

# Life prediction method of rolling bearing based on CNN-LSTM-AM

Wanqing Huang<sup>1</sup>, Yang Chen<sup>2</sup>, Yongqi Chen<sup>3</sup>, Tao Zhang<sup>4</sup>

College of Science and Technology, Ningbo University, Ningbo, 315211, P. R. China

<sup>2</sup>Corresponding author

**E-mail:** <sup>1</sup>2246356498@qq.com, <sup>2</sup>chenyang@nbu.edu.cn, <sup>3</sup>chenyongqi@nbu.edu.cn, <sup>4</sup>7762752193@qq.com

Received 13 November 2023; accepted 21 April 2024; published online 23 May 2024

DOI <https://doi.org/10.21595/jve.2024.23793>



Copyright © 2024 Wanqing Huang, et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Abstract.** Bearing is the key component to determine the health of machinery, and it is of great significance to monitor its working status in real time and predict its remaining useful life. In recent years, the RUL prediction method based on deep learning has been widely used and achieved good prediction results. Here, a bearing life prediction method based on convolution neural network (CNN), long short term memory (LSTM) and attention mechanism (AM) is proposed. First of all, the time domain and frequency domain features of the original vibration signals of rolling bearings are extracted, and the extracted feature set is normalized as the input of CNN. The main function of CNN is to extract spatial features and reduce the dimension of the data. Then, using LSTM to extract the information that may be ignored by CNN, the feature information extracted by CNN-LSTM is input to the attention mechanism for weighting, and the key information is screened. And then more accurately represent the degradation characteristics of the equipment, and finally get the bearing remaining life. The performance of the model is verified by two sets of public data sets, and the experimental results show that it is compared with the CNN-LSTM method. The root mean square error (RMSE) index based on CNN-LSTM-AM method is reduced by 14.6 % and 13.8 % respectively, and the score index is increased by 2.0 % and 2.4 % respectively. The results show that the proposed method has higher accuracy in bearing RUL prediction.

**Keywords:** remaining useful life, convolution neural network, long-short time memory network, attention mechanism.

## 1. Introduction

As an important part of rotating machinery, the health status of rolling bearings is directly related to the working state of the whole rotating machinery [1]. According to statistics, in the industrial rotating machinery equipment using rolling bearings, about 51 % of the mechanical failures are related to bearing damage [2]. Once the bearing failure occurs, it will seriously affect the safety of the mechanical system, and even cause serious economic losses [3-4]. Therefore, the prediction of remaining useful life of rolling bearings is an important guarantee for equipment operation and maintenance. Reasonable maintenance can be carried out by using the predicted rolling bearing remaining useful life, so as to avoid major safety accidents [5-6].

The main bearing RUL prediction methods based on vibration signals are model-based method and data-driven method [7]. The former can accurately predict the residual life of rolling bearings, but it has some disadvantages such as difficult modeling and low prediction accuracy for complex mechanical systems [8-10]. The data-driven method can study the mapping law between data features and residual life through machine learning method, which has the advantages of good universality and high economy, so it has become the focus of current RUL prediction [11]. Data-driven methods based on shallow machine learning, such as support vector machine (SVM) [12], artificial neural network (ANN) [13], k-nearest neighbor (KNN) [14], Markov model [15] et al. need to use complex signal processing techniques to extract features in advance before the mapping between fault evolution and health state can be obtained. There is a lack of strength in

the analysis of massive and complex data. On the other hand, the deep neural network model based on deep learning can deeply mine the feature information contained in massive data without complex signal preprocessing technology. Wang et al. [16] use recurrent convolution neural network (RCNN) to predict the remaining service life of machines. Liu et al. [17] proposed a method of remaining service life prediction based on LSTM and RNN. An et al. [18] introduced a hybrid model that combines CNN with superimposed bi-directional and unidirectional LSTM networks to deal with sequence data in remaining useful life prediction tasks. Ren et al. [19] proposed a multi-scale dense gate recursive unit network for predicting the remaining service life of the Internet of things. Yu et al. [20] proposed a similarity curve matching method based on bi-directional RNN and automatic encoder, which can be used to estimate the remaining service life of mechanical systems.

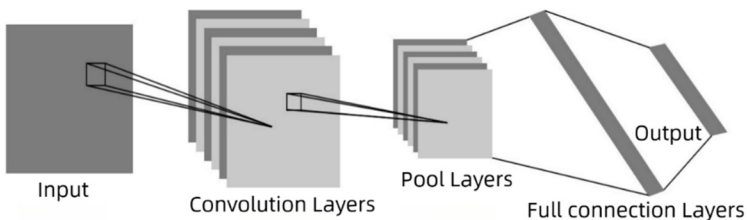
Based on the analysis of the above research, it is found that there are two limitations: the time dependence in different degradation states and the different effects of different characteristics on the results of residual life prediction are not taken into account. The key features that have great influence on the results of residual life prediction should be found.

In order to improve the prediction accuracy of bearing residual life, a bearing life prediction model based on CNN-LSTM combined with attention mechanism is proposed. The main innovations are as follows: 1) this method uses CNN to extract local spatial information, obtains time characteristic information through LSTM, and fully mines the potential feature information between data. 2) The attention mechanism is used to assign different weights to the characteristic information extracted by CNN-LSTM, highlight the key feature information, filter the useless information, express the degraded feature information of the equipment more accurately, and finally predict the remaining life of the bearing.

## 2. Theory

### 2.1. CNN

Convolution neural network is a kind of artificial neural network based on the biological “receptive field” mechanism, which is essentially a mapping from input to output, and can deal with image, voice, text and other data. it is widely used in image recognition, speech processing and fault diagnosis [21].



**Fig. 1.** Schematic diagram of CNN structure

As shown in Fig. 1 is a typical CNN model structure [22], which consists of an input layer, a number of convolution layers, pooling layers, a fully connected layer and an output layer, while the convolution layer with feature extraction function is the core of CNN. The convolution layer convolution the input data through a number of different convolution kernels, coupled with bias to extract local features, which can enhance the significant features of the original signal. The specific calculation model of convolution layer is as follows:

$$x_j^l = f \left( \sum_{i \in M_j} x_j^{l-1} * k_{ij}^l + b_j^l \right), \quad (1)$$

where  $x_j^l$  is the  $j$ th characteristic of the output of the  $l$ th convolution layer,  $x_j^{l-1}$  is the output of the previous convolution layer,  $k_{ij}^l$  is the kernel matrix of the convolution layer,  $b_j^l$  is the offset of the convolution of the  $l$  layer; \* represents the convolution operation;  $f(\cdot)$  is the activation function.

The function of the pooling layer is to reduce the dimension of the features extracted by the convolution layer, reducing the amount of data while retaining useful information. CNN has three sampling operations, the most common of which is maximum pooling. The mathematical formula is as follows:

$$y_i^{l+1}(j) = \max_{k \in D_j} \{x_i^l(k)\}, \quad (2)$$

where  $y_i^{l+1}(j)$  is the output value of the pooled layer,  $x_i^l(k)$  is the output value of the pooled layer,  $D_j$  is the width of the pooled core region.

## 2.2. LSTM

When dealing with sequence data, cyclic neural network is the most commonly used, while RNN is easy to cause gradient disappearance or gradient explosion when the input sequence is too long.

Long-term memory network is a special kind of RNN, which can learn long-term dependence and effectively solve the gradient problem in RNN. LSTM realizes the information processing in RNN through the gate control unit, and uses the switching degree of the door to determine the importance of the information to read, write or erase. The schematic diagram of the cellular structure of LSTM neurons is shown in Fig. 2.

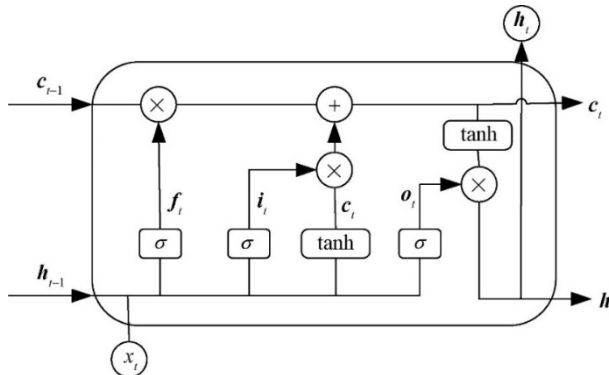


Fig. 2. Unit structure diagram of LSTM

The function and calculation process of each door of LSTM are as follows:

– The function of the forgetting door is to filter the information output from the previous state, that is, to delete the unimportant information:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \quad (3)$$

where  $f_t$  is the output state of the  $t$ -moment amnesia gate,  $\sigma(\cdot)$  is the Sigmoid activation function,  $W_f$  is the weight matrix,  $h_{t-1}$  is the output of the previous moment,  $x_t$  is the input of the current moment,  $b_f$  is the bias of the amnesia gate.

– The input gate determines whether the new information about the current state is useful and adds it to the cellular state:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \quad (4)$$

$$\tilde{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c), \quad (5)$$

where  $i_t$  is the output value of the  $t$ -time input gate,  $W_i$  is the weight matrix of the input gate,  $b_i$  is the bias matrix of the input gate,  $\tilde{c}_t$  is the cell state information of the input gate,  $\tanh(\cdot)$  is the activation function,  $W_c$  is the weight of the cell state,  $b_c$  is the bias of the cell state.

– To update the status is to calculate the state value of the memory unit at the current time:

$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t, \quad (6)$$

where  $c_t$ ,  $c_{t-1}$  are the cell state information at  $t$ -time and the last time, respectively.

– Outputs choose important information from the current state as output:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \quad (7)$$

$$h_t = o_t * \tanh(c_t), \quad (8)$$

where  $o_t$  is the output value of  $t$ -moment,  $W_o$  is the weight matrix of the output gate,  $b_o$  is the bias matrix of the output gate,  $h_t$  is the output information of the current hidden state.

### 2.3. Attention mechanism

Attention mechanism is originated from the study of human vision, was proposed and applied by Bahdanau et al. [23] in 2014. In the case of limited computing power, Attention Mechanism, as a resource allocation scheme, can highlight important information. When the neural network is used to process a large amount of information, only the key information input can be selected for processing, so as to improve the efficiency of the neural network. In essence, the Attention mechanism is the weighted summation of the value of the elements in the Source, while Query and Key are used to calculate the weight coefficients of the corresponding Value. The specific calculation process is as follows:

(1) Calculate the similarity or correlation between Query and Key.

(2) Normalize the score of step (1) by introducing Softmax, the original calculated score is sorted into a probability distribution in which the sum of the weights of all elements is 1, and the weights of important elements can be more highlighted.

(3) Weight and sum the Value according to the weight coefficient, and the calculation formula is as follows:

$$Attention(Q, K, V) = V \cdot \text{softmax}\left(\frac{K^T Q}{\sqrt{D_K}}\right), \quad (9)$$

where  $Q$ ,  $K$  and  $V$  are the matrices composed of query vector, key vector and value vector, respectively.  $D_K$  is the vector dimension of  $Q$ ,  $\text{softmax}(\cdot)$  is a normalized function.

### 3. The proposed method

In order to further improve the accuracy of rolling bearing RUL prediction and the efficiency of model training, a rolling bearing RUL prediction model based on CNN-LSTM-AM is established. The algorithm flow of the rolling bearing RUL prediction model proposed in this paper is shown in Fig. 3.

The specific steps for life prediction are as follows: Firstly, 14 time domain and 5 frequency domain features are extracted from the original vibration signal of rolling bearing to form feature vector, which is divided into training set and test set according to the proportion of 7:3.

In order to eliminate the influence of different dimensions and normalize it, the data is scaled

to the range of  $[0, 1]$ , and then the standardized deformed time series data set is extracted by convolution neural network to extract the local spatial state features and feature dimensionality reduction, and the output data is obtained.

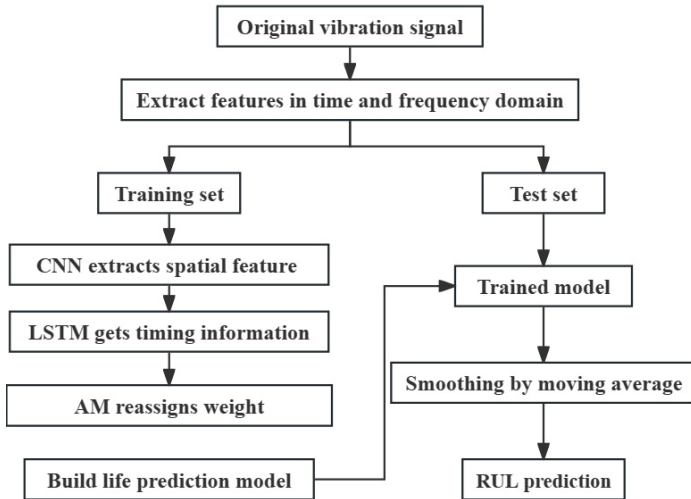


Fig. 3. Life prediction process of rolling bearing

Taking the output data of CNN as the input of LSTM network, we can learn the potential time relationship of sequence data through LSTM layer, mine its deep features effectively, and extract information from time series data that may be ignored by CNN.

Input the output data of the hidden layer of the LSTM network into the Attention layer, use the attention mechanism to calculate the correlation between each Q and K, then normalize and redistribute the weights through the softmax function, and add the weights according to the weight coefficients.

Finally, the output of the attention mechanism is transformed into one-dimensional data through the flat layer, and the full connection layer is used to map the bearing degradation characteristics to the remaining service life label, the trend quantitative health index is obtained, and the life prediction model is established, so as to realize the RUL prediction of rolling bearings.

## 4. Experimental verification

### 4.1. IMS Bearing datasets

In order to verify the effectiveness of the proposed method, the rolling bearing life cycle degradation data set published by the University of Cincinnati in PHM data Challenge in 2012 was selected [24]. The data set is collected from the PRONOSTIA test rig, which is shown in Fig. 4 as the structure diagram of the test rig. The platform is composed of AC motor, hydraulic loading system and PCB353B33 acceleration sensor. The rotational speed of the motor is 2000 rpm and the test bearing type is Rexnord ZA-2115 double-row bearing. The two accelerometers collect the bearing vibration signals on the X-axis and Y-axis respectively. The sampling frequency of the sensor is 20 KHz, and each sampling time is 1 s. Each file contains 20480 data of sampling points.

The experimental platform collects the life cycle degradation data of 17 bearings under 3 working conditions, and the bearing data set under each working condition is described in Table 1. In this paper, the life data of rolling bearings under condition 1 are selected for experimental verification.

PHM2012 rolling bearing data includes data in both horizontal and vertical directions.

However, Singleton et al. [25] found that the horizontal vibration signal can provide more degradation information than the vertical vibration signal, so this paper only uses the rolling bearing vibration data collected in the horizontal direction for the test.

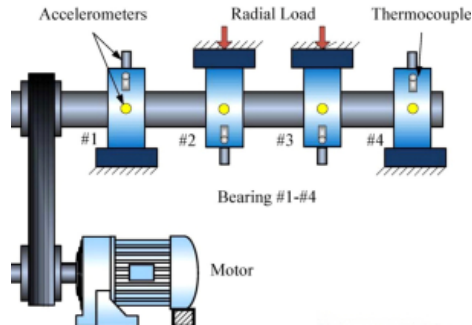


Fig. 4. PRONOSTIA test rig

Table 1. Specific description of the dataset

Working condition	Bearing number	Rotational speed / (r.min <sup>-1</sup> )	Load / N
Condition 1	1-1, 1-2, 1-3, 1-4, 1-5, 1-6, 1-7	4000	1800
Condition 2	2-1, 2-2, 2-3, 2-4, 2-5, 2-6, 2-7	4200	1650
Condition 3	3-1, 3-2, 3-3	5000	1500

Taking bearing1-3 in the IMS data set as an example, the full life cycle time domain waveform of the vibration signal of bearing 1-3 in the horizontal direction is shown in Fig. 5. It can be seen from Fig. 5 that the amplitude of the time domain vibration signal of the rolling bearing 1-3 increases gradually with the passage of running time, which is beneficial to the extraction of health feature information with the trend of degradation. At the same time, the root mean square value (RMS) of bearing1-3 is extracted, and its degradation trend is shown in Fig. 6. By observing Fig. 6, we can see that the RMS value of bearing 1-3 increases slowly from 0, and then increases rapidly, which well reflects the degradation trend of bearing 1-3 in the process of operation. Therefore, this paper selects 14 time domain features, such as average, amplitude and peak, and 5 frequency domain features, such as root mean square frequency and average frequency, as the input of the life prediction model.

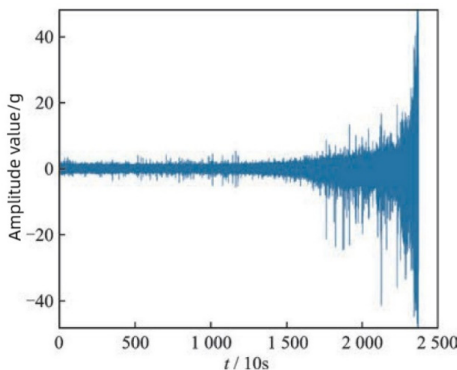


Fig. 5. Bearing 1-3 full life cycle time domain waveform

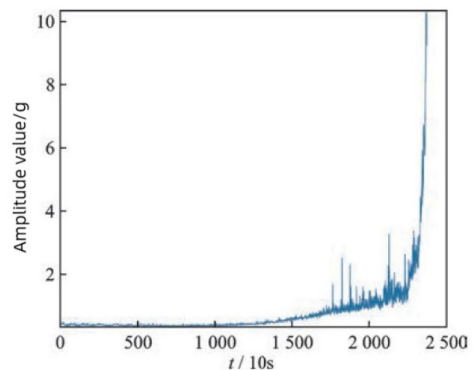


Fig. 6. Trend of root mean square of bearing 1-3

#### 4.2. Performance degradation index

In order to quantitatively evaluate the effect of the model for bearing RUL prediction, this

paper takes the root mean square error (RMSE) and Scoring function between the predicted value and the real value of rolling bearing remaining life as the evaluation index. Among them, RMSE can accurately reflect the average deviation between the predicted value and the real value. The smaller the RMSE, the better the performance of the model. The calculation formula is as follows:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N [f(i) - \hat{f}(i)]^2}, \quad (10)$$

$$Er_i = 100 \times \frac{R_i - \hat{R}_i}{R_i}, \quad (11)$$

$$A_i = \begin{cases} \exp^{-\ln(0.5) \cdot (Er_i/5)}, & Er_i \leq 0, \\ \exp^{-\ln(0.5) \cdot (Er_i/20)}, & Er_i > 0, \end{cases} \quad (12)$$

$$Score = \frac{1}{N} \sum_{i=1}^N (A_i), \quad (13)$$

where  $N$  is the number of samples,  $f(i)$  is the true value of the current sample,  $\hat{f}(i)$  is the predicted value of the model,  $R_i$  is the actual value of the remaining life of the bearing,  $\hat{R}_i$  is the predicted value of the remaining life of the bearing,  $Er_i$  is the percentage error of the  $i$ th sample.

## 5. Analysis of experimental results

After many times of debugging, the specific structure of the CNN-LSTM network model is set as follows: the model includes three convolution attention layers, three pooling layers, three LSTM layers and three full connection layers. The first convolution layer uses a larger convolution kernel. In order to extract short-term local features with a larger field of vision, the rest of the convolution layer uses a small convolution kernel in order to reduce the parameters of the network, facilitate deepening the network and restrain over-fitting.

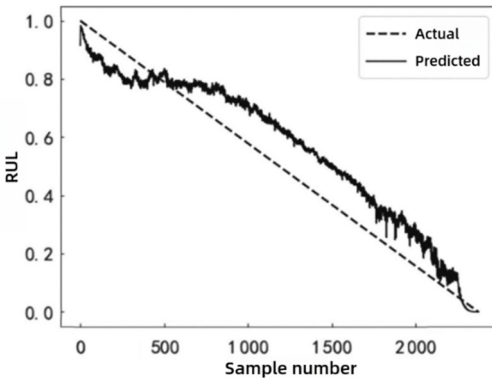


Fig. 7. Prediction of life degradation trend of CNN-LSTM

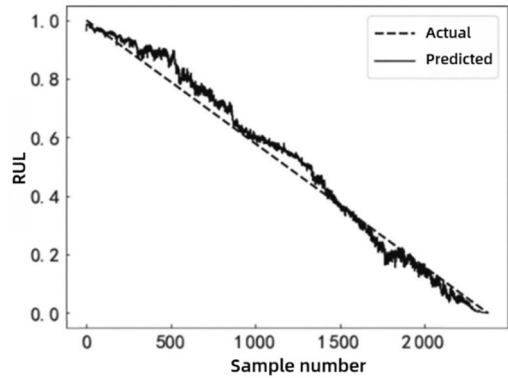


Fig. 8. Prediction of life degradation trend of CNN-LSTM-AM

In order to verify the effectiveness of the model proposed in this paper, bearing 1-3 is taken as an example. Fig. 7 and Fig. 8 show the life prediction trend charts of CNN-LSTM and CNN-LSTM-AM respectively. It can be seen from Fig. 7 that the life prediction trend of CNN-LSTM model can reflect the bearing degradation state, but the prediction accuracy needs to be improved. After adding the self-attention mechanism, we can see that the predicted value is closer to the actual value in Fig. 8, which verifies that the self-attention mechanism can improve the performance of the model and improve the prediction accuracy.

In order to explain the advantages and disadvantages of the proposed model more effectively,

the prediction results of this model are compared with those of the other three deep learning methods LSTM-AM, CNN-LSTM and CNN-BiLSTM, and the performance of the model is evaluated by RMSE and Score. The specific results are shown in Table 2.

It can be seen from Table 2 that compared with other methods, the prediction Score value of the proposed method is the lowest and the RMSE score is the highest. The comparison between CNN-LSTM- AM and CNN-LSTM algorithm shows that under the premise of increasing attention mechanism, the RMSE value of the proposed method is 14.6 % lower than that of CNN-LSTM algorithm, the Score is increased by 2.0 %, and the computing time of the model is 3.1 % faster. By assigning different attention weights to different stages of bearing degradation, the attention mechanism makes the model pay more attention to the important sequence information, which can improve the accuracy of RUL prediction and improve the computational efficiency.

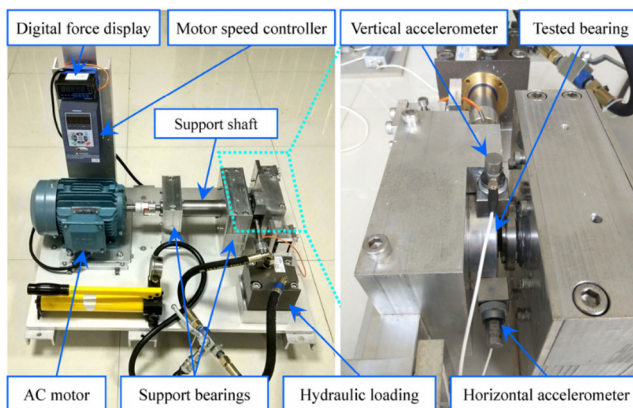
**Table 2.** Comparison of performance indexes of prediction results of each method

Method	LSTM-AM	CNN-LSTM	CNN-BiLSTM	CNN-LSTM-AM
Bearing 1-1	915	957	980	933
Bearing 1-2	812	852	879	819
Bearing 1-3	729	768	794	745
Score	0.923	0.931	0.945	0.950
RMSE	0.176	0.178	0.164	0.152

Comparing the results of CNN-LSTM-AM and LSTM-AM prediction method, we can see that the RMSE value of the proposed method is 12.5 % lower than that of LSTM-AM algorithm, and the Score is 2.9 % higher. The results show that the CNN-LSTM model can mine the deep feature network and make full use of the context information of the sequence data to complete the accurate prediction of bearing RUL; comparing the prediction results of CNN-LSTM-AM and CNN-BiLSTM model, it can be concluded that the RMSE value of the proposed method is reduced by 7.3 %, and the Score is increased by 1.4 %, which fully verifies the reliability and superiority of the method proposed in this paper in bearing residual life prediction.

### 5.1. XJTU-SY Bearing datasets

In order to further illustrate the applicability of the proposed method, a control experiment was carried out by using the data set of XJTU-SY rolling bearing accelerated life test. The bearing accelerated life testing platform used in this test is designed by the joint laboratory and manufactured by Shenyang science and technology. The platform consists of an AC motor, a motor speed controller, a rotating shaft, a supporting bearing, a hydraulic loading system and a test bearing, as shown in Fig. 9.



**Fig. 9.** Testbed of rolling element bearings



The adjustable working conditions of the test platform mainly include radial force and rotational speed, in which the radial force is generated by the hydraulic loading system, which acts on the bearing housing of the test bearing, and the rotational speed is set and adjusted by the speed controller of the AC motor. The test bearing is LDK UER204 rolling bearing, and its related parameters are shown in Table 3. A total of three types of working conditions are designed, as shown in Table 4, there are 5 bearings under each type of working conditions.

In order to obtain the vibration signal of the whole life cycle of the bearing, two unidirectional accelerometers are fixed on the horizontal and vertical direction of the test bearing through the magnetic seat, respectively. In the experiment, DT9837 portable dynamic signal collector is used to collect vibration signals. In the experiment, the sampling frequency is 25.6 kHz, the sampling interval is 1 min, and the sampling time is 1.28 s.

**Table 3.** Parameters of the tested bearings

Parameter	Value	Parameter	Value
Outer race diameter	39.80 mm	Inner race diameter	29.30 mm
Bearing mean diameter	34.55 mm	Ball diameter	7.92 mm
Number of balls	8	Contact angle	0°
Load rating (static)	6.65 kN	Load rating (dynamic)	12.82 kN

**Table 4.** Operating condition

Operating condition	Bearing number	Rotational speed / (r.min <sup>-1</sup> )	Load / kN
Condition 1	1-1, 1-2, 1-3, 1-4, 1-5	2100	12
Condition 2	2-1, 2-2, 2-3, 2-4, 2-5	2250	11
Condition 3	3-1, 3-2, 3-3, 3-4, 3-5	2400	10

Because the load is applied in the horizontal direction, the vibration signal in this direction can contain more degradation information, so the horizontal vibration signal is selected for spectrum and envelope spectrum analysis. In the whole life cycle of bearing operation, the signal processing method is used to extract features from time domain and frequency domain respectively to reflect the degradation process from normal state to serious fault. Lei Yaguo, et al. [26] study found that with the passage of running time, the characteristic amplitude shows an increasing trend, which better reflects the degradation process of the bearing. Similarly, 14 time domain features and 5 frequency domain features are selected as the input of the life prediction model.

In order to evaluate the prediction effect of the proposed method, the prediction results of this method are compared with those of LSTM-AM, CNN-LSTM and CNN-BiLSTM algorithms. RMSE and Score are used to evaluate the performance of the model. The specific results are shown in Table 5.

**Table 5.** Comparison of performance indexes of prediction results of each method

Method	LSTM-AM	CNN-LSTM	CNN-BiLSTM	CNN-LSTM-AM
Bearing 1-1	185	187	174	168
Bearing 2-1	319	312	297	285
Bearing 3-1	965	958	934	923
Score	0.897	0.916	0.925	0.939
RMSE	0.194	0.188	0.176	0.162

It can be seen from Table 5 that compared with the other three methods, the prediction Score value of the proposed method is the lowest and the RMSE score is the highest. The comparison between CNN-LSTM-AM and CNN-LSTM algorithm shows that under the premise of increasing attention mechanism, the RMSE value of the proposed method is 13.8 % lower than that of CNN-LSTM algorithm, the score is 2.4 % higher, and the computing time of the model is 5.6 % faster. By assigning different attention weights to different stages of bearing degradation, the attention mechanism makes the model pay more attention to the important sequence information, which can improve the accuracy of RUL prediction and improve the computational efficiency.

Comparing the results of CNN-LSTM-AM and LSTM-AM prediction method, we can see that the RMSE value of the proposed method is 16.5 % lower than that of LSTM-AM algorithm, and the score is 4.7 % higher. The results prove that CNN-LSTM model can mine deep features and make full use of the context information of sequence data to complete the accurate prediction of bearing RUL. Comparing the prediction results of CNN-LSTM-AM and CNN-BiLSTM model, it can be concluded that the RMSE value of the proposed method is reduced by 7.9 %, and the score is increased by 1.5 %, which verifies the effectiveness of the proposed method in rolling bearing life prediction.

## 6. Conclusions

In this paper, a rolling bearing RUL prediction method based on deep learning is proposed, and the effectiveness of the method is verified by PHM2012 bearing degradation data set. The following conclusions are drawn: 1) The local internal information is extracted by CNN, the deep features are mined, and the deep features are input into the LSTM network, which combines the advantages of LSTM network with autonomous learning ability to time information series. The established CNN-LSTM model can effectively deal with time series information, which avoids the disadvantages of traditional algorithms that require a lot of experience of experts. 2) The attention mechanism is used to assign different weights to the feature information extracted by CNN-LSTM, highlight the key feature information and filter the useless information to represent the degraded feature information of bearing more accurately, so as to effectively improve the accuracy of RUL prediction model. Two sets of open data sets are used to verify the method, and the results show that the method has good adaptability. 3) The bearing degradation data sets used in this paper are all constant speed data. However, there are a large number of variable speed and load conditions in industrial production, and rolling bearing life prediction is a difficult problem in rolling bearing life prediction under variable working conditions. The next step will be to study the residual life prediction of rolling bearings under the complex background of variable working conditions.

The prediction method based on CNN-LSTM-AM model provides a strategy for extracting bearing deep fault features and provides a certain reference for rolling bearing life prediction. The applicability of the proposed method to other similar faults of rolling bearings still needs to be deeply studied on a large number of actual bearing vibration signals. In order to further reduce the prediction error, it will also be studied from the point of view of transfer learning to further improve the accuracy of rolling bearing RUL prediction.

## Acknowledgements

National Natural Science Foundation of China (61773225) and sponsored by Zhejiang Natural Science Foundation (LY20A010012). (1) IMS bearing datasets are generated by the NSF I/UCR Center for Intelligent Maintenance Systems (IMS – [www.imscenter.net](http://www.imscenter.net)) with support from Rexnord Corp. in Milwaukee, WI. (2) XJTU-SY bearing datasets are provided by the Institute of Design Science and Basic Component at Xi'an Jiaotong University, and the Changxing Sumyoung Technology Co., Ltd. The download link is <http://biaowang.tech/xjtu-sy-bearing-datasets>.

## Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

## Author contributions

Wanqing Huang: conceptualization, data curation, formal analysis, investigation, validation and writing-original draft preparation. Yang Chen: conceptualization, funding acquisition,

methodology, supervision and writing-review and editing. Yongqi Chen: conceptualization, methodology, project administration, resources and software. Tao Zhang: investigation, project administration, resources and visualization.

## Conflict of interest

The authors declare that they have no conflict of interest.

## References

- [1] W. Huang, Y. Chen, Y. Chen, T. Zhang, F. Yu, and X. Mao, "A bearing fault diagnosis method based on VMD-HPE," in *Biologically Inspired Cognitive Architectures 2023*, pp. 390–405, Feb. 2024, [https://doi.org/10.1007/978-3-031-50381-8\\_40](https://doi.org/10.1007/978-3-031-50381-8_40)
- [2] M. M. Manjurul Islam and J.-M. Kim, "Reliable multiple combined fault diagnosis of bearings using heterogeneous feature models and multiclass support vector Machines," *Reliability Engineering and System Safety*, Vol. 184, pp. 55–66, Apr. 2019, <https://doi.org/10.1016/j.res.2018.02.012>
- [3] T. Han, L. Zhang, Z. Yin, and A. C. C. Tan, "Rolling bearing fault diagnosis with combined convolutional neural networks and support vector machine," *Measurement*, Vol. 177, p. 109022, Jun. 2021, <https://doi.org/10.1016/j.measurement.2021.109022>
- [4] L. Zhang and R. Wu, "Performance degradation assessment of rolling bearings based on the entropy energy ratio of adaptive frequency band," *Journal of Vibration and Shock*, Vol. 40, No. 2, pp. 63–71, 2021.
- [5] S. Xiang, Y. Qin, C. Zhu, Y. Wang, and H. Chen, "LSTM networks based on attention ordered neurons for gear remaining life prediction," *ISA Transactions*, Vol. 106, pp. 343–354, Nov. 2020, <https://doi.org/10.1016/j.isatra.2020.06.023>
- [6] X. Zhai, X. Wei, and J. Yang, "A comparative study on the data-driven based prognostic approaches for RUL of rolling bearings," in *2019 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 1751–1755, Dec. 2019, <https://doi.org/10.1109/ssci44817.2019.9002764>
- [7] Y. Qin, S. Xiang, Y. Chai, and H. Chen, "Macroscopic-microscopic attention in LSTM networks based on fusion features for gear remaining life prediction," *IEEE Transactions on Industrial Electronics*, Vol. 67, No. 12, pp. 10865–10875, Dec. 2020, <https://doi.org/10.1109/tie.2019.2959492>
- [8] S. Sahoo, S. P. Parida, and P. C. Jena, "Dynamic response of a laminated hybrid composite cantilever beam with multiple cracks and moving mass," *Structural Engineering and Mechanics*, Vol. 87, No. 6, pp. 529–540, Sep. 2023, <https://doi.org/10.12989/sem.2023.87.6.529>
- [9] S. P. Parida, P. C. Jena, and R. R. Dash, "Dynamics of rectangular laminated composite plates with selective layer-wise fillering rested on elastic foundation using higher-order layer-wise theory," *Journal of Vibration and Control*, Vol. 29, No. 23-24, pp. 5598–5615, Nov. 2022, <https://doi.org/10.1177/10775463221138353>
- [10] S. P. Parida and P. C. Jena, "Free and forced vibration analysis of flyash/graphene filled laminated composite plates using higher order shear deformation theory," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, Vol. 236, No. 9, pp. 4648–4659, Oct. 2021, <https://doi.org/10.1177/09544062211053181>
- [11] L. Guo, Y. Lei, N. Li, T. Yan, and N. Li, "Machinery health indicator construction based on convolutional neural networks considering trend burr," *Neurocomputing*, Vol. 292, pp. 142–150, May 2018, <https://doi.org/10.1016/j.neucom.2018.02.083>
- [12] Y. Wang, C. Jiang, and L. Su, "Fault diagnosis method of rolling bearing based on variational mode decomposition algorithm of parameter optimization and support vector machine," in *Lecture Notes in Electrical Engineering*, Singapore: Springer Nature Singapore, 2022, pp. 760–776, [https://doi.org/10.1007/978-981-19-6226-4\\_73](https://doi.org/10.1007/978-981-19-6226-4_73)
- [13] S. G. Kumbhar, R. G. Desavale, and N. V. Dharwadkar, "Fault size diagnosis of rolling element bearing using artificial neural network and dimension theory," *Neural Computing and Applications*, Vol. 33, No. 23, pp. 16079–16093, Jun. 2021, <https://doi.org/10.1007/s00521-021-06228-8>
- [14] H. S. Kumar and S. H. Manjunath, "Use of empirical mode decomposition and K – nearest neighbour classifier for rolling element bearing fault diagnosis," *Materials Today: Proceedings*, Vol. 52, No. 3, pp. 796–801, Jan. 2022, <https://doi.org/10.1016/j.matpr.2021.10.152>

- [15] J. Li, X. Zhang, X. Zhou, and L. Lu, "Reliability assessment of wind turbine bearing based on the degradation-Hidden-Markov model," *Renewable Energy*, Vol. 132, pp. 1076–1087, Mar. 2019, <https://doi.org/10.1016/j.renene.2018.08.048>
- [16] B. Wang, Y. Lei, T. Yan, N. Li, and L. Guo, "Recurrent convolutional neural network: a new framework for remaining useful life prediction of machinery," *Neurocomputing*, Vol. 379, pp. 117–129, Feb. 2020, <https://doi.org/10.1016/j.neucom.2019.10.064>
- [17] J. Liu, Q. Li, W. Chen, Y. Yan, Y. Qiu, and T. Cao, "Remaining useful life prediction of PEMFC based on long short-term memory recurrent neural networks," *International Journal of Hydrogen Energy*, Vol. 44, No. 11, pp. 5470–5480, Feb. 2019, <https://doi.org/10.1016/j.ijhydene.2018.10.042>
- [18] Q. An, Z. Tao, X. Xu, M. El Mansori, and M. Chen, "A data-driven model for milling tool remaining useful life prediction with convolutional and stacked LSTM network," *Measurement*, Vol. 154, p. 107461, Mar. 2020, <https://doi.org/10.1016/j.measurement.2019.107461>
- [19] L. Ren, X. Cheng, X. Wang, J. Cui, and L. Zhang, "Multi-scale dense gate recurrent unit networks for bearing remaining useful life prediction," *Future Generation Computer Systems*, Vol. 94, pp. 601–609, May 2019, <https://doi.org/10.1016/j.future.2018.12.009>
- [20] W. Yu, I. Y. Kim, and C. Mechefske, "An improved similarity-based prognostic algorithm for RUL estimation using an RNN autoencoder scheme," *Reliability Engineering and System Safety*, Vol. 199, p. 106926, Jul. 2020, <https://doi.org/10.1016/j.ress.2020.106926>
- [21] K. Zhao, J. Zhang, S. Chen, P. Wen, W. Ping, and S. Zhao, "Remaining useful life prediction method based on convolutional neural network and long short-term memory neural network," in *2023 Prognostics and Health Management Conference (PHM)*, pp. 336–343, May 2023, <https://doi.org/10.1109/phm58589.2023.00068>
- [22] P. Yao et al., "Intelligent rolling bearing imbalanced fault diagnosis based on Mel-frequency cepstrum coefficient and convolutional neural networks," *Measurement*, Vol. 205, p. 112143, Dec. 2022, <https://doi.org/10.1016/j.measurement.2022.112143>
- [23] D. Bahdanau, K. Cho, and Y. Beengio, "Neural machine translation by jointly learning to align and translate," in *International Conference on Learning Representation*, 2015.
- [24] R. G. K. M. E. R. M. Z. P. Nectoux and C. Varnier, "PRONOSTIA: an experimental platform for bearings accelerated degradation tests," in *IEEE International Conference on Prognostics and Health Management*, 2012.
- [25] R. K. Singleton, E. G. Strangas, and S. Aviyente, "Extended Kalman filtering for remaining-useful-life estimation of bearings," *IEEE Transactions on Industrial Electronics*, Vol. 62, No. 3, pp. 1781–1790, Mar. 2015, <https://doi.org/10.1109/tie.2014.2336616>
- [26] B. Wang, Y. Lei, N. Li, and N. Li, "A hybrid prognostics approach for estimating remaining useful life of rolling element bearings," *IEEE Transactions on Reliability*, Vol. 69, No. 1, pp. 401–412, Mar. 2020, <https://doi.org/10.1109/tr.2018.2882682>



**Wanqing Huang** studies in Ningbo University for Master's degree. His research interests include bearing fault diagnosis and life prediction.



**Yang Chen** received her Ph.D. degree in Zhejiang University, Hangzhou, China, in 2011. Now she works in Ningbo University. Her research interests include fault diagnosis, control algorithm.



**Yongqi Chen** received his Ph.D. degree in Tongji university, Shanghai, China, in 2012. Now he is a professor in Ningbo University. His research interests include fault diagnosis, intelligent algorithm, statistical learning.



**Tao Zhang** studies in Ningbo University for Master's degree. His research interests include fault diagnosis and statistical learning.