# CNN-LSTM-Based Model for Predicting the Remaining Useful Life of Rolling Bearings

Xiaopeng Yu[1], Hao Zhang[1], Fukai Zhao[1], Dong Zhen[1(✉)], Kiuhua Lu[2], and Wei Hu[3]

[1] Tianjin Key Laboratory of Power Transmission and Safety Technology for New Energy Vehicles, School of Mechanical Engineering, Hebei University of Technology, Tianjin 300131, China
`d.zhen@hebut.edu.cn`
[2] WorldTech Intelligence Technology (Tianjin) Com., Ltd., Tianjin 300131, China
[3] World Transmission Technology (Tianjin) Co., Ltd., Tianjin 300131, China

**Abstract.** To achieve accurate prediction of the remaining useful life of rolling bearings, a prediction method combining convolutional neural network (CNN) and long and short-term memory network (LSTM) is proposed. Firstly, the feature parameters in the time domain frequency domain are extracted, and the degradation trend is determined to be characterized using segmentation function, then the feature parameters are normalized as the input of CNN, and the information is extracted using CNN, and then these deep-level features are input to LSTM These deep-level features are then fed into the LSTM for prediction using the predict function to achieve the goal of rolling bearing life prediction. To verify the rationality of the proposed method, the PHM2012 dataset was used and its whole life cycle vibration data was substituted into the proposed method. The experimental results show that the proposed method has a good fitting effect, and the prediction results are close to the real-life value.

**Keywords:** Rolling bearings · Life prediction · Deep learning · Convolutional neural networks · Long and short-term memory networks

## 1 Introduction

At present, rolling bearings are widely used in many rotating machinery and equipment, playing an important role in ensuring the safe and reliable operation of equipment, and once a failure occurs, it will lead to a series of negative effects, such as prolonging downtime and causing malicious accidents. Therefore, accurate prediction of the remaining service life of bearings is of great significance for preventive maintenance decisions on rotating machinery. Existing failure prediction and health management methods can be divided into three main categories: physical model-based methods, data-driven methods, and a mixture of the two [1]. Of these, data-driven methods, which model degradation characteristics based on historical sensor data, have a wide range of applications, while deep learning, as a type of data-driven method, has been used in various fields.

The traditional statistically driven approach is significantly influenced by the choice of model and is susceptible to engineering experience and subjective factors. Also,

statistical analysis models are less generalizable for different research objectives and require repeated modelling. In contrast, machine learning has powerful data processing capabilities and does not require exact physical models and expert prior knowledge, thus deep learning shows promise in the field of remaining life prediction.

In terms of deep learning, Lecun [2] first proposed convolutional neural networks. Later, Li et al. [3] used deep convolutional neural networks for RUL prediction of aero engines and used Dropout technique to prevent overfitting and obtained good prediction performance.

For the life prediction of bearings, the vibration signal of bearings contains a lot of timing information, so the CNN which is insensitive to timing information does not fully exploit its ability. The most suitable one for processing timing information in deep learning is recurrent neural network [4], and WANG et al. [5] used recurrent convolutional neural network for predicting the remaining service life of machines, breaking the limitation of CNN. However, RNNs suffer from gradient disappearance when the computational load is high [6], and LSTM, as an improvement of RNN, can solve this phenomenon well. Wang et al. [7] predict the remaining life of a bearing by using manually extracted time-domain, frequency-domain, and time-frequency-domain features as the input to an LSTM network.

## 2 Theoretical Foundations

### 2.1 Convolutional Neural Networks

Convolutional neural network is a feed-forward neural network containing convolutional operations, which excels in feature extraction and is widely used in speech recognition, image classification, machinery fault diagnosis and other fields.

Figure 1 shows a typical CNN is made up of a stack of multiple convolutional and pooling layers, consisting of five parts: an input layer, a convolutional layer, a pooling layer, a fully connected layer and an output layer. The convolutional and pooling layers are usually used in pairs for convolution and dimensionality reduction of the input feature
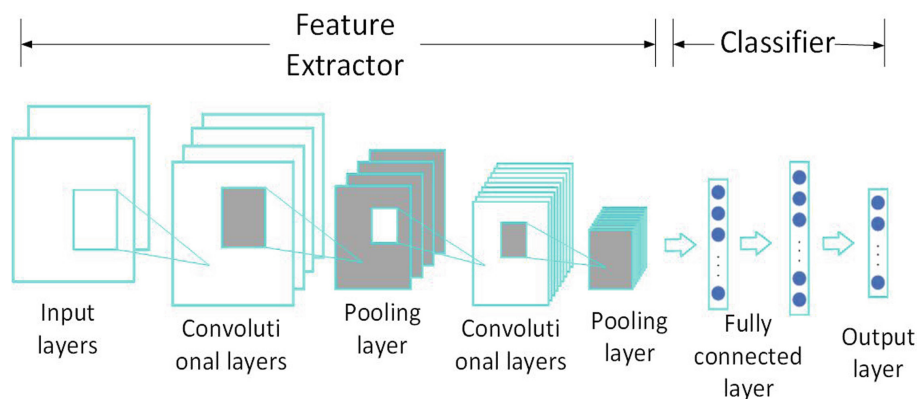


**Fig. 1.** CNN network structure

information, while the fully connected and output layers are used for the output of the model training results.

The convolutional neural network operation is a convolutional kernel traversing the entire input sequence data to produce a higher level, more abstract feature space; secondly, the pooling layer compresses each generated feature for secondary feature extraction, dimensionality reduction and selection of higher-level important features; finally, new sequence features are generated as input to the next convolutional layer and pooling layer. The specific operations of the convolution and pooling layers are as follows.

1) Convolutional layer: A portion of the data is selected for calculation by sliding the size of the convolutional kernel window, and the result of the convolution is the feature map. Usually, a convolutional layer has multiple convolutional kernels, resulting in multiple feature maps, and the weights of the same convolutional kernel are shared. This feature reduces the number of network connections, reduces model complexity and lowers system memory expenditure. Assuming that the input to the CNN model is X, Eq. (1) defines the formula for calculating the output of the convolution layer.

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

where: $C_n$ is the nth feature map output from the convolution layer; $\sigma$ is the activation function; $W_n$ is the weight matrix of the nth convolution kernel of the current convolution layer; $b_n$ is the bias of the nth convolution kernel of the current convolution layer; and $\otimes$ is the convolution operation.

2) Pooling layer: The pooling layer is mainly used to carry out the down sampling operation to achieve the purpose of reducing the network parameters and speeding up the calculation speed. There are generally four types of pooling functions, namely maximum pooling, average pooling, global average pooling and global maximum pooling. Equation (2) defines the pooling function when the maximum pooling operation is used.

$$P_n = max C_n \tag{2}$$

where: $P_n$ is the output of the pooling layer; $C_n$ is the input of the pooling layer.

## 2.2  Long Short-Term Memory Neural Networks

During a convolutional neural network operation, the state always propagates from front to back, which means that in a CNN network, information flows only in one direction. At each computational step, the CNN only considers the current input and ignores previous degradation information. Consequently, convolutional neural networks are good at extracting data features but are unable to model the backward and forward correlation of different machine degradation states. The RNN model can retain the model's memory of input patterns, which can effectively compensate for the shortcomings of CNNs. LSTM is an improved RNN that introduces memory units, effectively overcomes the gradient

disappearance problem, and solves the problem of long-term dependence that RNNs cannot learn.

The hidden layer structure of the LSTM network is the long- and short-term memory block, which consists of three control gates and a cellular structure, Fig. 2 shows the specific structure of the LSTM network. $f_t$, $i_t$, $o_t$ are the forgetting gates, input gates and output gates respectively. The LSTM controls the flow of information in the time series through the action of these three gates, thereby better capturing the long-term dependence problem in the sequence. This allows for better capture of long-term dependency problems in the sequence and efficient processing of sequence data.
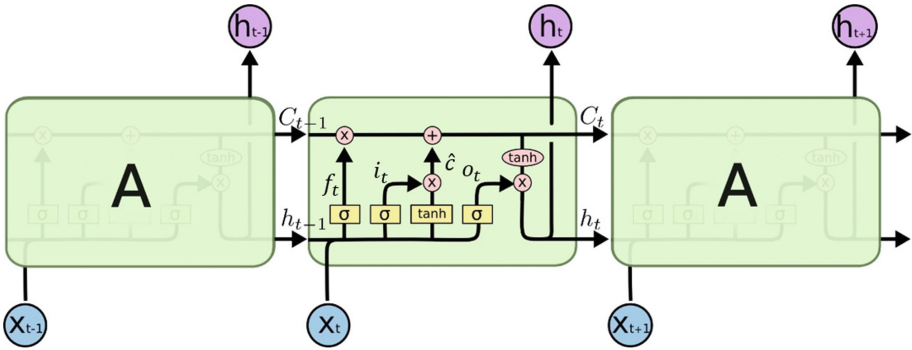


**Fig. 2.** LSTM network structure

Equations (3), (4), (5), (6), (7) and (8) define the update steps of the LSTM.

1) Calculate the value of the forgetting gate $f_t$. The LSTM controls the memory cell state by calculating the forgetting gate $f_t$ from the input vector $[h_{t-1}, X_t]$ consisting of the previous moment's output $h_{t-1}$ and the current moment's input $x_t$ together

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \tag{3}$$

2) Calculate the value of the input gate $i_t$. The input gate determines which values are used to update through the sigmoid function, which controls the effect of the current data input on the state value of the memory cell

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \tag{4}$$

3) Calculate the candidate state $\hat{c}_t$, which will be generated by the tanh activation function before generating the new candidate state information value $\hat{c}_t$. And $\hat{c}_t$ is the result of the joint action in two-time steps

$$\hat{c}_t = \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \tag{5}$$

4) Calculate the value $c_t$ of the memory cell at the current moment

$$c_t = f_t \otimes c_{t-1} + i_t \otimes \hat{c}_t \tag{6}$$

5) Calculate the value of the output gate $o_t$. The main function of the output gate is to control the message output

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \tag{7}$$

6) LSTM memory cell output

$$h_t = o_t \tanh(c_t) \tag{8}$$

where: $h_{t-1}$ is the output at the previous moment; $W_{xf}$, $W_{xi}$, $W_{xc}$, $W_{xo}$ are the weights of forgetting gates, input gates, memory units and output gates of the hidden layer at moment t; $W_{hf}$, $W_{hi}$, $W_{hc}$, $W_{ho}$ are the weights of forgetting gates, input gates, memory units and output gates of the hidden layer between moments $t - 1$ and t, respectively $b_c$ is the memory node bias; $b_i$, $b_f$, $b_o$ correspond to the bias vectors of the three multiplication gates respectively; $\sigma$ is the activation function, generally using the sigmoid function, taking values from 0 to 1.

The long and short-term memory neural network is trained by a time-based back-propagation algorithm where errors are back-propagated through the time dimension. The training allows the network to achieve feature extraction of time series data, thus allowing information on the degradation process of the bearing in the time domain signal to be reflected and therefore more accurate time series prediction.

## 3    RUL Prediction Methods

### 3.1    Data Pre-processing

For the CNN-LSTM neural network model, input and output data are required. The original vibration signal of the bearing has too many disturbing features and is not obvious, so it is not suitable to be used as the input of the network model directly. The time domain and frequency domain features of the vibration signal can reflect the degradation state of the bearing. The common time domain features are: peak value, root mean square value, peak-to-peak value, cliffness indicator, etc.; while the frequency domain features are: Centre frequency, average frequency, etc. Under normal operating conditions, the bearing condition monitoring signal is normally distributed, while the cliffness indicator changes when the signal deviates from the normal distribution, and the magnitude of the change represents the degradation of the bearing.

Figure 3 shows the root mean square values and root square magnitudes of bearing1–1 for the data used in this paper.
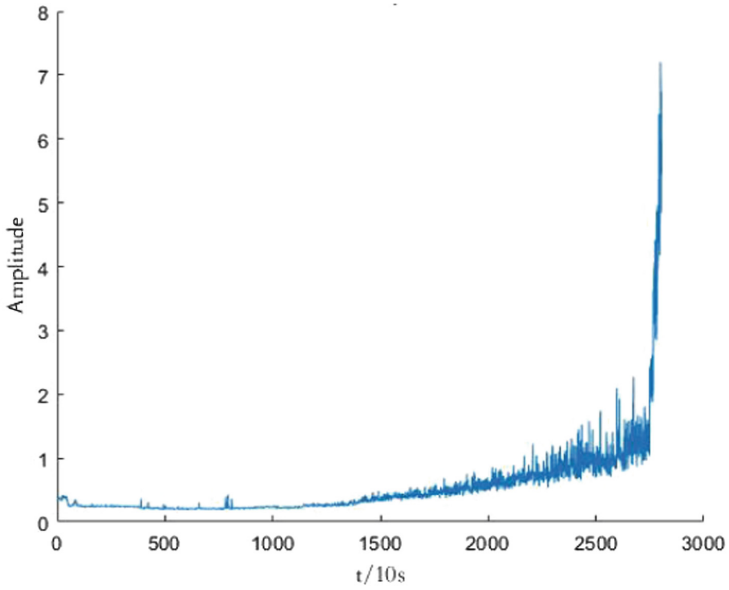
(a)
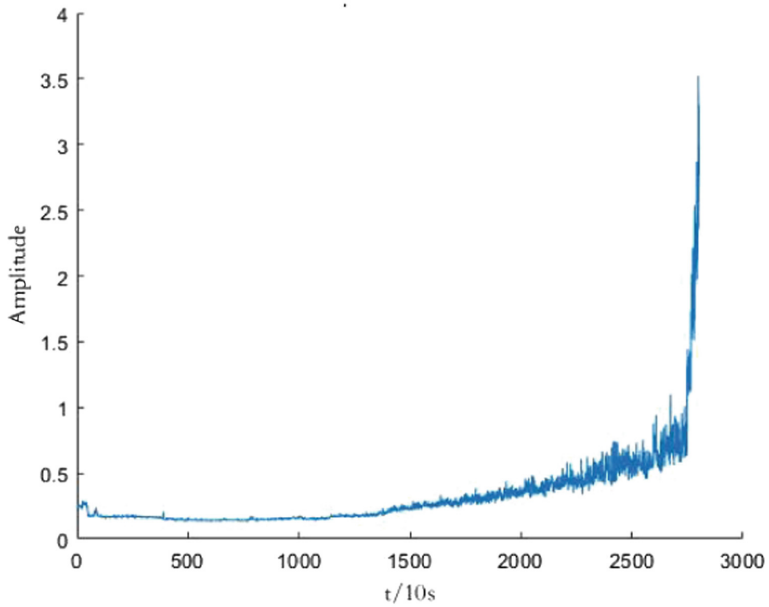


(b)

**Fig. 3.** (a) Bearing1–1 Root mean square values of vibration signals. (b) Bearing1–1 Root square amplitude of vibration signal

It is easy to see from Fig. 3 that as the bearing degrades, its indicators such as root mean square value and root square amplitude also show a clear trend of increasing amplitude, indicating that it is feasible to extract the characteristic information in the time and frequency domain for use as life prediction.

Therefore, in this paper, 12 time-domain frequency-domain features, namely mean value, root mean square value, root square amplitude, cliff value, peak-to-peak value, waveform indicator, peak indicator, pulse indicator, margin indicator, cliff indicator, mean frequency and root mean square frequency, are selected as inputs to the neural network model.

To facilitate the calculation, the remaining bearing life is mapped between 0 and 1. When RUL = 1, the bearing is intact, when RUL = 0, the bearing is damaged and needs to be replaced or repaired immediately. Equation (9) defines the formula for calculating RUL

$$RUL = \frac{X - i}{X} \tag{9}$$

The degradation process of most bearings can be reduced to two processes, the normal operation phase, and the degradation phase [8], so the remaining life of the bearing can be considered as a segmental function, with the point at which the vibration amplitude begins to consistently exceed a threshold value defined as the moment at
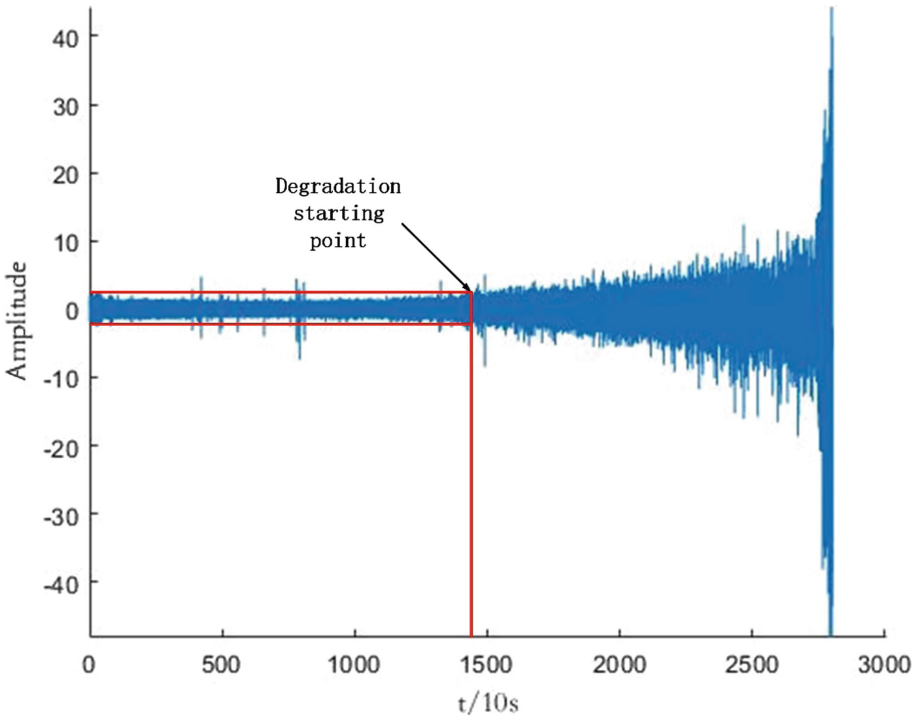


**Fig. 4.** Bearing1-1 Full life vibration signal

which degradation has begun. Figure 4 shows the 1424th sampling point of bearing1–1 is defined as the moment of the start of degradation.

## 3.2 CNN-LSTM Model Building

To give full play to the spatial feature extraction ability of CNN and the temporal feature extraction ability of LSTM, this paper combines the two to propose a CNN-LSTM model to predict the granted service life of bearings, Fig. 5 shows the flow block diagram.
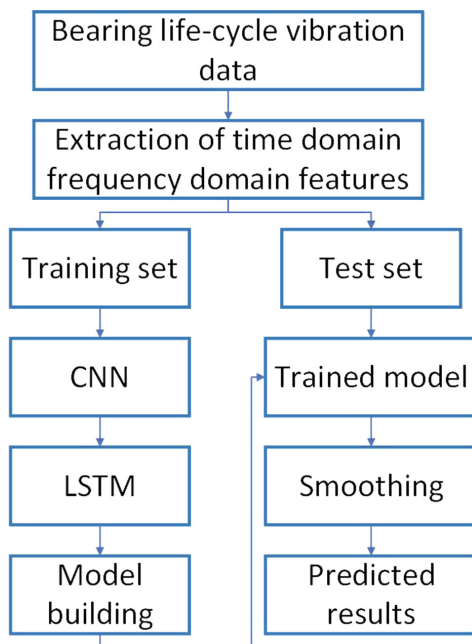


**Fig. 5.** Prediction flow chart

As shown in the figure, the vibration data of the bearing is first pre-processed to extract its time and frequency domain features and divide the training and test sets according to a certain ratio, after which the feature values are normalized so that the values are mapped to between 0 and 1 and used as feature inputs. The training set is then fed into a CNN-LSTM network for training, and the test data is fed into the trained network model for testing, and the prediction results are smoothed to produce the final RUL prediction results for the bearings.

Figure 6 shows the specific network structure used in this paper, Table 1 shows the parameters.
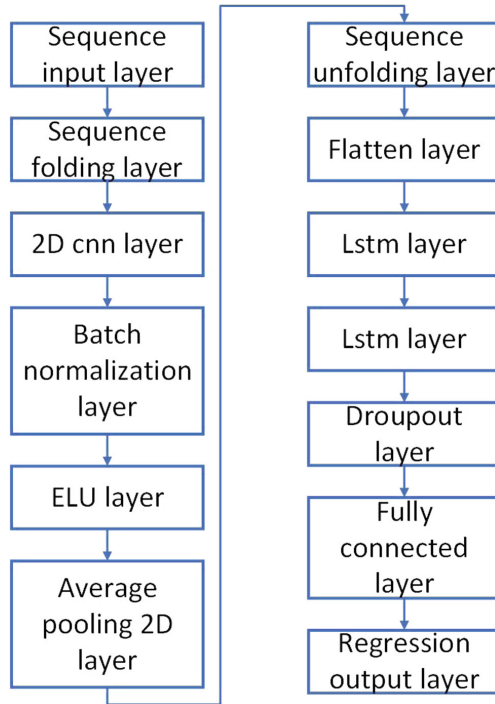
**Fig. 6.** CNN-LSTM network structure

**Table 1.** Network model parameters

| Network parameters | CNN-LSTM |
|---|---|
| Input dimensions | 12 |
| Output dimensions | 1 |
| Number of network layers | 13 |
| Learning rate | 0.005 |
| Time step | 10 |
| Rate of decline in learning rates | 0.45 |

The network model in this paper has a 9:1 split between the training and test sets, uses the mean square error MSE as the cost function, and uses the adam algorithm with a batch size of 200 and a learning rate of 0.005 as the optimizer.

## 3.3  Smoothing of Prediction Results

For better prediction results, the predicted data will be smoothed, because the prediction results obtained using deep learning methods will have large local fluctuations, which

will lead to bias in the prediction results, so this paper uses an exponential method with a window size of 30 to smooth the predicted data.

## 4 Experimental Validations

### 4.1 Description of the Dataset

In this paper, the validity of the CNN-LSTM model for predicting the remaining service life of rolling bearings is verified by using the publicly available bearing full life cycle degradation dataset from the IEEE PHM 2012 Data Challenge.

The dataset was provided by the FEMTO-ST Institute in France using the PRONOS-TIA experimental platform, which uses two acceleration sensors to collect acceleration data in both horizontal and vertical directions at a sampling frequency of 25.6 kHz, with each sampling lasting 10 s and sampling at ten-second intervals, respectively. The bearing is considered damaged when the vibration amplitude exceeds 20 g. From the literature [9] it is known that the horizontal vibration signal contains more degradation information and is more suitable for the RUL prediction of the bearing, so the horizontal vibration signal is chosen for model training and testing in this paper.

### 4.2 Evaluation Indicators

To quantitatively analyses the goodness of the model prediction results, evaluation indicators need to be introduced. In this paper, root mean square error and score function is chosen as evaluation indicators, which are defined as follows.

(1) RMSE value: The root mean square error represents the deviation between the predicted and true values. The weights assigned to the overestimation and underestimation of bearing life are the same, which means that the final evaluation result is the same when the life is overestimated and underestimated, if the difference is the same. Equation (10) defines the specific formula for calculating RMSE

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

where: N is the number of samples, $f(i)$ is the true RUL label of the sample, and $\hat{f}(i)$ is the predicted RUL value.

(2) scoring function: scoring function is provided by the 2012 PHM data challenge, used to evaluate the prediction of good or bad results of the function. Unlike the RMSE, the scoring function has different evaluation scores for the overestimation and underestimation of the results. As the underestimation of the bearing life is more likely to improve the safety of the equipment, the scoring function will give a higher score for the predicted life than the actual life, Eqs. (11), (12) and (13) define the specific calculation process of the score function.

$$E_{ri} = \frac{ActRUL - PreRUL}{ActRUL} \tag{11}$$

$$A_i = \begin{cases} exp^{-\ln(0.5)*\left(\frac{E_{ri}}{5}\right)}, E_{ri} \leq 0 \\ exp^{\ln(0.5)*\left(\frac{E_{ri}}{20}\right)}, E_{ri} > 0 \end{cases} \qquad (12)$$

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

where: $E_{ri}$ is the percentage error of the ith sample, ActRUL is the actual RLU value of the bearing, PreRUL is the predicted RUL value of the bearing, and Score is the final score.

### 4.3  Analysis of Results

To verify the validity of the model, Bearing1–3 is selected as the test set in this paper to analyses its prediction results. Figure 7 shows the time domain diagram of the vibration signals of the bearings in the test set.
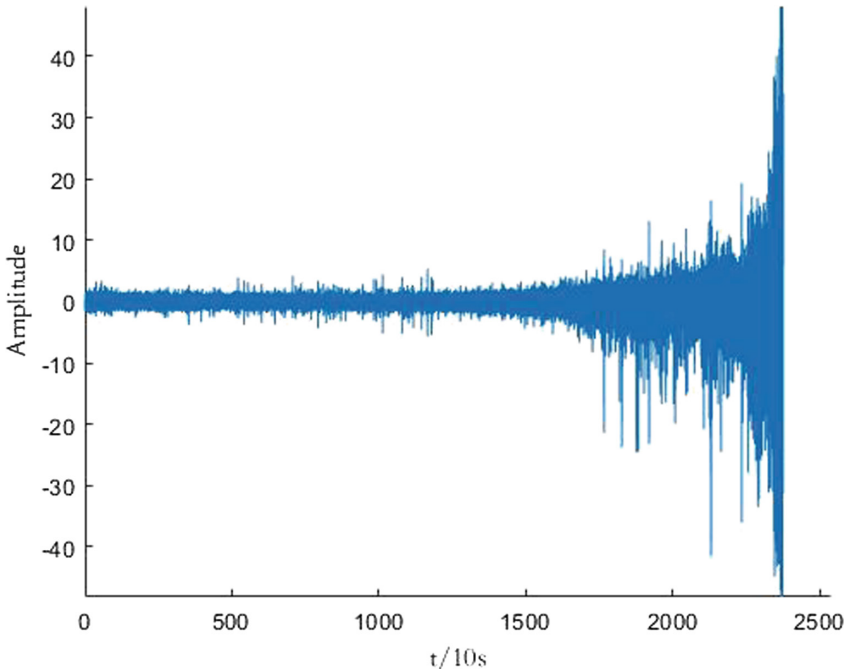


**Fig. 7.**  Bearing1–3 Time domain diagram of the vibration signal

The predictions from the model output were smoothed and Fig. 8 shows the comparison of the processed results with the original data.
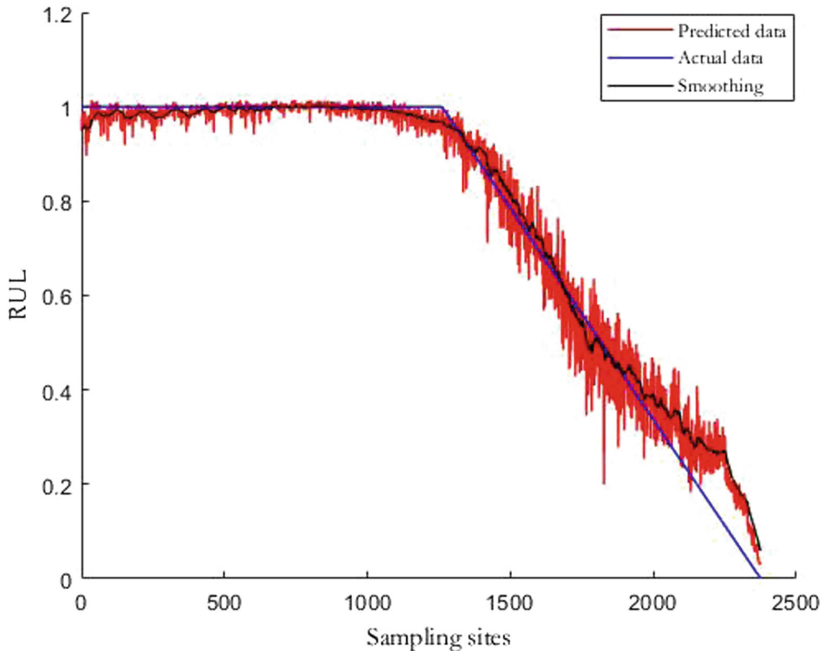
**Fig. 8.** Predicted results

As can be seen from Fig. 8, the prediction results of the CNN-LSTM model accurately match the degradation trend of the bearings, but there are local fluctuations in the prediction, and after exponential smoothing the fluctuations are significantly reduced, which can better match the degradation trend, which proves the feasibility of this smoothing method.

For the rolling bearing RUL prediction problem, it is quite difficult to obtain very accurate RUL prediction results throughout the life cycle of the bearing. The CNN-LSTM network model proposed in this paper can provide an accurate indication of the degradation trend of the bearing and act as an early warning.

To verify the superiority of the method in this paper, the prediction results were compared with the literature [10] using the same dataset, comparing their RMSE values and Score scores, and Table 2 shows the results.

**Table 2.** Comparison of predicted results

| Test bearings | CNN | | GRU | | CNN-GRU | | CNN-LSTM | |
|---|---|---|---|---|---|---|---|---|
| | RMSE | Score | RMSE | Score | RMSE | Score | RMSE | Score |
| Score | 0.092 | 0.964 | 0.153 | 0.965 | 0.082 | 0.974 | 0.044 | 0.975 |

As can be seen from Table 2, the proposed method in this paper has the lowest RMSE and highest Score compared with other methods, so the CNN-LSTM model can

effectively predict the RUL of the bearings and has a smaller prediction error than the other three models, which proves the effectiveness of the model in this paper.

## 5 Conclusion

In response to the problem that the remaining useful life of bearings is difficult to predict, this paper proposes a CNN-LSTM method for predicting the RUL of bearings and verifies the effectiveness of the method with the PHM2012 bearing degradation dataset. The specific contents are.

1) Pre-process the vibration signals, select 12 time-domain frequency-domain signals that have a great influence on the prediction results, and input them into the prediction model after normalization to improve the prediction accuracy.
2) Construct a network structure, build a CNN-LSTM network, and adjust the model parameters to achieve the best prediction results.
3) Select suitable evaluation indexes and compare the prediction results of this model with those of other models to verify the effectiveness and superiority of this model.

## References

1. Man, S.K., Tan, A., Mathew, J.: A review on prognostic techniques for non-stationary and non-linear rotating systems. Mech. Syst. Sig. Process. 62–63(oct.), 1–20 (2015)
2. Lecun, Y., Bottou, L.: Gradient-based learning applied to document recognition. Proc. IEEE **86**(11), 2278–2324 (1998)
3. Li, X., Ding, Q., Sun, J.Q.: Remaining useful life estimation in prognostics using deep convolution neural networks. Reliab. Eng. Syst. Saf. **172**(APR.), 1–11 (2018)
4. Zhang, J., Wang, P., Yan, R., et al.: Long short-term memory for machine remaining life prediction. J. Manuf. Syst. **48**, 78–86 (2018)
5. Wang, B., Lei, Y., Yan, T., et al.: Recurrent convolutional neural network: a new framework for remaining useful life prediction of machinery. Neurocomputing **379**, 117–129 (2020)
6. Hochreiter, S.: Untersuchungen zu dynamischen neuronalen Netzen (1991)
7. Wang, F., Liu, X., Deng, G., Yu, X., Li, H., Han, Q.: Remaining life prediction method for rolling bearing based on the long short-term memory network. Neural Process. Lett. **50**(3), 2437–2454 (2019). https://doi.org/10.1007/s11063-019-10016-w
8. Li, X., Zhang, W., Ding, Q.: Deep learning-based remaining useful life estimation of bearings using multi-scale feature extraction. Reliab. Eng. Syst. Saf. **182**(FEB.), 208–218 (2019)
9. Soualhi, A., Medjaher, K., Zerhouni, N.: Bearing health monitoring based on Hilbert-Huang transform, support vector machine and regression. IEEE Trans. Instrum. Meas. **64**(1), 52–62 (2014)
10. Zhao, G.Q.: A remaining useful life prediction of rolling element bearings based on convolutional neural network and bidirectional GRU. Qingdao University of Technology (2021)