Convolution neural network based particle filtering for remaining useful life prediction of rolling bearing

Xiyang Liu[1] [ID]**, Guo Chen[2], Zhenjie Cheng[3], Xunkai Wei[4] and Hao Wang[4]**[AQ: 6]

Abstract
Aiming at the problem of remaining useful life prediction of rolling bearing in aero engine, a data-driven prediction method based on deep learning and particle filter is proposed. Initially, only the vibration data of rolling bearing in normal stage are trained by the deep convolution neural network. According to the feature distance between normal and degraded samples, the evolution features during the whole lifetime are extracted adaptively, and the health index of rolling bearing is constructed. Then, the alarm and failure threshold are determined by unsupervised clustering algorithm. Combined with the extracted feature, remaining useful life of rolling bearing is tracked and predicted by particle filter algorithm based on four parameter exponential model. Finally, the effectiveness of the proposed method is verified by three groups of whole lifetime test data of rolling bearings. Results show that the degradation feature extracted by deep learning method has higher prediction accuracy of 2.19%, 0.93%, and 1.43% respectively than RMS values, and has more stable performance and less influenced by the number of particles or resampling methods, which can better reflect the evolution trend of rolling bearing than the traditional feature.

Introduction

As an important supporting component, rolling bearing has a significant impact on the service life, safety and reliability of aeroengine. Evaluating the reliability of rolling bearing, perceiving the degradation and accurately predicting its remaining service life can provide evidence for making the preventive maintenance decision,[1] reduce maintenance costs and downtime, and improve the safety guarantee of aeroengine.

The remaining useful life (RUL) prediction of rolling bearing generally includes four stages: data acquisition, construction of health indicators (HIs), division of health stages (HSs), and prediction of remaining life.[2] Among them, the construction of health indicators plays an important role. The selection of reasonable

indicators can precisely reflect the degradation trend of bearings, greatly improve the performance of the RUL prediction model. The quality of feature extraction directly determines the results of prediction.

[1]College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing, China
[2]College of General Aviation and Flight, Nanjing University of Aeronautics and Astronautics, Nanjing, China
[3]School of Engineering, Hong Kong University of Science and Technology, Hong Kong, China
[4]Beijing Aeronautical Technology Research Center, Beijing, China

Corresponding author:
Guo Chen, College of General Aviation and Flight, Nanjing University of Aeronautics and Astronautics, Liyang 213300, China.
Email: idescg@163.com[AQ: 1]

Aiming at the construction of health indicators for the prediction of RUL of rolling bearings, scholars have conducted many related researches. The time-domain and frequency-domain characteristics of bearings, such as root mean square (RMS) or peak,[3] wavelet coefficients,[4] Hilbert spectrum,[5] etc., are the most commonly used characteristic indexes in bearing RUL estimation. In addition, some researches constructed new health indicators according to the statistical characteristics of bearing fault signals. For example, Kumar et al.[6] obtained the degradation trend of bearing through a variety of entropy features, and selected the degradation features based on the mixed metric health indicators; Kong and Yang[7] characterized the signal energy by the mean absolute extreme value of the vibration signal of rolling bearing, and combined with its root mean square (RMS-MAVE) as the degradation characteristic index. Other studies construct health indicators by fusing multiple physical indicators or multiple sensor signals. For example, Zhao et al.[8] constructed high-dimensional features through multi-scale time-frequency representation (TFR) obtained by S-transform and Gaussian pyramid, and then combined with principal component analysis (PCA) and linear discriminant analysis (LDA) for two-step supervised dimensionality reduction. Wang et al.[9] estimated the degradation by calculating the Mahala Nobis distance between multiple statistics of the bearing and the health stage through time-domain feature fusion, and combined with Kalman filter to predict the remaining life of the bearing. However, the health indicators used in the above researches are still traditional manual features based on time domain or frequency domain, which has certain requirements for professional skills and knowledge reserve, and can only reflect some unilateral characteristics of bearing vibration signal. Therefore, they have specific limitations in the intelligent and sufficient extraction of features and the comprehensive reflection of the bearing deterioration trend.

Deep learning method has great advantages in feature adaptive extraction, which has been widely concerned and has become a hot research direction in recent years.[10] The following convolution neural network (CNN) method, with its unique advantages in large-scale data feature extraction, has been taken general attention and gains meaningful achievements. The fundamental thought of CNN is to enable the network acquiring the ability of learning spatial hierarchical structure of patterns and has translation invariance through convolution and pooling operations.[11] The deep CNN network (DCNN) developed on the basis of CNN has more powerful learning ability than the shallow network, which is widely applied into pattern recognition, fault diagnosis, and other fields.

Despite its popularity and numerous applications in RUL estimation of bearing, deep learning has still certain limitation on the number of samples. A large number of degradation and failure samples is needed in the training of deep model, which is difficult to obtain in practice. Moreover, deep learning method is highly dependent on historical data, and its stability remains further discussion. The vibration signal of rolling bearing has strong non-stationarity,[12,13] which poses a certain challenge to the performance of deep model.

Bayesian filtering is proposed to resolve the limitation of prediction stability, and the particle filter (PF) algorithm developed on this basis has been widely used.[14] The PF method can dynamically adjust the model parameters under non-stationary conditions by using the optimal Bayesian estimation and Monte Carlo algorithm, realize multi-step advance prediction with high stability, which is quite appropriate for the prediction of nonlinear and non-Gaussian signals.

In view of the present situation, the deep learning method is combined with particle filter in this paper to estimate the remaining useful life of rolling bearing. By using the strong feature extraction ability of deep learning, a deep convolution neural network (DCNN) is developed to adaptively extract the health indicators of rolling bearings, which retains the essential characteristics of bearing degradation. Combined with the unique advantages of particle filter algorithm in long-term prediction of nonlinear non-Gaussian signals, the tracking prediction of the remaining life of rolling bearings is realized. The whole lifetime vibration data based on bearing fatigue testing machine is acquired to verify the effectiveness of the proposed model.

This innovative of this paper is mainly reflected in the following aspects:

1. Only samples on normal stage are trained to develop the model, and the characteristic distance between normal and degradation samples is extracted by the model and described as the bearing damage evolution stage, which is more in line with the reality of the scarcity of fault samples.
2. Aiming at the problem that the traditional time or frequency characteristics can hardly reflect the essential characteristics of bearing degradation comprehensively and intelligently, an efficient deep learning model is developed to construct the degradation index adaptively, and the particle filter algorithm is introduced to solve the problem of huge distribution difference between train and test datasets and improve prediction stability.
3. The degeneration and failure thresholds of bearing are often vaguely defined by human observation.[15] In this paper, the damage evolution stage is adaptively divided by unsupervised learning, which is more effective and convenient.
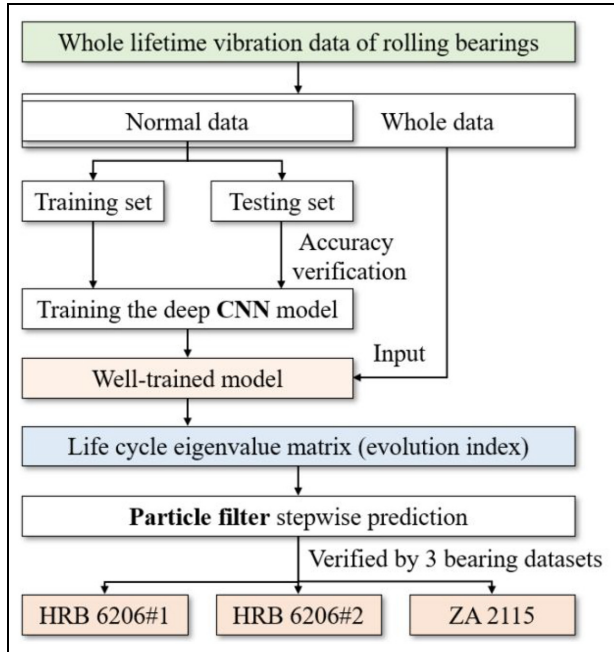
Figure 1. Prediction process of residual life of rolling bearing based on deep learning and particle filter.

The residual life prediction process of rolling bearing based on deep learning and particle filter proposed in this paper is shown in Figure 1.

Adaptive extraction of new deterioration index of rolling bearing based on vibration monitoring data

Vibration monitoring data of whole lifetime fatigue test of rolling bearing

The experimental method is adopted in this paper to obtain the real whole lifetime fatigue vibration monitoring data of rolling bearing. The test equipment includes bearing life enhancement test machine (ABLT-1A, Hangzhou Bearing Test and Research Center Co., Ltd, China), acceleration sensor (AI002, Yangzhou Jingming Technology Co., Ltd, China), dynamic signal test and analysis system (JM5937, Yangzhou Jingming Technology Co., Ltd, China), etc. The bearing life enhancement test machine is mainly used for rolling bearing fatigue life accelerating test, the external structure, and sensor placement of which are shown in Figure 2.

The bearings used in the test are deep groove ball bearings (HRB 6206, Harbin Bearing Manufacturing Co., Ltd, China). The appearance and corresponding parameters of the bearing are shown in Figure 3(a) and Table 1. The bearings were used for accelerated fatigue test, which means the test was accelerated by increasing

radial load (6.25 kN of each bearing) and rotating speed to shorten the bearing life. The test was carried out twice, and finally two failed rolling bearings were obtained. As shown in Figure 3(b) and (c), the inner race fault occurred in #1 bearing, and the mixed fault of inner race and ball occurred in #2 bearing. Relevant test parameters are shown in Table 2.

After twice tests of the same type bearing, 300 and 650 data files are obtained respectively. Each file contains the vibration signal of the bearing at a certain time step, and all files contain the vary trend of the bearing whole life stage from its normal operation to the complete failure. Figure 4 shows the comparison of vibration signals of #1 bearing in different degradation stages. It can be seen from the figure that the amplitude of bearing acceleration increases continuously from normal to the final failure stage, which has obvious evolution progress of characteristics.

Extraction of rolling bearing deterioration index based on deep convolution neural network

Deep convolution neural network model. Convolution neural network is a kind of multi-layer perceptron network that uses the convolution of weight sharing, which is capable of digging more information by increasing the depth of the network.[16] As one of the most commonly used methods in deep learning, compared with other neural networks, the greatest advantage of deep convolution neural network is that it can automatically learn more abstract features from data, and calculate the features for classification or regression. The parameters that need to be updated in the process of CNN network propagation include input value x_i, weight value ω_i, bias b, and convolution kernel, etc. Activation functions such as ReLU, sigmoid, and tanh are used to introduce nonlinearity into neural networks. The convolution operation process is:

$$y_j = f\left(\sum_{i=1}^{n} \omega_{ij} x_i + b_j\right) \tag{1}$$

Where x_i is the input signal, y_j is the output signal, ω_{ij} is the weight of the neuron j connected to the input signal x_i, and b_j is bias. $f(\cdot)$ is the activation function, which used in this paper is ReLU:

$$f(y_j) = max(0, y_j) \tag{2}$$

The internal transmission structure of neural network unit is shown in Figure 5.

$\omega_i \, X f_i = \omega_i * X + b_i * f_i \, \omega_i \, p \omega_i \, p \, i^{th} \, x_i \, y_i = \gamma \frac{x_i - \mu}{\sqrt{\sigma^2 + \varepsilon}} + \beta$ $\mu \, x_i \, \sigma^2 \, \varepsilon \, \beta \, \gamma \, \omega^k \, b^k \, x^{k-1} \, y^k = \omega^k x^{k-1} + b^k$ [AQ: 2]

The construction of each layer of the network has a great impact on the bearing degradation features
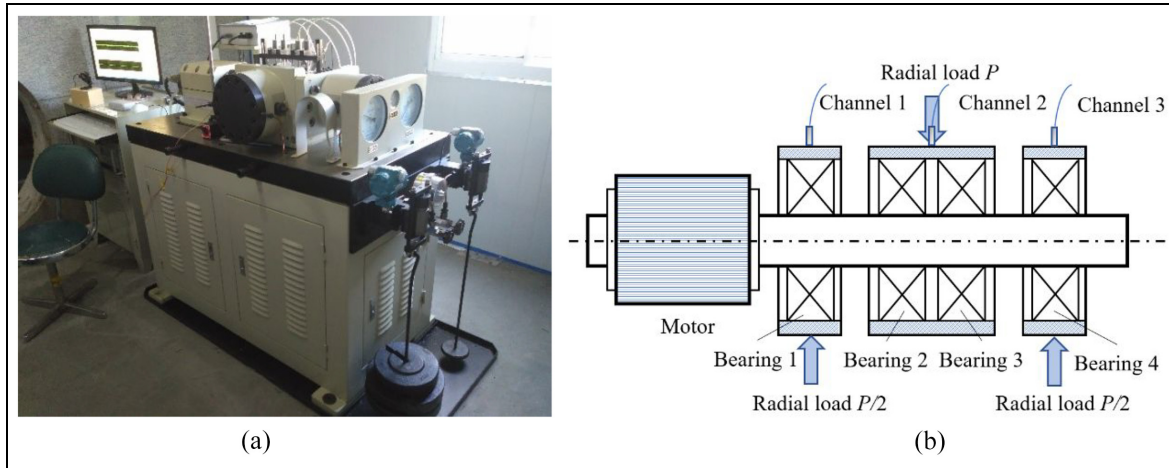
Figure 2. Illustration of bearing life enhancement testing machine: (a) ABLT-1A bearing test machine and (b) sensor placement.

Table 1. Main parameters of HRB 6206 deep groove ball bearing.

Bore diameter	Outer diameter	Width	Diameter of balls	Pitch diameter	Number of balls	Contact angle
30 mm	62 mm	16 mm	9.5 mm	46 mm	9	0°

Table 2. Relevant parameters of the test.

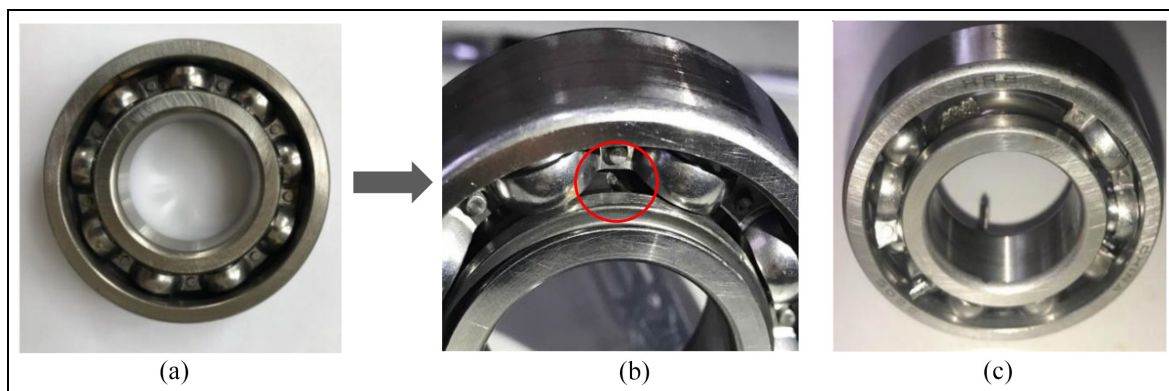| Type of bearing	Test duration	Sampling rate	Sampling time	Rotating speed	Save interval of data	Preset defect	Eventually failure type
HRB 6206#1	30 h	32 kHz	2 s	11,500 r/min	6 min	No	Inner race fault
HRB 6206#2	32.5 h	50 kHz	1 s	11,500 r/min	3 min	No	Mixed fault of inner race and ball

Figure 3. Appearance comparison of HRB 6206 rolling bearing before and after the test: (a) normal, (b) inner race fault, and (c) inner race and ball mixed fault.

extraction. A deeper model means better nonlinear expression ability, which can learn more complex transformations, thus can fit more complex feature inputs. However, network with excessive depth may lead to gradient instability, which degrades the network and reduces the performance of the model. In order to determine the most appropriate network construction, this paper takes a convolution-pooling-activation layer
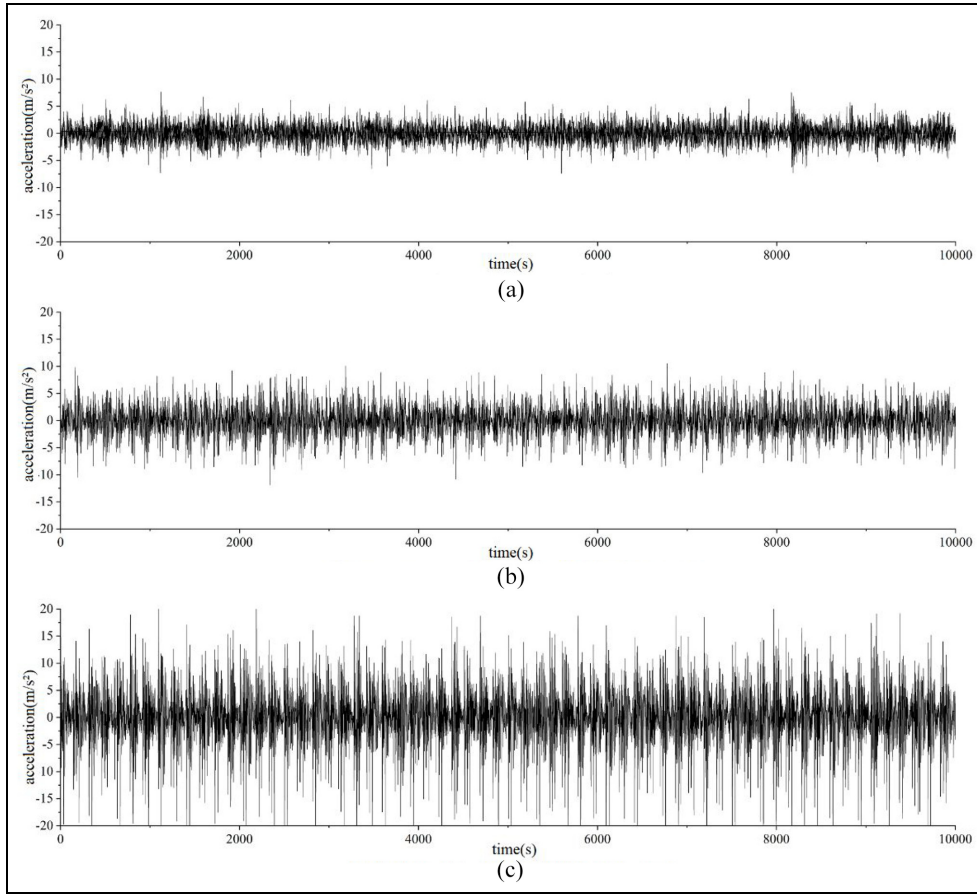
Figure 4. Amplitude of bearing acceleration signals under different states (6206#1): (a) normal state (at the 100th time step), (b) degenerate state (at the 250th time step), and (c) failure state (at the 290th time step).
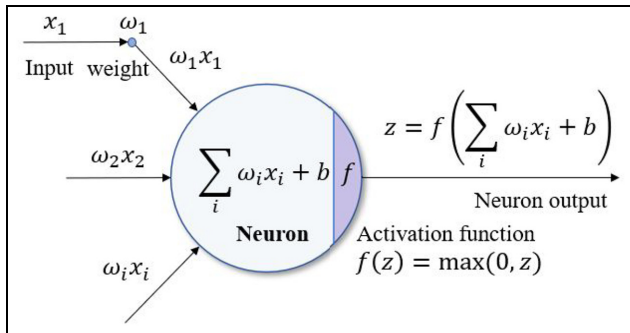

Figure 5. Unit structure of neural network.

Figure 6. Effect of convolution-pooling-activation layers group number on network performance.

as the basic nonlinear transformation module, by gradually deepening the module, training time, and the loss of the network in the verification set are recorded, according to which the best combination is selected. Based on the two bearing whole lifetime data sets, the influence of different nonlinear module groups on network loss and training speed are shown in Figure 6.

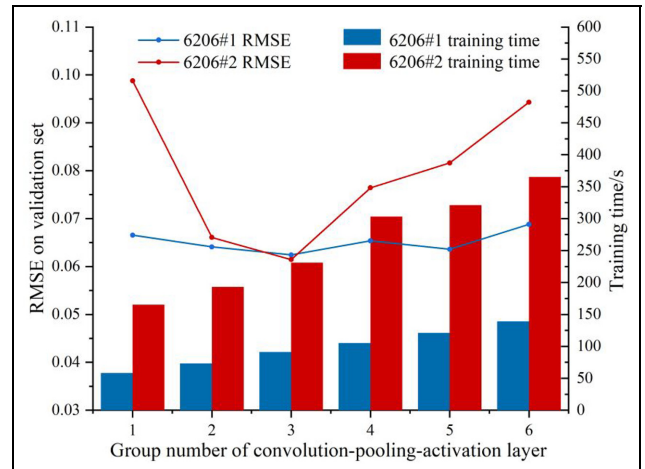As can be seen from the figure that, with the convolution-pooling-activation layer group increases, the loss of the network on the verification set decreases significantly and reaching the lowest at the three groups. After that, with the increase of network depth, the gradient instability leads to over fitting, which

Table 3. Parameters of CNN network.

Structure	Input layer	P1	C1	P2	C2	P3	C3	F1	Output
Parameter	$64 \times 32 \times 1$	2×2	3×3, 16	2×2	3×3, 32	2×2	3×3, 32	1	1
Output	$64 \times 32 \times 1$	$32 \times 16 \times 1$	$32 \times 16 \times 16$	$16 \times 8 \times 16$	$16 \times 8 \times 32$	$8 \times 4 \times 32$	$8 \times 4 \times 32$	$1 \times 1 \times 1$	1
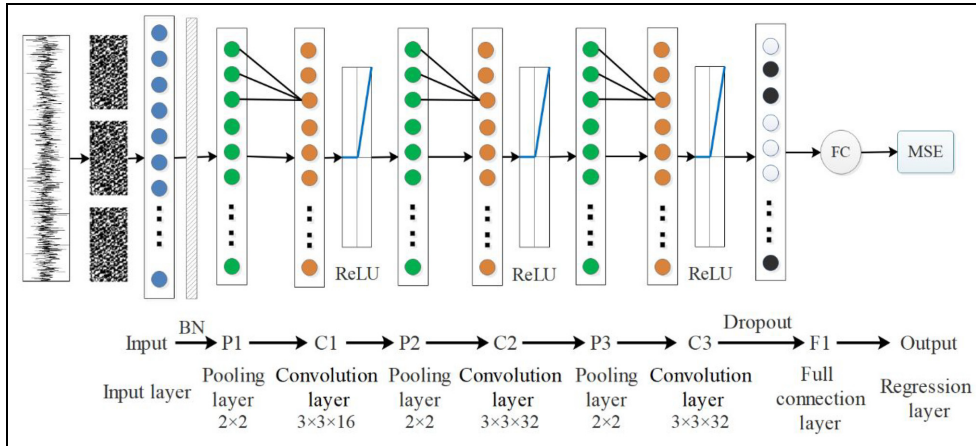
Figure 7. Structure of CNN network.

makes the loss of the network increases gradually, and the complexity of the network leads to the obvious increase of training time. Therefore, this paper selects three groups of nonlinear modules to form the main part of the network, that is, to develop a deep convolution network model which is composed of input layer, three convolution layers, three pooling layers, three ReLU layers, dropout layer ($p = 0.2$), BN layer, full connection layer, and regression layer. Specific parameters inside the network are determined through variables control and comparative analysis, namely, the parameters are selected to minimize the network training loss. The CNN network structure is shown in Figure 7. The corresponding network parameters are shown in Table 3.

The whole lifetime of rolling bearing is divided into three stages: normal stage, degradation stage, and failure stage. Before using particle filter to predict the remaining useful life, it is indispensable to extract indicators that can reflect the evolution stage of bearing. Deep learning method can adaptively extract the deep features of samples through supervised learning and eliminate redundant features without prior experience, which it is widely used in the field of feature engineering.

Relying on the powerful feature extraction ability of deep learning, this paper proposes a new degradation index extraction method of rolling bearing based on deep convolution neural network. By developing a deep convolution neural network model, the samples in the normal stage of rolling bearing are training under supervision, and then the whole lifetime data are input into the well-trained model successively, and finally the whole bearing deterioration feature corresponding to each time point is obtained. Theoretically, with the failure evolution of bearing, the difference between its feature value and normal value will increase. In this paper, the corresponding exponential model is developed by the characteristic distance between normal and abnormal samples, and the residual life is estimated by particle filter method.

Taking 6206#1 bearing whole lifetime data as an example, the process of extracting new deterioration index of rolling bearing based on deep CNN can be divided into four steps:

Step 1: The acceleration signals of rolling bearing from normal stage to failure are acquired every 6 min, and finally 300 time point samples in total are obtained. Each time point is composed of 65,536 vibration points.
Step 2: The normal stage data is selected according to the variation characteristics of bearing vibration value, then it is labeled and divided into training set and test set. The first 50 time points in the normal stage are chosen as training set, each of which is divided into 32 parts and converted into a data matrix with shape 64 × 32 for CNN training, which
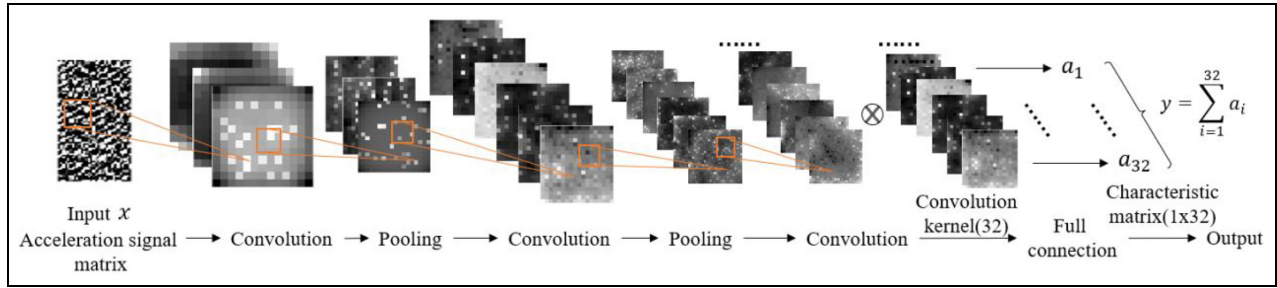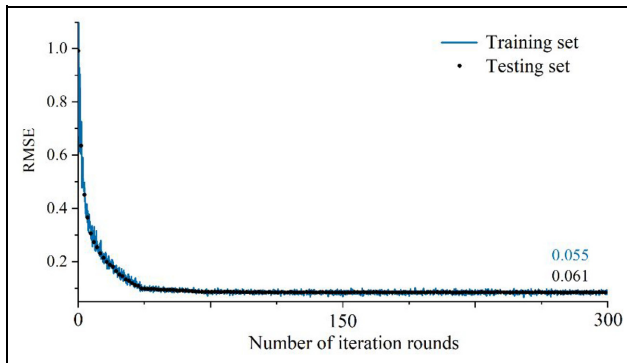
Figure 8. Extraction process of evolution index.

Figure 9. Convergence process of 6206#1 normal sample during training.

Figure 10. Evolution trend of 6206#1 bearing whole life time of CNN extracted feature.

are pixels with different gray levels according to the value of the signal data. In order to facilitate training, the label step size is normalized at 50 time points, namely, the step size is set to 0.02, representing the subtle feature changes of bearing damage in the normal stage.

Step 3: The label step size is set to 0.02 according to the normalization of 50 time points, representing the subtle feature changes in the process of bearing damage evolution in normal stage.

Step 4: The model with well-trained parameters is saved, and the samples at all time points from normal to failure are input into the model. Each sample x returns a characteristic value y. Finally, 300 characteristic values are obtained corresponding to the bearing evolution state at 300 time points. The acquisition process of evolution index y is shown in Figure 8.

Feature extraction results. The CNN models are trained based on the normal stage data of 6206#1 and 6206#2 separately according to the above theory, among which the #1 data convergence process is shown in Figure 9. The root-mean-square-error values on the final training set and test set of the model are 0.055 and 0.061
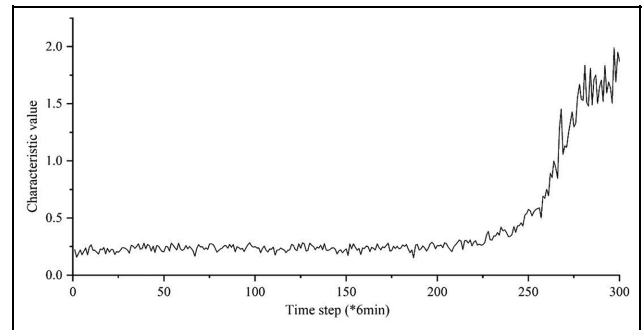
respectively, which are far less than its step size 0.2. The result proves that the network can effectively extract the deterioration characteristics of bearings.

The whole lifetime data of 6206#1 bearing under 300 time points are input into the well-trained model in sequence, and the change of CNN extracted values corresponding to 300 time points are extracted and shown in Figure 10. As can be seen from the figure that, the values begin to rise significantly when the bearing enters the degradation period, and fluctuate violently in a high range at the stage of failure, which proves that the change of CNN extracted features with time can convincingly reflect the degradation trend of the bearing.

Estimation of rolling bearing RUL based on particle filter and new deterioration index

The process of particle filter algorithm

Particle filter is an approximate Bayesian filtering algorithm based on Monte Carlo simulation. The main concept of PF is to approximate the probability density function (PDF) of the system random variables by discrete random sampling points, and replace the integral operation with the average value of sample to obtain the minimum variance estimation of the state. Based on Bayesian theory and sequential importance

sampling (SIS) algorithm, particle filter shows significant advantages in model parameter estimation of nonlinear and non-Gaussian systems, and has been widely used in the field of life prediction, such as residual life prediction of lithium battery,[17,18] crack propagation prediction of planetary gear system,[19] health monitoring of tensegrity,[20] etc.

The basic process of particle filter algorithm is as follows:

Assuming that the system satisfies the relationship between state x_k and observed value z_k under discrete time series t_k:

$$z_k = h_k(x_k, v_k) \tag{3}$$

where h_k is the observation function of the system, v_k is the observation noise, and the process of particle filter is as follows:

(1) Initialization of particle set. The particle swarm $\left\{ x_0^{(i)} \right\}_{i=1}^{N}$ is generated by using a priori probability distribution $p(x_0)$ when $k = 0$, and the particle weight $w_0^{(i)} = 1/N$;

(2) Sequence importance sampling. Sample the particle $\left\{ x_k^{(i)} \right\}_{i=1}^{N}$ in the importance distribution function at time k, and the particle set at this time is:

$$\left\{ x_{0:k}^{(i)} \right\} = \left\{ x_k^{(i)}, x_{0:k-1}^{(i)} \right\}_{i=1}^{N} \tag{4}$$

After obtaining the observation value z_k at time k, the importance weight estimation value can be obtained:

$$\omega_k^{(i)} = \omega_{k-1}^{(i)} \frac{p(z_k|x_k^{(i)})p(x_k^{(i)}|x_{k-1}^{(i)})}{q(x_k^{(i)}|x_{k-1}^{(i)}, z_{1:k})} \tag{5}$$

The particle weight is normalized:

$$\tilde{\omega}_k^{(i)} = \frac{\omega_k^{(i)}}{\sum\limits_{i=1}^{N} \omega_k^{(i)}} \tag{6}$$

(3) Resampling. By resampling the particles and probability density function represented by corresponding weights, the particles with smaller weights are reduced and those with larger weights are copied to obtain a new particle set $\left\{ x_k^{(i)} \right\}_{j=1}^{N}$, and all particles have the same weight $1/N$.

(4) The mean value of resampled particles is used for state estimation, as shown in the equation (9).

$$x_{0:k}^{est} = \sum_{i=1}^{N} x_{0:k}^{(i)} \tilde{\omega}_k^{(i)} \tag{7}$$

Where $x_{0:k}$ is the state variable at time $0:k$, $z_{1:k}$ is the observed value at time k, $p(x_0)$ is the initial probability density function, $q(x_k|x_{0:k-1}, z_{1:k})$ is the importance function, $p(x_k|x_{k-1})$ is the probability density value of system state transition; $p(z_k|x_k)$ refers to the value of the observed likelihood probability density of the state in the system.

Determination of degradation threshold and failure threshold

Alarm threshold and failure threshold are critical factors in the RUL prediction of rolling bearing. They are always determined by manual observation in traditional mind which lack of scientific basis and not conducive to the adaptive intelligent diagnosis of rolling bearing. Therefore, a k-means clustering method is proposed in this paper to automatically divide the evolution stage of rolling bearing, and the ordinate corresponding to the starting point of each class is taken as the end threshold of the previous stage.

The k-means clustering is an efficient unsupervised learning algorithm, which divides data into various regions by iterative solution, and the distance between each region is used as the standard of similarity measure. The solution objective is to minimize the sum of distances, that is:

$$\min \sum_{i=1}^{k} \sum_{x \in C_i} \|x - \mu_i\|_2^2 \tag{8}$$

Where x is the feature value, C_i is the cluster partition, μ_i is the centroid of C_i, of which the expression is,

$$\mu_i = \frac{1}{|C_i|} \sum_{x \in C_i} x \tag{9}$$

Different deterioration degrees of rolling bearing have different meanings for the actual evolution monitoring and maintenance decision, thus it is necessary to divide the fault stage in detail. In this paper, the clustering of rolling bearing evolution stage is set as four categories, and the Euclidean distance judgment criterion is adopted. Finally, the evolution stages division of the two groups of rolling bearings based on RMS and CNN feature are obtained, as shown in Figures 11 and 12.

As can be seen from Figures 11 and 12, the k-means clustering method can adaptively divide the evolution characteristics of rolling bearing into four stages: normal, early and late degradation, and failure based on the unsupervised data-driven method. In this paper, the
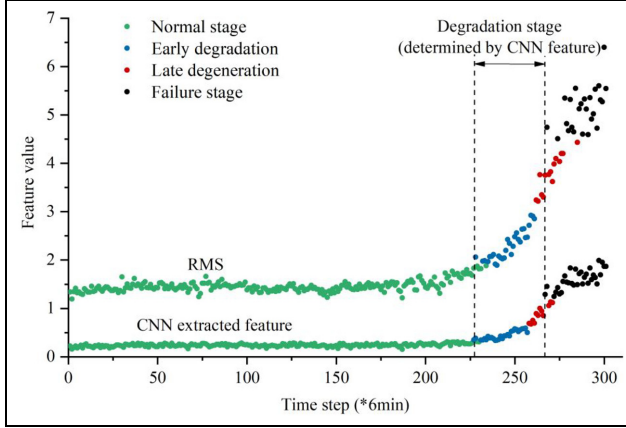
Figure 11. Evolution stages division of RMS and CNN extracted feature of HRB 6206#1 bearing after k-means clustering.
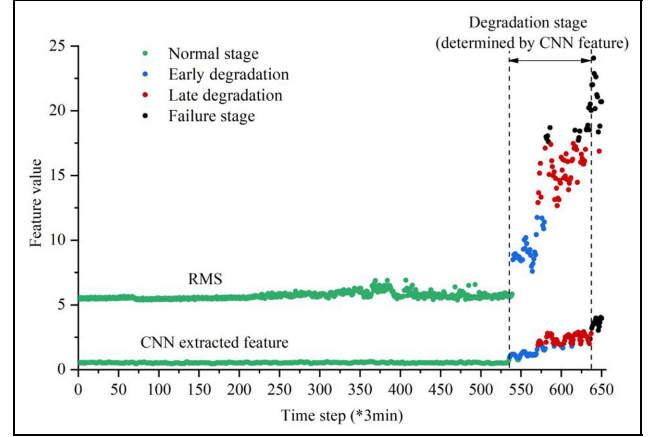

Figure 12. Evolution stages division of RMS and CNN extracted feature of HRB 6206#2 bearing after k-means clustering.

Table 4. The degradation and failure thresholds of whole lifetime evolution data of HRB6206 bearing determined by *k*-means clustering algorithm.

Data set	6206#1		6206#2	
Feature category	RMS	CNN	RMS	CNN
Degradation starting point	228	227	540	536
Degradation threshold	2.063	0.349	8.496	0.970
Failure starting point	268	267	582	638
Failure threshold	4.743	1.287	17.86	3.209

first points of different colors are set as the starting point of this evolution stage, and the starting point of degradation stage and failure stage based on CNN characteristics are obtained, as shown by the dotted line in figures above. The corresponding feature values under those times are the degradation threshold and failure threshold of rolling bearing respectively.

By the comparisons of Figures 11 and 12, it can be seen that compared with RMS, the CNN extracted features have fewer misjudgment points in different stages and have higher prediction accuracy. In particular, in Figure 12, the judgment between degradation stage and failure stage based on RMS is quite vague. In contrast, CNN has a clearer evolution trend.

According to the clustering results, the degradation threshold and failure threshold of two groups of rolling bearing data are extracted, as shown in Table 4. It can also be seen from the table that the feature extracted by CNN is earlier than RMS in judging degradation and failure, which proves that this method can find faults earlier and is of great significance for early diagnosis of bearing damage.

Method verification based on whole lifetime test of rolling bearing

Assuming that the characteristic value of the whole lifetime vibration signal of the rolling bearing conforms to the equation: $y = a \times e^{bT} + c \times e^{dT}$, where T is the time step, a, b, c, d contains noise which is proposed to be Gaussian white noise, and the state vector of the prediction model is:

$$x(T) = [a(T), b(T), c(T), d(T)] \qquad (10)$$

The state equation is:

$$\begin{cases} a(T+1) = a(T) + w_a(T), w_a \sim N(0, \sigma_a) \\ b(T+1) = b(t) + w_b(T), w_b \sim N(0, \sigma_b) \\ c(T+1) = c(T) + w_c(T), w_c \sim N(0, \sigma_c) \\ d(T+1) = d(T) + w_d(T), w_d \sim N(0, \sigma_d) \end{cases} \qquad (11)$$

The observation equation is:

$$\begin{aligned} y(T) = {} & a(T) \times e^{b(T) \times T} + c(T) \times e^{d(T) \times T} \\ & + v(T), v(T) \sim N(0, \sigma_v) \end{aligned} \qquad (12)$$

The process of bearing residual lifetime prediction based on particle filter algorithm is as follows:

(1) The acceleration signal of rolling bearing is acquainted, and the characteristic value reflecting the bearing state is extracted by convolution neural network;

(2) Set the alarm threshold and failure threshold. When the characteristic value reaches the alarm threshold at time T, it is marked as the beginning of degradation, and T is the starting time of the RUL prediction of the bearing. When the characteristic value reaches the failure threshold, the bearing is judged as failure and the prediction is ended;

(3) A four-parameter-exponential model is developed, and the initial parameters a_0, b_0, c_0, d_0 of the model are obtained by fitting the previous n characteristic data before degradation by least square fitting;

(4) The particle filter algorithm is used to predict the characteristic value at time $T + 1$ forward and compare it with the failure threshold. If it is less than the failure threshold, continue to predict the characteristic value at time $T + 2$ until it is over the threshold, then stop the prediction and record the time as T_r, and predict the RUL $= T_r - T$ at this time;

(5) Continue to collect the vibration signal, extract the characteristic value at time $T + 1$, also take the previous n characteristic data at that time, and repeat steps (3) and (4) to obtain the RUL at time $T + 1$;

(6) When the extracted characteristic value (also called observed value) exceeds the given failure threshold, the prediction algorithm terminates, and the predicted rolling bearing RUL curve from initial degradation to complete failure is obtained. The predicted RUL is compared with the real RUL to obtain the error value of PF tracking prediction.

The error of bearing RUL prediction can be expressed by mean absolute error (MAE) and root mean square error (RMSE), and their expressions are as follows:

$$MAE = \frac{1}{N}\sum_{i=1}^{m}(RUL_p(i) - RUL_r(i)) \qquad (13)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{m}(RUL_p(i) - RUL_r(i))^2} \qquad (14)$$

Where RUL_p and RUL_r are the predicted RUL and real RUL value of rolling bearing respectively, m is the number of predicted time step points. In order to prove the advantages of CNN extracted feature in rolling bearing RUL prediction, it is compared with the root mean square (RMS) value which is commonly used in bearing RUL prediction.[21,22] The RMS and CNN extracted features of the whole lifetime vibration data of rolling bearing are extracted respectively, then the two values are used to track and predict bearing RUL by the particle filter method, and finally the prediction results are compared. The discrete expression of RMS is:

$$x_{RMS} = \sqrt{\sum_{i=1}^{n} x_i^2/n} \qquad (15)$$

The physical meaning of features extracted by CNN can be explained as follows: The CNN extracted feature corresponds to the normalized label during training, that is, the time step of bearing damage evolution. Since the evolution is positively correlated with the damage situation of the bearing, the CNN extracted feature can also be interpreted as the damage degree of the bearing, which mapped to the specific size of the damage. When the extent of damage reaches a certain degree, namely the failure threshold, the bearing can be determined as failure.

As for the 6206#1 bearing, the particle filter algorithm is used to track and estimate the observed feature values between time steps 200–220. The 227th point and the 267th point are determined as the starting point of degradation and failure based on CNN feature respectively by clustering method, and the corresponding feature values under those times are set as degradation threshold and failure threshold respectively. In this paper, the change of feature value is gradually predicted from the 227th time step until the failure threshold is reached, and the RULs under 41 time steps are finally obtained. The RMS value is predicted under the same period as a contrast. The resampling method is set as residual resampling, and the number of particles is 1000.

The tracking prediction process of CNN extracted feature at the 238th time step is shown in Figure 13. It can be seen from the figure that the predicted value and the observed value have a high degree of coincidence, and the distance of intersections between them and the failure threshold is close, which proves high accuracy of prediction. The RUL prediction error of the step is $2 \times 6\,min$, which is within an acceptable range compared with the 30 h total lifetime of the fatigue strengthening test of rolling bearing.

As for the 6206#2 bearing, the observed feature values are tracked and estimated by PF algorithm between time steps 500–535. The changes of CNN feature and RMS in the early degradation stage are predicted
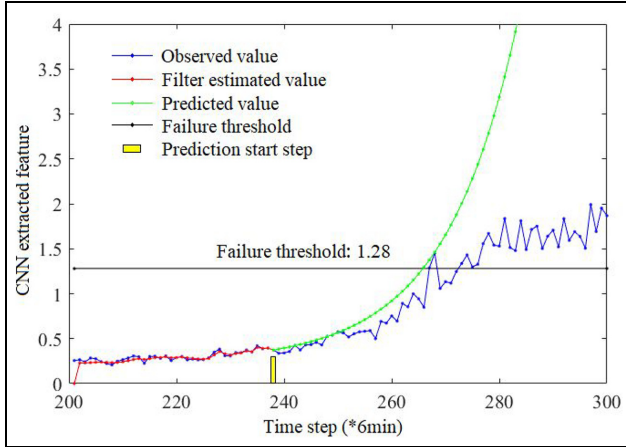
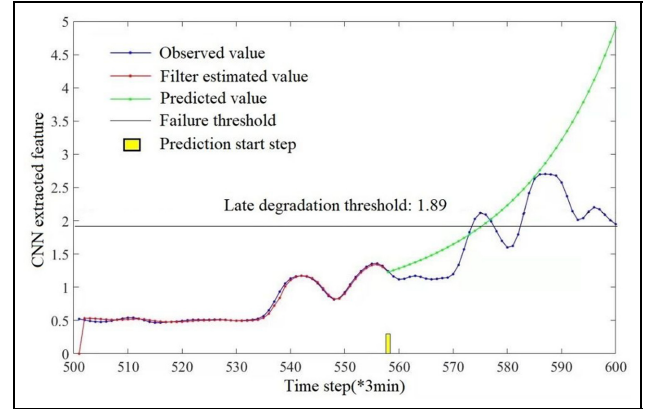**Figure 13.** Tracking prediction of CNN feature at 238th time step (6206#1).

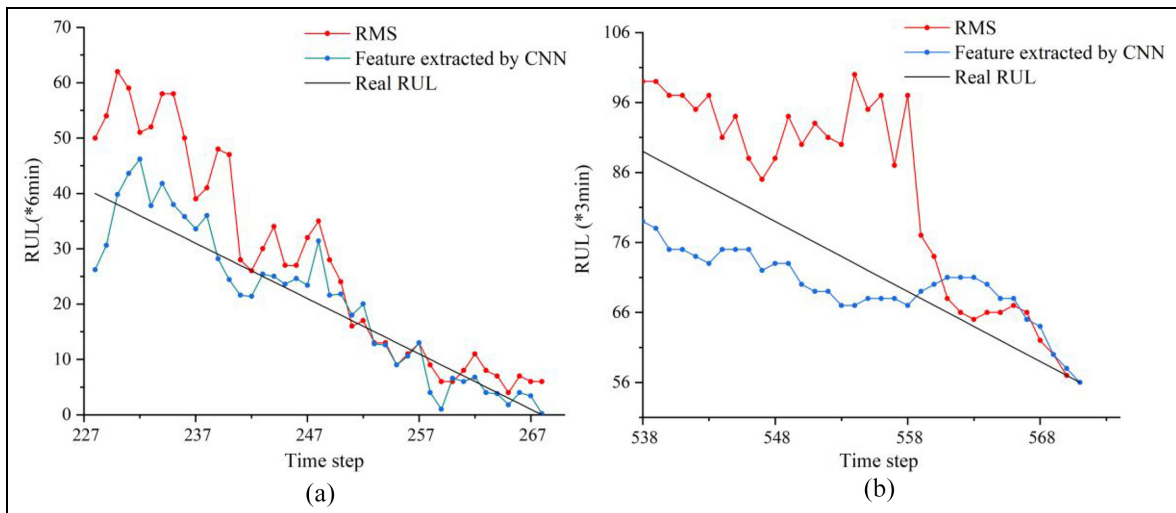**Figure 14.** Tracking prediction of CNN feature at 558th time step (6206#2).

Figure 15. RUL predicted results based on particle filter: (a) HRB 6206#1 and (b) HRB 6206#2.

gradually from the 536th point, and finally the RULs under 34 time steps are obtained. Among them, the CNN feature prediction results at the 558th time step are shown in Figure 14. The RUL prediction error of the step is 3×3 min, which also means high prediction accuracy.

According to the above method, the bearing RULs within the degradation range are tracked and predicted step by step by particle filter, and the two groups of RUL prediction results based on the RMS and CNN extracted feature are obtained respectively, as shown in Figure 15.

It can be seen from the figures that the prediction result based on CNN extracted feature is closer to the real value than those based on RMS. Especially for #2 bearing, the curve of RMS fluctuates violently in the degradation stage, which produces large errors in the

updating process of the model, resulting in a decline of prediction accuracy. On the contrary, the method based on CNN feature has a relatively stable change trend, which can better adapt to the adjustment of model parameters, leading to better prediction results than RMS.

RUL prediction based on ZA-2115 bearing

In order to examine the ability of CNN model to extract the whole lifetime characteristics of rolling bearings under diverse bearing models and working conditions, the whole lifetime vibration data of rolling bearings in IMS (Intelligent Maintenance Systems) laboratory of University of Cincinnati (UC, USA) are used for comparative study. The test rig is shown in Figure 15(a). Rexnord ZA-2115 rolling bearings are
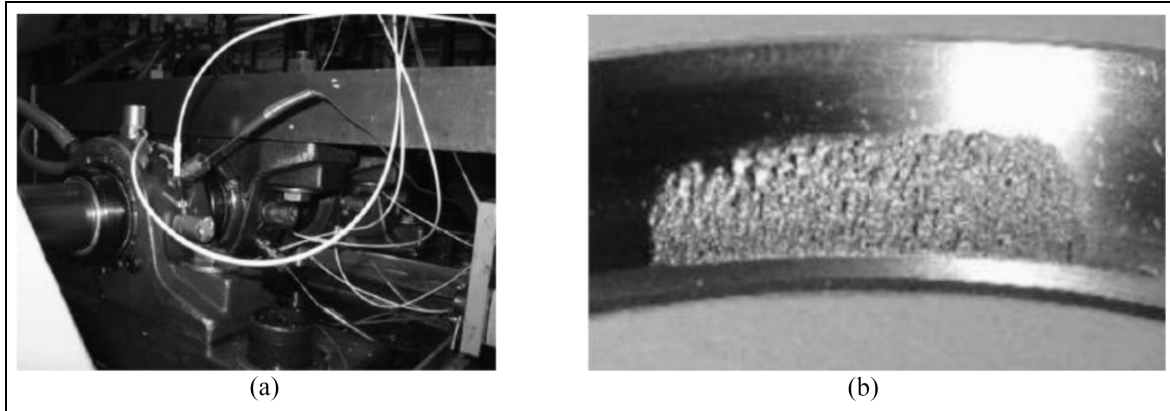
Figure 16. Whole life time test of Rexnord ZA-2115 rolling bearing: (a) bearing test rig and (b) outer race failure of ZA-2115 bearing.
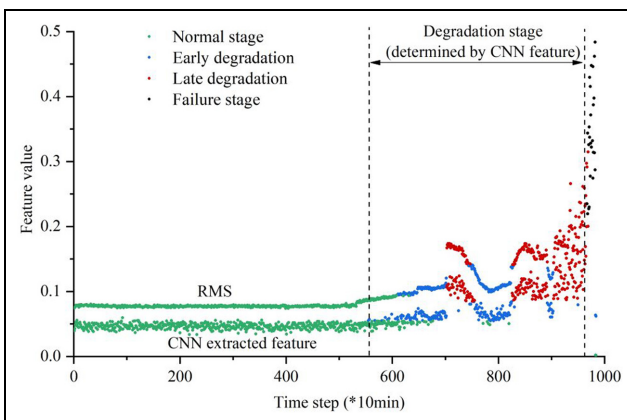

Figure 17. Evolution stages division of RMS and CNN feature of ZA-2115 bearing after k-means clustering.

Figure 18. RUL predicted results of ZA-2115 bearing.

used in the exceeding designed lifetime test, and PCB 353b33 acceleration sensors are used for monitoring. The test lasts 164 h, the data storage time is 10 min, and the sampling frequency is 20 kHz. Outer race failure of one of the bearings occurred at the end of the test, as shown in Figure 16(b).

The CNN feature and RMS under 984 time steps are extracted to track the degradation process of the bearing after k-means clustering method, as shown in Figure 17.

According to the clustering results, the degradation threshold and failure threshold of ZA-2115 rolling bearing data are determined, as shown in Table 5.

It can be seen from Figure 16 that the RMS and CNN feature still have similar upward trend in the whole life stage. Significantly, the two evolution curves began to show slow upward trends at about the 550th point, indicating slight degradation of bearing which is captured exactly by CNN. The starting point of failure threshold determined by k-means clustering based on
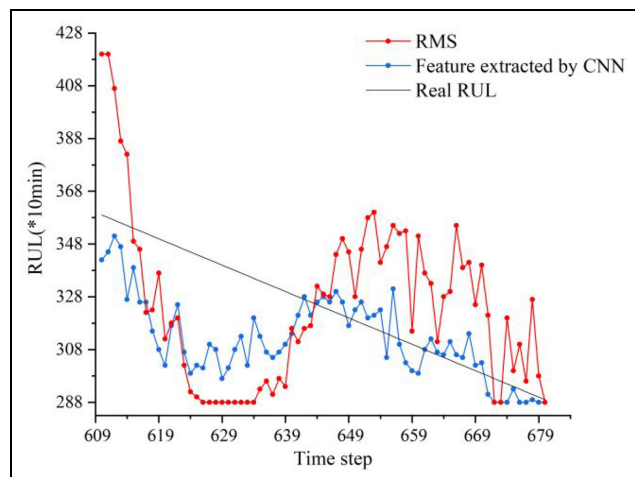
Table 5. The degradation and failure thresholds of whole lifetime evolution data of ZA-2115 bearing determined by k-means clustering algorithm.

Data set	ZA-2115	
Feature category	RMS	CNN
Degradation starting point	609	555
Degradation threshold	0.095	0.055
Failure starting point	968	967
Failure threshold	0.344	0.235

CNN feature is 56 time steps earlier than RMS and closer to the real inflection point, which is proved that the CNN method can find faults earlier and predict the RUL more accurately.

Based on the above thresholds, PF step-by-step tracking is carried out for RUL prediction within the

Table 6. Comparison of RUL prediction accuracies by particle filter based on RMS and CNN feature.

Feature extraction method	HRB 6206#1		HRB 6206#2		Rexnord ZA-2115	
	RMS	CNN	RMS	CNN	RMS	CNN
RMSE of fitting curve	0.2372	0.1053	0.9624	0.2259	0.0234	0.0177
RMSE of prediction results	14.25	7.45	12.78	6.73	34.52	20.49
Normalized RMSE of prediction results	4.67%	2.48%	1.97%	1.04%	3.51%	2.08%

Table 7. Comparison of normalized MAE by particle filter based on RMS (6206#1).

Number of particles	200	500	1000	2000	5000	10000
Polynomial resampling	5.042%	5.124%	4.503%	5.550%	5.422%	5.618%
Random resampling	5.327%	5.016%	4.453%	5.251%	5.764%	5.586%
Residual resampling	5.010%	4.921%	4.364%	5.289%	5.460%	5.390%
System resampling	5.054%	4.763%	4.408%	4.693%	4.693%	4.693%

early degradation range, and RUL trend prediction results based on RMS and CNN extracted features are obtained respectively, as shown in Figure 18.

Comparing the prediction results of ZA-2115 bearing and 6206 bearings, it is found that the average error and error fluctuation of ZA-2115 based on CNN are significantly greater than those of 6206. The reason is that the prediction stage (71 time steps × 10 min) of ZA-2115 is relatively larger than 6206#1 (41 time steps × 6 min) and 6206#2 (35 time steps × 3 min), which has more iterations by particle filter and results in error accumulation. In addition, the rising trend of values in the prediction stage of 6206 bearings are more stable than that of ZA-2115, which are more conducive to the updating of exponential model parameters.

Comparative analysis of test results

Comparative analysis of three groups of experimental data. In order to specifically compare the RUL prediction ability of the features of different characteristics in different test data, the RMSE of the least squares fitting curve, the RMSE of the prediction results, and the RMSE of the prediction results normalized by whole lifetime are compared, as shown in Table 6.

As can be seen from the above comparison:

(1) For three whole life experimental data sets, the bearing degradation characteristics extracted by CNN have higher prediction accuracy of remaining useful life than RMS. Among them, the prediction accuracy of HRB 6206 series bearing data set is higher than that of Rexnord ZA-2115 bearing data set. The reason is that the latter has more and larger prediction time steps, leading to more iterations by particle

filter and error accumulation. In addition, the characteristic ascending trend of ZA-2115 is steeper, which has an adverse impact on the updating and prediction of index model parameters.

(2) As for the whole life test data of HRB 6206 series bearings, the RMSE of the model fitting and RUL prediction results of CNN extracted features are significantly lower than the that of RMS, which proves the effectiveness of the CNN feature in the RUL prediction of bearing based on particle filter. Due to the various types of faults and other effects such as working conditions and noise, #1 bearing data characteristic curve is more stable, leading to higher prediction accuracy than that of #2 bearing.

In order to investigate the influence of parameters on particle filter algorithm and the stability of the model, this paper compares the effects of diverse resampling methods on the accuracy of the model under distinct particle numbers. The comparison of RUL tracking prediction results based on CNN feature and RMS of 6206#1 bearing data are shown in Tables 7 and 8.

As can be seen from Tables 7 and 8:

(1) When the number of particles is 1000, the average errors obtained by particle filter using RMS and CNN feature are both relatively low, and the error of system resampling is lower than that of other resampling methods;

(2) Compared with RMS, all the errors of the bearing damage evolution characteristics extracted by CNN are lower than RMS after particle filter tracking and prediction, and the accuracies

Table 8. Comparison of normalized MAE by particle filter based on CNN feature (6206#1).

Number of particles	200	500	1000	2000	5000	10000
Polynomial resampling	2.241%	2.284%	2.216%	2.203%	2.235%	2.216%
Random resampling	2.297%	2.216%	2.210%	2.266%	2.253%	2.247%
Residual resampling	2.415%	2.210%	2.197%	2.266%	2.241%	2.259%
System resampling	2.235%	2.197%	2.129%	2.179%	2.203%	2.236%

are less influenced by the number of particles or resampling method, which proves better performance of CNN method.

Comparative analysis with single deep learning model. For the CNN and PF combined model, CNN plays the role of degradation feature extraction and dimension reduction, while PF model is responsible for RUL prediction based on degradation features. However, CNN itself has the ability of regression prediction and is also widely used in the field of bearing residual life prediction,[23] while those trainings are usually based on both normal and abnormal data. In order to explore the difference between single CNN regression prediction and CNN + PF combined prediction based only on normal data training, the feature of whole life time data is extracted after normal data training by the same model as above, and the CNN regression model is developed to predict the data in the degradation stage. The comparison of prediction results between the single CNN regression and CNN + PF is shown in Figure 19.

As can be seen from Figure 19, due to the huge difference in data distribution between normal stage and degeneration stage, the single CNN regression model fails to predict the evolution trend of abnormal data after training only based on normal data. On the contrary, the particle filter method can solve this problem, for it can predict the prior estimate at current time according to the optimal estimate of the last time by constructing an exponential equation with very similar evolution characteristics, and achieve gradual correction by using the observation equation, so as to obtain the prediction result of the remaining life that is closest to the reality. The comparison results show that the combination model of CNN and PF is feasible.

Summary

(1) Through the comparison of three different whole lifetime test results of rolling bearings, it is proved that the bearing degradation features extracted by convolutional neural network have higher prediction accuracy of remaining useful life than the traditional feature index
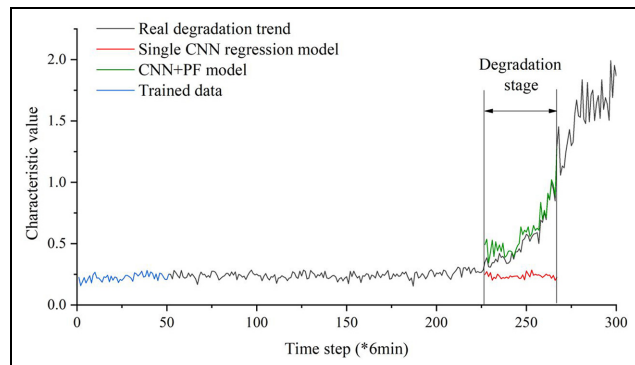

Figure 19. Extracted evolution trend based on single CNN regression model and CNN + PF model.

RMS, which shows that the bearing degradation features extracted based on deep learning method are more sufficient and can better reflect the RUL situation of bearings;

(2) By comparing the particle filter prediction errors of the two characteristics under various sampling methods and particle numbers, it is found that the overall error is lower when the particle number is 1000, and the system resampling method has the best performance. Among them, the CNN characteristic error is significantly lower than RMS;

(3) All the errors of the bearing damage evolution characteristics extracted by CNN are lower than that of RMS after particle filter tracking and prediction, and the accuracies are less influenced by the number of particles or resampling method, which proves better performance of deep learning method;

(4) Compared with the traditional deep regression model, the CNN and PF combined model can predict the evolution of degradation state accurately only based on normal data, which is of great significance in practical engineering application.