the method addressed the problem of long-term dependence in Recurrent Neural Networks.

Although the above-mentioned data-driven techniques in the literature have yielded promising results in the application of RUL prediction for mechanical equipment, they generally have the following limitations: 1) The above-mentioned bearing life prediction methods usually require human manual extraction of time domain, frequency domain and statistical features [23-25], then screening and fusion, and finally regression prediction by machine learning, etc. In this process, the operation of extracting features is tedious and requires a lot of effort from researchers, and even important information may be lost. 2) For the complex and huge data, it is difficult for a single model structure to extract complex features sincerely [26]. 3) A single deep learning approach has difficulties deriving probability distribution functions that embody the uncertainty features of the remaining useful life. In view of the problems mentioned above, this article implements a combination of a statistical data-driven method and an approach based on deep learning to achieve RUL prediction for bearings.

In CNN networks, although CNN has powerful feature extraction capability, CNN cannot fully extract temporal features because the length of the convolutional kernel of CNN is fixed [27]. And because the input information flows in one direction, CNN cannot characterize the before-and-after correlation of degraded states [28]. In LSTM networks, although LSTM is suitable for processing time series [29], however, ordinary LSTM merely seizes the temporal dependence of signal in a sole temporal direction [30], which leads to room for improvement in prediction accuracy. A RUL prediction method by CNN and BiLSTM is presented using the raw vibration signal of the bearing as the input for RUL prediction. In bearing degeneration, the collected vibration signals are time-dependent, while they belong to time series data so that the BiLSTM can be used for RUL prediction. BiLSTM consists of both forward and backward LSTMs. BiLSTM can obtain contextual semantic information better than LSTM, so the advantage of processing time-series data by BiLSTM is used to input the spatial features extracted by CNN into the BiLSTM network further to extract temporal features for the construction of the health index. The CNN-BiLSTM model can eliminate the steps of feature extraction, screening, and fusion in traditional methods by taking advantage of its automatic extraction of temporal and spatial features. Premeditating that the health index degradation procedure is non-monotonic. At the same time, the Wiener process can characterize the non-monotonic degradation procedure. Therefore, we finally use the Wiener process to predict the RUL of the constructed health index. The Bayesian method is employed to assess the unknown parameters of the Wiener process.

Due to the mathematical properties of the Wiener process to characterize non-monotonic degradation processes, it is widely used in reliability assessment and RUL analysis of mechanical products [31]. Wang [32] studied the modeling of degradation data by a non-simultaneous Wiener process and proposed a pseudo-likelihood approach to model parameters estimation. Joseph and Yu [33] used the Wiener process to model the reliability of window wiper switches and develop reliability improvement strategies. Pan et al. [34] used the Wiener process model with truncated normal distribution to characterize the degradation process of the system. The analytical expressions for the PDF and the reliability function were derived. Wang and Feng [35] proposed a state-space model based on the Wiener process to describe the degradation process of lithium-ion batteries; the maximum likelihood estimation and unscented particle filter algorithm are employed for parameters estimation. Cai et al. [36] integrated the current and historical degradation data by the Wiener process for the RUL re-prediction for the subsea Christmas tree system. Zhang et al. [37] presented a review of the research and application of the Wiener process in modeling degradation data and RUL prediction.

The main contributions of this work are summarized as follows

(1) Contrast the health index with the traditional manual extraction of features from the vibration signal's time domain, frequency domain, and time-frequency domain. The CNN-BiLSTM prediction model constructed in this paper automatically extracts deep representative features in space and time from the original vibration signals of the bearings, without the need for manual feature extraction and selection.

(2) Unlike most prediction methods based on artificial intelligence, this paper uses the Wiener process to characterize the non-monotonic degradation process to

portray the non-linear health index of the bearing, and calculates the first-hitting-time to reach the distribution of the bearing's RUL. Furthermore, the confidence interval of the bearing's RUL is constructed to quantify the uncertainty in the bearing degradation process.

The remaining parts are organized as follows: Section 2 presents CNN and BiLSTM. Section 3 describes in detail the process of the proposed method for RUL prediction. Section 4 describes how the Wiener process can be used for bearing RUL prediction. Section 5 validates the proposed method through the bearing degradation experiment. And finally, conclusions are drawn, and future perspectives are presented in Section 6.

2. Basic Theory of Neural Networks

2.1. CNN

CNN is an effective feedforward neural network for feature extraction and pattern recognition first proposed by Lecun and Bengio [38] in deep learning research. CNN models have powerful image feature extraction capabilities, which use a few parameters to catch the spatial features of the input. Then the captured features are combined to form high-level data features, which are eventually fed into the fully-connected layers to achieve prediction. Usually, CNN consists of the input layer, convolution layer, pooling layer (subsampled layer), and fully connected layer. CNN structure is demonstrated in Figure 1.

2.1.1 Convolution Layer

The role of the convolution layer is to perform the convolution operation on the local region of the input data and the convolution kernel. Assuming that the input to the CNN model is X the output of the convolution layer is calculated as Eq. (1)

$$C_n = \sigma(W_n \otimes X + b_n) \tag{1}$$

where n is the number of convolution kernels, C_n denotes the nth feature map of the convolution layer output, $\sigma(\cdot)$ is the activation function, W_n represents the convolution kernel, b_n represents the bias term, and $\otimes$ is the convolution operator.

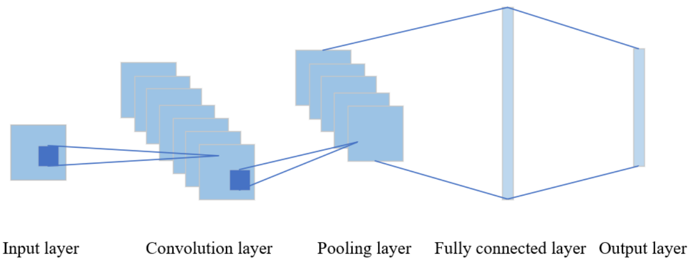

Figure 1. The typical CNN structure

2.1.2 Pooling Layer

The role of the pooling layer is to diminish the size of the feature maps and the computational complexity. Pooling layer can keep the most required information while reducing the dimensionality of features. Common pooling operations are average pooling and maximum pooling. In this paper, we use the more widely used maximum pooling, which is denoted as Eq. (2)

$$P_n = maxC_n \tag{2}$$

where P_n is the output of the pooling layer, C_n denotes the input for the pooling layer.

2.1.3 Fully Connected Layer

Each output neuron of the fully connected layer (FC layer) is connected with the input node. The input feature is combined with the FC layer, and then the activation function is used to get the predicted value. The FC layer is expressed as Eq. (3)

$$F = f\left(W_{fc}X_{fc} + b_{fc}\right) \tag{3}$$

where X_{fc} is the input feature maps of FC layer, W_{fc} represents the weights of fully connected layer, and b_{fc} denotes bias terms of FC layer.

2.2. BiLSTM

In CNN, since the information flows only in one direction, the CNN only considers the input of this time, while ignoring the previously degraded information during the computation. Consequently, CNN cannot characterize the before-and-after correlation of degraded states [39]. LSTM is a neural network with memory function, a variant of recurrent neural network (RNN). LSTM has excellent capability in the temporal data processing. LSTM also is widely used in natural language processing (NLP) and other fields. LSTM uses input gate, output gate, and forget gate to achieve control of the information [40]. A single LSTM neuron is shown in Figure 2.

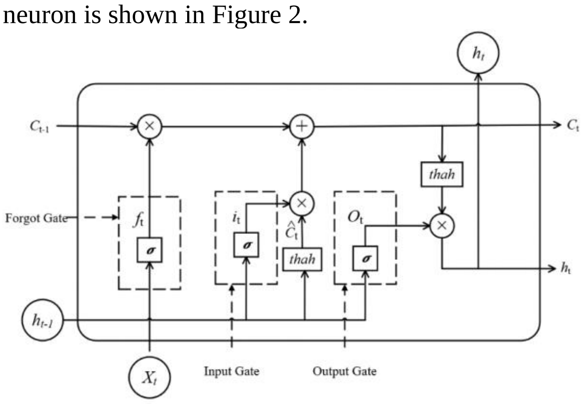

Figure 2. Structure of LSTM neuron

In Figure 2, σ denotes the activation function. LSTM adds a new cell state and three gate structures based on RNN [41]. The cell state is responsible for storing information. The activation function tanh and dot multiplication operations together form the gate structure.

The ordinary LSTM merely seizes the signal's temporal dependence in a single temporal direction. This paper integrates a bidirectional recurrent network architecture with LSTM into BiLSTM, which can learn hidden features in forwarding and backward directions [29]. BiLSTM is similar to the LSTM computational process. BiLSTM adds the inverse operation to LSTM to process the sequence in both forward and backward directions, and its network expansion diagram is shown in Figure 3. First, BiLSTM inverts the input sequence, then stacks the forward and backward LSTM in the same way as LSTM, and finally outputs them. Since the evolution of bearings failure evolves over time, there is a robust correlation between its features before and after, and the BiLSTM model can catch the past and future information at the same time, so BiLSTM is better at extracting the degraded features of bearings, thereby improving the accuracy of prediction.

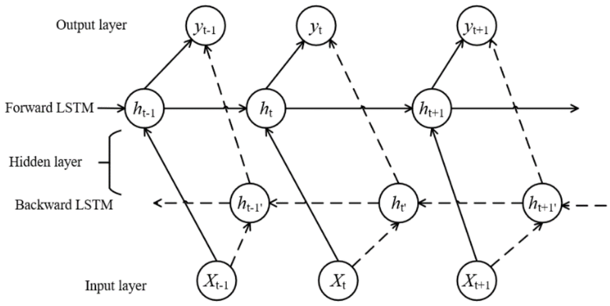

Figure 3. BiLSTM network expansion diagram

3. Rolling Bearings RUL Prediction Method

Figure 4 displays the process of rolling bearings RUL prediction method. The specific steps are as follows:

Step 1: Data acquisition. Acquisition of historical vibration signals from a large number of rolling bearings.

Step 2: Dividing the data set into a training set for training and a test set for testing, then normalizing the training set and the test set. The normalized vibration signal is used as the input of the CNN-BiLSTM model for deep feature extraction, and the degradation value is used as the label for the output of the model. The degradation value is a ratio. Assuming a specific bearing has a life of 22,000s, when the bearing operates to 11,000s, the degradation value is equal to 0.5.

Step 3: Set the CNN and BiLSTM network's network parameters. The CNN-BiLSTM model is used to extract spatial and temporal features from the input signal, and the output is performed through a FC layer to achieve the construction of health index based on the lifetime percentage. The network structure of the CNN-BiLTSM model is shown in Figure. 5. The network has 15 layers, including one input layer, three 1D convolutional layers, three maximum pooling layers for spatial feature extraction, two BiLSTM layers for temporal feature extraction, one Flatten layer for spreading operation, and one Dropout layer for overfitting prevention, and four fully connected layers for output.

Step 4: Test set validation. The test set data are input into the established model to obtain predicted degradation values and plot bearing degradation curve.

Step 5: According to the performance degradation trend of rolling bearings, the Wiener process predicts the RUL of rolling bearings.

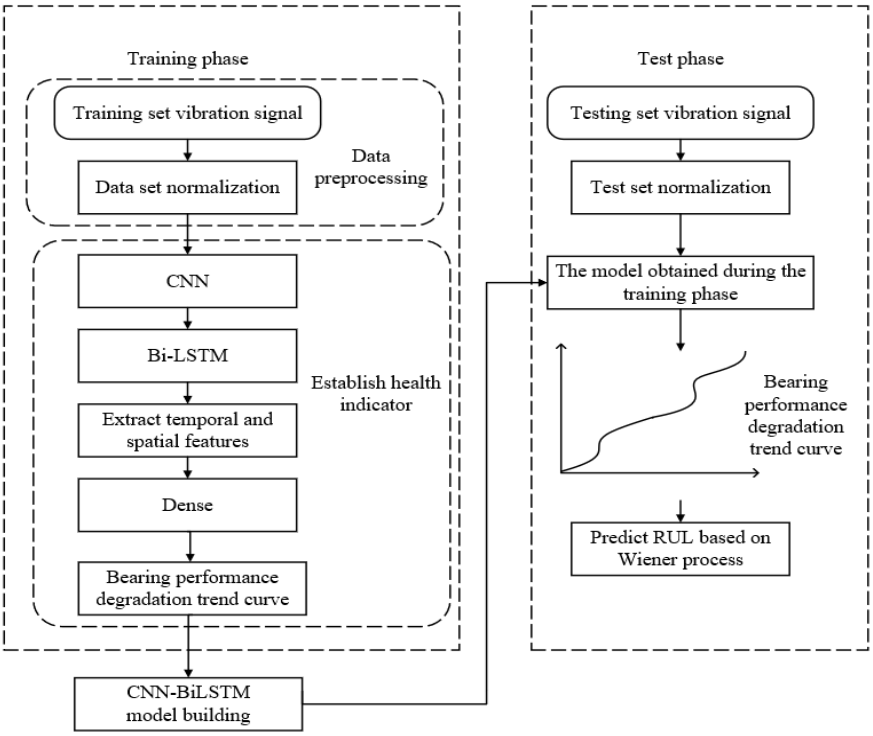

Figure 4. The process of rolling bearings RUL prediction method

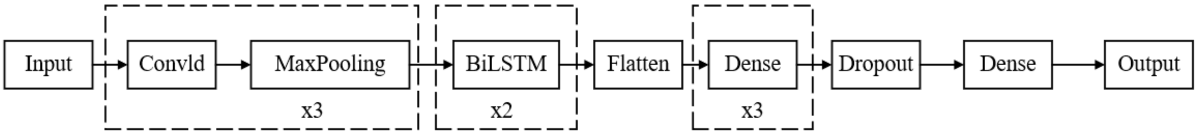

Figure 5. CNN-BiLTSM model network structure

4. Rolling Bearings RUL Prediction Based on Wiener Process

4.1. The Degradation Process

The advantage of the Wiener process over the Gamma process and the inverse Gaussian process is the ability to model non-monotonic degenerate processes. The Wiener process $\{Y(t)\,; t > 0\}$ could define as Eq. (4)

$$f(\Delta Y(t)|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi\Delta\tau(t)}} exp\left(-\frac{\left(\Delta Y(t) - \mu\Delta\tau(t)\right)^2}{2\sigma^2\Delta\tau(t)}\right) \quad (4)$$

Supposing that the performance degradation process of products obeys the Wiener process and is the failure threshold, the product's life time can be defined as the time of first reaching C: $T = inf\{t : Y\,(t) \geq C\}$. The life time T obeys the IG distribution: $T \sim IG(C/\mu, C^2/\sigma^2)$. The PDF and the CDF of T are Eqs. (5) and (6)

$$f(t|\mu,\sigma) = \frac{C}{\sqrt{2\pi t}\sigma t} exp\left(-\frac{(\mu t - C)^2}{2\sigma^2 t}\right) \quad (5)$$

$$F(t|\mu,\sigma) = \Phi\left(\frac{\mu t - C}{\sigma\sqrt{t}}\right) + exp\left(\frac{2\mu C}{\sigma^2}\right)\Phi\left(\frac{-C - \mu t}{\sigma\sqrt{t}}\right) \quad (6)$$

The remaining useful life T_k at any time T_0 can be defined as: $T = inf\{t : Y\,(T_0 + T_k) \geq C\}$. Define L_0 as the actual cumulative observed degradation value of the product's performance up to the time T_0. The PDF and CDF of the RUL T_0 are Eqs. (7) and (8)

$$f(t_k|\mu,\sigma) = \frac{C - L_0}{\sqrt{2\pi t_k}\sigma t_k} exp\left(-\frac{(\mu t_k - C + L_0)^2}{2\sigma^2 t_k}\right) \quad (7)$$

$$F(t_k|\mu,\sigma) = \Phi\left(\frac{\mu t_k - C + L_0}{\sigma\sqrt{t_k}}\right) + exp\left(\frac{2\mu(C - L_0)}{\sigma^2}\right)\Phi\left(\frac{-C - \mu t_k + L_0}{\sigma\sqrt{t_k}}\right) \quad (8)$$

4.2. Parameter Estimation

When the Wiener process is used to characterize the degradation process, the degradation increment Δd_{ij} is normally distributed: $\Delta d_{ij} \sim N\left(\mu\Delta t_{ij}, \sigma^2\Delta t_{ij}\right)$ and $\Delta t_{ij} = t_{ij} - t_{i,j-1}$. The sample number is denoted by i and the discrete observation times is characterized by j. The likelihood function of the degradation process can be expressed as Eq. (9)

$$L(D|\mu,\sigma) = \prod_{i=1}^{N}\prod_{j=2}^{M} f\left(\Delta d_{ij}|\mu,\sigma\right) = \prod_{i=1}^{N}\prod_{j=2}^{M}\frac{1}{\sigma\sqrt{2\pi\Delta t_{ij}}} exp\left(-\frac{\left(\Delta d_{ij} - \mu\Delta t_{ij}\right)^2}{2\sigma^2\Delta t_{ij}}\right) \quad (9)$$

where, $f(\cdot)$ is the PDF of the normal distribution.

Assuming that the joint prior distribution of the model is $\pi(\theta)$=π(μ,σ) , the joint posterior distribution of the parameters to be estimated in the model via Bayesian method can be expressed as Eq. (10)

$$p(\mu,\sigma|D) \propto \pi(\theta)L(D|\mu,\sigma) = \pi(\mu,\sigma)\prod_{i=1}^{N}\prod_{j=2}^{M}\frac{1}{\sigma\sqrt{2\pi\Delta t_{ij}}} exp\left(-\frac{\left(\Delta d_{ij} - \mu\Delta t_{ij}\right)^2}{2\sigma^2\Delta t_{ij}}\right) \quad (10)$$

5. Experiment Verification

5.1. Description of Data

The data set for this experiment comes from the IEEE 2012 PHM Data Challenge [42], and it is measured by the PRONOSTIA experimental platform at FEMTO-ST Institution in France. The bearing life degradation experimental platform is shown in Figure 6.

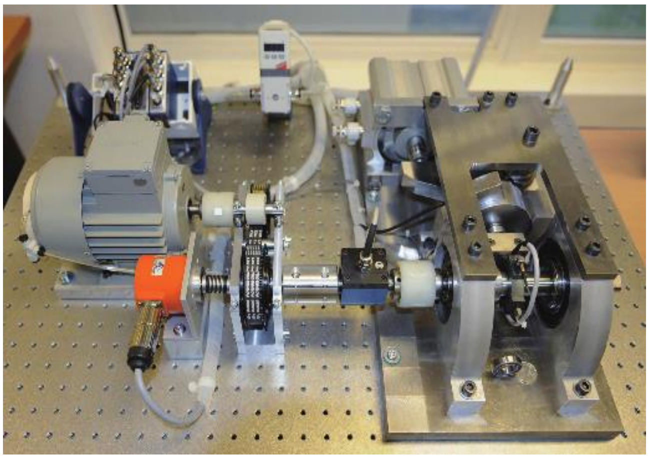

Figure 6. PRONOSTIA experimental platform

The dataset includes vibration signals for the full life cycle of 17 rolling bearings under three different operating conditions, as shown in Table 1. The data is collected every 10s, the sampling frequency is 25.6kHz, and one sampling time is 0.1s, so 2560 vibration accelerations can be obtained for each acquisition, as shown in Figure 7. Bearing failure is considered when the measured value of vibration acceleration reaches a predetermined threshold value. The dataset contains vibration acceleration signals in both horizontal and vertical directions. It is compared with the vertical vibration signal, and the horizontal vibration signal has more useful information [43], so only the horizontal vibration signal is used for the experiments in this paper. The bearing data named Bearing1_1 and Bearing1_2 under working condition 1 are chosen as the training set for the experiment to simulate the extreme lack of training data, containing 3674 sampling points. The test set is selected from the bearing data named Bearing1_3, which contains 2375 sampling points. For each experimental sample (X_i, Y_i), X_i is the value of the 2560th vibration acceleration acquired for the i th time. Y_i is the degree of degradation corresponding to this acquisition that the ratio between the current moment and the failure moment.

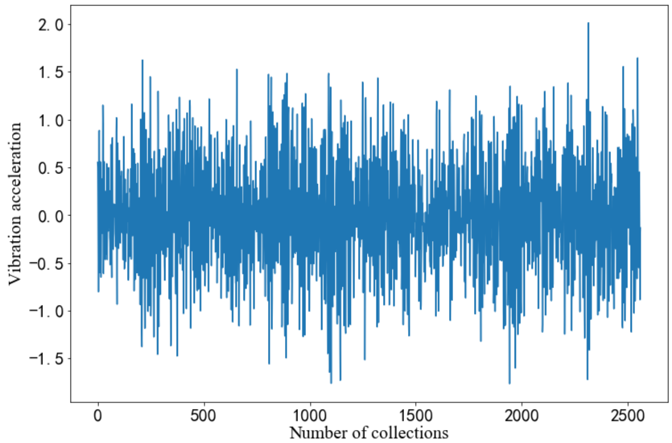

Figure 7. Vibration signal of the first sampling point in Bearing1_1

Table 1. Detailed information on working conditions

Working condition	Rotating speed(rpm)	Load
Working condition1	1800	4000
Working condition2	1650	4200
Working condition3	1500	5000

5.2. CNN-BiLSTM Model Parameter Setting

The input data of this experiment is a two-dimensional vibration signal of (2560, 1), and 2560 samples are selected each time for the experiment. The model uses the adam algorithm as the training parameter of the optimization algorithm, mean square error (MSE) is chosen as loss function, Rectified Linear Unit (ReLU) is used as the activation function. The activation function of the final output layer of the network uses the sigmoid activation function to control the final output result between [0, 1]. The overall structure of the model is shown in Figure 5, and the detailed parameters configuration of CNN-BiLSTM model network structure are summarized in Table 2. To avoid overfitting, the Dropout technique is used during the training process, and each experiment is conducted 50 epochs. MSE is used to evaluate the training effect, can well reflect the actual situation of the predicted value error, expressed as Eq. (11)

$$MSE = \frac{1}{N}\sum_{t=1}^{N}\left(y_t - \hat{y}_t\right)^2 \tag{11}$$

where y_t and $\hat{y}_t$ are the true and predicted values of bearing degradation, respectively. N is the total number of samples.

5.3. Experimental results and analysis

The CNN-BiLSTM model was fitted on the test set to obtain the fitting results of the actual degradation values and the predicted degradation values, and the results are shown in Figure 8.

Table 2. Parameters configuration of model network structure

Layer	size	Number	Activation function
Conv1	64×1	16	Relu
Pooling1	2×1	—	—
Conv2	3×1	32	Relu
Pooling2	2×1	—	—
Conv3	3×1	64	Relu
Pooling3	2×1	—	—
BiLSTM	—	100	tanh
BiLSTM	—	50	tanh
Dense1	—	128	Relu
Dense2	—	64	Relu
Dense3	—	32	Relu
Dense4	—	1	sigmoid

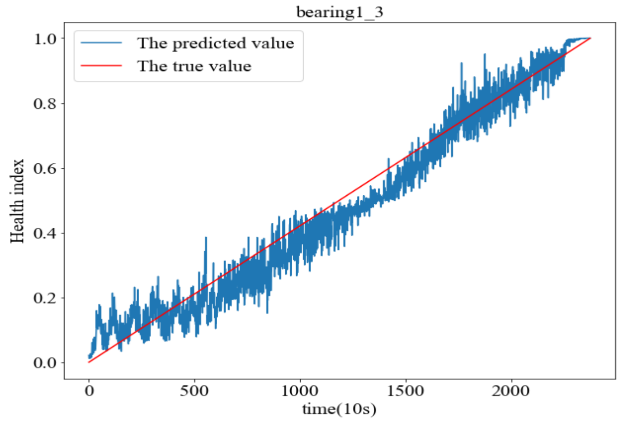

Figure 8. Fitting results on the test set

In Figure 8, the horizontal coordinate indicates the running time of Bearing1_3, the vertical coordinate indicates the health index of Bearing1_3, the actual degradation value is denoted by a slash, and the predicted degradation value is represented by a curve. As is shown in Figure 8, the CNN-BiLSTM model can better fit the degradation degree of the bearing, but the prediction results of some adjacent time points differ significantly. The prediction accuracy changes significantly with the local fluctuations of the samples, so the moving average (MA) method is used to do smoothing on the prediction results to reduce the impact of oscillations [44], and the moving window size is set to 20. The results after MA is shown in Figure 9.

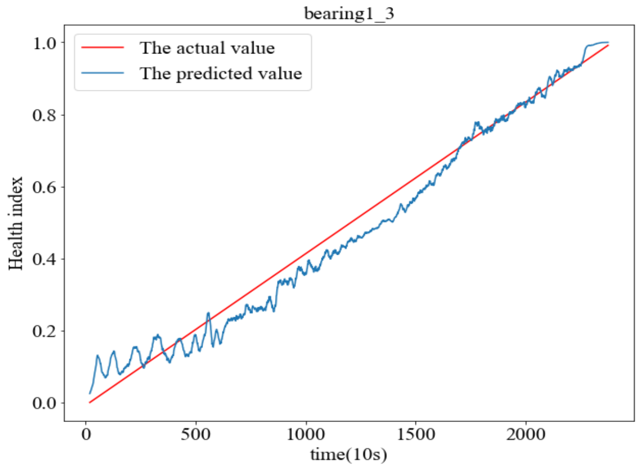

Figure 9. After MA result on the test set

From Figure 9, we can see the predicted value curve after noise reduction is closer to the actual remaining life line. To further validate the model effects, the experimental results are compared with those of four network models, SVR, MLP, CNN, BiLSTM, and CNN-LSTM, using the same experimental data and setting the same network parameters. For the purpose of quantitatively evaluating the prediction effect of the model, root mean square error (RMSE), mean absolute error (MAE), and correlation index (R2) were used as an evaluation index for each of the three models, can be expressed as Eqs. (12-14)

$$RMSE = \sqrt{MSE} \tag{12}$$

$$MAE = \frac{1}{N}\sum_{t=1}^{N}\left|y_t - \hat{y_t}\right| \tag{13}$$

$$R^2 = 1 - \frac{\sum_{t=1}^{N}\left(y_t - \hat{y_t}\right)^2}{\sum_{t=1}^{N}(y_t - \bar{y_t})^2} \tag{14}$$

where $\bar{y_t}$ is the average value of bearing degradation.

Each model is trained five times to take the average as the prediction result, and the final comparison of the validation outcome of the five models on the test set is shown in Table 3.

Table 3. Model effect comparison

Network model	Average RMSE	Average MAE	Average R2
SVR	0.194	0.155	—
MLP	0.157	0.115	—
CNN	0.072	0.061	89.05%
BiLSTM	0.093	0.079	81.86%
CNN-LSTM	0.067	0.056	90.48%
Proposed method	**0.061**	**0.051**	**92.13%**

In Table 3, As can be seen, the RMSE and MAE of the proposed method are the smallest, and R2 is the largest on the test set of the experiment. The proposed method reduced the average RMSE by 9.0% and the average MAE by 8.9% contrasted with the CNN-LSTM model, reduced the average RMSE by 15.3% and the average MAE by 16.4% contrasted with the CNN model, and reduced the average RMSE by 68.6% and the average MAE by 67.1% than the SVR model.This result shows that: 1) traditional machine learning methods have a relatively large gap in prediction accuracy compared with deep learning methods, which proves the superiority of deep learning methods. 2) Coupled models have better prediction results compared with single models. 3) The CNN-BiLSTM model can adequately extract deep features in the time and space of vibration signals, while the BiLSTM model is more effective and superior than the LSTM model. (4) The effectiveness and superiority of the CNN-BiLSTM model compared with other models are verified, and the CNN-BiLSTM model is more appropriate for the RUL prediction of bearings.

Representation of the CNN-BiLSTM model's predicted degradation value through health index, the health index of Bearing1_3 has substituted into the Wiener process to analyze the RUL distributions, and the Bayesian method is used for parameters estimation. Since the illustrative example used in this paper has no prior information, the uninformative prior distribution is used to avoid subjective factors' effect on the reliability results. The uninformative prior distribution of the Wiener process parameters is shown as: $\mu \sim U(0,1), \sigma \sim U(0,1)$. The 20000 samples of model parameters of joint posterior distribution are generated by the Markov chain Monte Carlo method.

The results for parameter estimation based on the degradation data up to 21000s are given in Table 4.

The PDFs of posterior distributions of the Wiener process model parameters are shown in Figure 10.

Table 4. Statistical summarization of parameters estimation

	Mean	Standard deviation	Confidence interval	
			2.5%	97.5%
μ	0.00102	0.0007221	0.00004691	0.002701
σ	0.04922	0.0007652	0.004775	0.05074

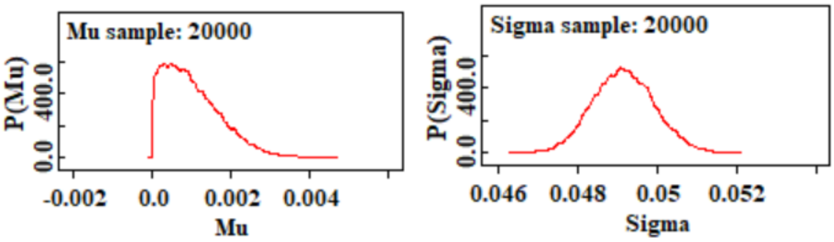

Figure 10. The PDFs of posterior distributions of model parameters: μ and σ

The results of the RUL prediction of Bearing 1_3 at 19000 and 21000 time points are summarized in Figure 11. As shown in Figure 11, it can be observed that as the number of observations grows, the predicted RUL is getting closer to the true RUL. These results can also validate the effectiveness of the proposed method in this paper.

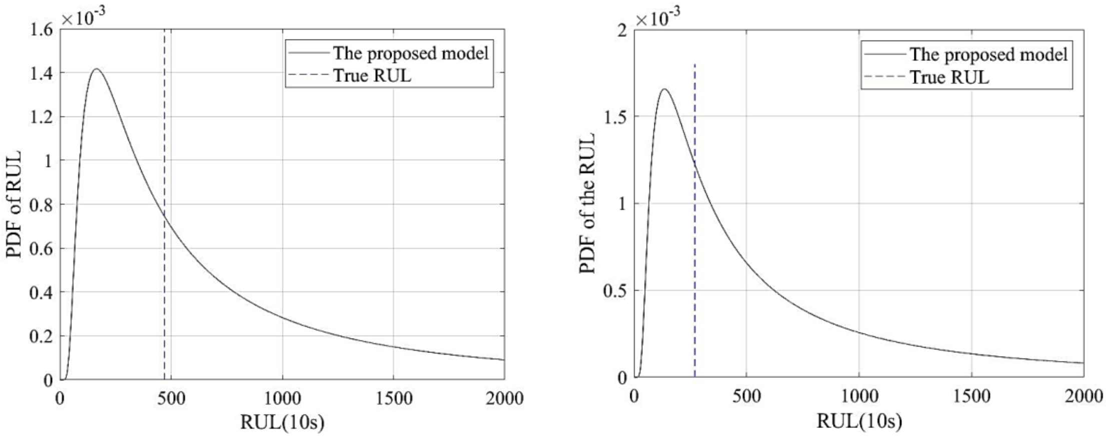

Figure 11. Results of RUL prediction for Bearing 1_3. (a) At time $t_{p1} = 19000s$. (b) At time $t_{p2} = 21000s$

6. Conclusion

In this paper, we propose a deep learning network model for rolling bearings RUL prediction: CNN-BiLSTM model. For rolling bearings, the vibration level acceleration signal in the bearings dataset is first used as input to build a CNN-BiLSTM network model to fully extract the deep features of the vibration signal in time and space. The established model is used to predict the degradation value of test set bearings. Then according to the relationship between degradation value and RUL, the Wiener process is used to obtain the RUL of the corresponding bearing. Finally, the proposed model was experimentally validated on the PHM2012 bearing degradation dataset and compared with some existing network models. The experimental results show that the rolling bearings degradation model based on CNN-BiLSTM can better fit the degradation curve of the bearing, and the Wiener process can accurately predict the RUL of bearings. It proves the effectiveness of the method.

In the future, since there is a large amount of noise in the original vibration signal and the data fluctuation range is large, the noise reduction process will be carried out in future work. And the proposed method in this paper is carried out in a single working condition, while there are a large number of variable loads working conditions in industrial production, so the subsequent research on RUL prediction under multiple working conditions will be carried out.

Acknowledgements

This research was supported by Sichuan Province Regional Innovation Cooperation Project under the contract No. 22QYCX0033, Scientific Research Starting Project of SWPU under the contract No.2019QHZ007.

References

[1] S. Patil, A. Patil, V. Handikherkar, S. Desai, V. M. Phalle, F. S. Kazi, Remaining useful life (RUL) prediction of rolling element bearing using random forest and gradient boosting technique, American Society of Mechanical Engineers, Available from: (https://doi.org/10.1115 /IMECE2018-87623), 2018.

[2] L. Chen, J. Tan, X. Jiang, Degradation trend prediction of rolling bearings based on TESPAR and LS-SVM algorithm, Coal Mine Machinery 38.8 (2017) 18–20.

[3] A. Rai, S.H. Upadhyay, The use of MD-CUMSUM and NARX neural network for anticipating the remaining useful life of bearings, Measurement 111 (2017) 397–410.

[4] C.G. Huang, H.Z. Huang, Y.F Li, A bidirectional LSTM prognostics method under multiple operational conditions, IEEE Transactions on Industrial Electronics 66.11 (2019) 8792–8802.

[5] R.K. Singleton, E.G. Strangas, S. Aviyente, Extended Kalman filtering for remaining-useful-life estimation of bearings, IEEE Transactions on Industrial Electronics 62.3 (2014) 1781–1790.

[6] Y. Liu, Q. Liu, C. Xie, F. Wei, Reliability assessment for multi-state systems with state transition dependency, Reliability Engineering & System Safety 188 (2019) 276–288.

[7] L. Han, C. Shi, J. Zhang, Prediction for remaining useful life of rolling bearings based on Bi-directional long short-term memory, Manufacturing Automation 42.5 (2020) 47–50.

[8] Z. Wang, J. Cao, X. Ma, H. Qiu, Y. Zhang, H. Fu, S. Krishnaswamy, An improved independent increment process degradation model with bilinear properties, Arabian Journal for Science and Engineering 42.7 (2017) 2927–2936.

[9] R. Yuan, M. Tang, H. Wang, H. Li, A reliability analysis method of accelerated performance degradation based on bayesian strategy, IEEE Access 7 (2019) 69047–169054.

[10] H. Xiao, S. Gao, L.H. Lee, Simulation budget allocation for simultaneously selecting the best and worst subsets, Automatica 84 (2017) 117–127.

[11] C.H. Hu, H. Pei, X.S. Si, D. B. Du, Z. N. Pang, X. Wang, A prognostic model based on DBN and diffusion process for degrading bearing, IEEE Transactions on Industrial Electronics 67.10 (2019) 8767–8777.

[12] Q. Zhai, Z. S. Ye, RUL prediction of deteriorating products using an adaptive Wiener process model, IEEE Transactions on Industrial Informatics 13.6 (2017) 2911–2921.

[13] N. Li, Y. Lei, T. Yan, N. Li, T. Han, A Wiener-process-model-based method for remaining useful life prediction considering unit-to-unit variability, IEEE Transactions on Industrial Electronics 66.3 (2018) 2092–2101.

[14] M. Abdel-Hameed, A gamma wear process, IEEE transactions on Reliability 24.2 (1975) 152–153.

[15] M.H. Ling, H.K.T. Ng, K.L. Tsui, Bayesian and likelihood inferences on remaining useful life in two-phase degradation models under gamma process, Reliability Engineering & System Safety 184 (2019) 77–85.

[16] J. Yu, Health condition monitoring of machines based on hidden Markov model and contribution analysis, IEEE Transactions on Instrumentation and Measurement 61.8 (2012) 2200–2211.

[17] Z. Liu, Q. Li, X. Liu, C. Mu, A hybrid LSSVR/HMM-based prognostic approach, Sensors 13.5 (2013) 5542–5560.

[18] M. Hasanipanah, A. Shahnazar, H. Bakhshandeh Amnieh, D. Jahed Armaghani, Prediction of air-overpressure caused by mine blasting using a new hybrid PSO–SVR model, Engineering with Computers 33.1 (2017) 23–31.

[19] B. Liu, Q. Ning, C. X. Liu, Q. Ai, P. He, Residual life prediction of rolling bearings based on continuous hidden Markov model and PSO-SVM, Journal of Computer Applications 39 (2019) 31–35.

[20] J. D. Zhang, Y. S. Zou, J. L. Deng, X. L. Zhang, Bearing remaining life prediction based on full Convolutional layer neural networks, China Mechanical Engineering 30.18 (2019) 2231–2235.

[21] J. Zhu, N. Chen, W. Peng, Estimation of bearing remaining useful life based on multiscale convolutional neural network, IEEE Transactions on Industrial Electronics 66.4 (2018) 3208–3216.

[22] S. Zheng, K. Ristovski, A. Farahat, C. Gutpta, Long short-term memory network for remaining useful life estimation, IEEE international conference on prognostics and health management (ICPHM). IEEE (2017) 88–95.

[23] F. Camci, K. Medjaher, N. Zerhouni, P. Nectoux, Feature evaluation for effective bearing prognostics, Quality and reliability engineering international 29.4 (2013) 477–486.

[24] H. Ocak, K. A Loparo, F. M. Discenzo, Online tracking of bearing wear using wavelet packet decomposition and probabilistic modeling: A method for bearing prognostics, Journal of sound and vibration 302 (2007) 951–961.

[25] L. Guo, N. Li, F. Jia, Y. Lei, J. Lin, A recurrent neural network-based health indicator for remaining useful life prediction of bearings, Neurocomputing 240 (2017) 98–109.

[26] D. Li, L. Yang, Remaining Useful Life Prediction of Lithium Battery Based on Sequential CNN–LSTM Method, Journal of Electrochemical Energy Conversion and Storage 18.4 (2021).

[27] Y. Cao, Y. Ding, M. Jia, R. Tian, A novel temporal convolutional network with residual self-attention mechanism for remaining useful life prediction of rolling bearings, Reliability Engineering & System Safety 215 (2021) 107813.

[28] B. Wang, Y. Lei, T. Yan, N. Li, L. Guo, Recurrent convolutional neural network: A new framework for remaining useful life prediction of machinery, Neurocomputing 379 (2020) 117–129.

[29] Y. Cheng, K. Hu, J. Wu, H. Zhu, X. Shao, A convolutional neural network-based degradation indicator construction and health prognosis using bidirectional long short-term memory network for rolling bearings, Advanced Engineering Informatics 48 (2021) 101247.

[30] X. Ma, J. Zhang, B. Du, C. Ding, L. Sun, Parallel architecture of convolutional bi-directional LSTM neural networks for network-wide metro ridership prediction, IEEE Transactions on Intelligent Transportation Systems 20.6 (2018) 2278–2288.

[31] S. Gan, Z. Song, Reliability Evaluation by a Dependent Competing Failure Model Including a Time-Varying Rate for Sudden Degradation Increments, Arabian Journal for Science and Engineering 46.10 (2021) 10137–10152.

[32] X. Wang, Semiparametric inference on a class of Wiener processes, Journal of Time Series Analysis 30.2 (2009) 179-207.

[33] V.R. Joseph, I.T. Yu, Reliability improvement experiments with degradation data, IEEE Transactions on Reliability 55.1 (2006) 149–157.

[34] D. Pan, J. B. Liu, F. Huang, J. Chao, A. Alsaedi, A Wiener process model with truncated normal distribution for reliability analysis, Applied Mathematical Modelling 50 (2017) 333–346.

[35] R. Wang, H. Feng, Lithium-ion batteries remaining useful life prediction using wiener process and unscented particle filter, Journal of Power Electronics 20.1 (2020) 270–278.

[36] B. Cai, H. Fan, X. Shao, Y. Liu, G. Liu, Z. Liu, R. Ji, Remaining useful life re-prediction methodology based on Wiener process: Subsea Christmas tree system as a case study, Computers & Industrial Engineering 151 (2021) 106983.

[37] Z. Zhang, X. Si, C. Hu, Y. Lei, Degradation data analysis and remaining useful life estimation: A review on Wiener-process-based methods, European Journal of Operational Research 271.3 (2018) 775–796.

[38] Y. LeCun, Y. Bengio, Convolutional networks for images, speech, and time series, The handbook of brain theory and neural networks 3361.10 (1995) 1995.

[39] G. Zhao, P. Jiang, T. Lin, Remaining life prediction of rolling bearing based on CNN-BiLSTM model with attention mechanism, Journal of Mechanical & Electrical Engineering 38.10 (2021) 1253–1260.

[40] Z. Cao, C. Ye, Prediction of bearing remaining useful life based on parallel CNN-SE-Bi-LSTM, Application Research of Computers 38.07 (2021) 2103-2107.

[41] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural computation 9.8 (1997) 1735–1780.

[42] P. Nectoux, R. Gouriveau, K. Medjaher, E. Ramasso, B. Chebel-Morello, N. Zerhouni, C. Varnier, PRONOSTIA: An experimental platform for bearings accelerated degradation tests, IEEE International Conference on Prognostics and Health Management, PHM'12. IEEE Catalog Number: CPF12PHM-CDR (2012) 1–8.

[43] A. Soualhi, K. Medjaher, N. Zerhouni, Bearing health monitoring based on Hilbert–Huang transform, support vector machine, and regression, IEEE Transactions on instrumentation and measurement 64.1 (2014) 52–62.

[44] Y. Hu, H. Li, X. Liao, E. Song, Performance Degradation Model and Prediction Method of Real-time Remaining Life for Wind Turbine Bearings, Proceedings of the CSEE 36.6 (2016) 1643–1649.