Prediction of remaining useful life of rolling bearing based on fractal dimension and convolutional neural network

Guorong Ding, Wenbo Wang and Jiaojiao Zhao

Abstract
In order to predict the remaining useful life (RUL) of rolling bearings in complex environmental conditions, a bearing RUL prediction method based on fractal dimension and one-dimensional convolutional neural network (1D-CNN) is proposed. This method uses fractal dimension to characterize the degeneration process of the rolling bearing and combines the features of time domain, frequency domain, wavelet packet domain, and entropy domain. Fractal dimension provides an analytical method for characterizing the complexity of vibration signals. The features extracted from different feature domains can complement each other’s advantages, reveal the degradation state of the bearing more comprehensively and achieve better performance. Then, the percentage of the remaining life of the bearing is used as the degradation tracking index of the rolling bearing. Based on the idea of stratified sampling, the training samples and test samples are constructed, and then a 6-layer CNN is constructed to train the model. Finally, three commonly used data sets of full-life bearings are used to verify the model, namely, IEEE prognostics and health management 2012 Data Challenge, IMS dataset, and XJTU-SY dataset. The experimental results show that, on the three experimental datasets, compared to the long short-term memory network (LSTM) and the extreme learning machine (ELM) methods, the prediction effect of the RUL of the bearing based on the fractal dimension and 1D-CNN proposed in our paper is better. Its mean absolute error and root mean square error (RMSE) and mean absolute percentage error (MAPE) have been reduced, and the correlation index ($R^2$), adjusted $R^2$, and relative accuracy (RA) have been improved, which can predict the RUL of the bearing more accurately.

Keywords
Remaining life prediction, fractal dimension, convolutional neural network, degradation feature extraction

Introduction
With the continuous advancement of science and technology, machinery and equipment are increasingly developing in the direction of large-scale, precise, intelligent, automated, and systematic. In order to meet the needs of production, machinery and equipment are becoming more and more intelligent and complex. However, the working environment is often harsh and changeable. The equipment will gradually age after long-term operation, and the remaining useful life (RUL) will gradually decrease, thereby greatly increasing the potential for failure.1 Driven by big data and artificial intelligence, research on fault prediction and prognostics and health management (PHM) technology has become more in-depth. This technology is not only traditional problem diagnosis, but also predictive performance analysis and intelligent maintenance of future states. Therefore, accurately predicting the RUL of the equipment is of great significance to the preventive maintenance decision-making of mechanical equipment.

Bearings are a key part of rotating electromechanical equipment, and their reliable operation can improve the safety and efficiency of modern production equipment. Generally speaking, bearings usually work in different environments, and different types of failures often occur during operation. If effective protective measures are not taken in time, it may cause the entire machine to malfunction and cause a lot of economic losses. Therefore, RUL prediction of bearings will be a very important task. It can provide early warning reports for equipment maintenance personnel and replace parts in advance, thereby reducing expensive unplanned maintenance and improving the safety of equipment operation.2

At present, the mainstream prediction methods of RUL mainly fall into three categories: prediction methods based on physical models, data-driven, and hybrid models.3–4 Essentially, these three methods can find applications in practical problems. However, the use of physical mechanisms and hybrid-based methods may be limited in practice, due to the lack of underlying physical knowledge in many
practical components and systems. Remaining useful life prediction methods based on physical mechanisms generally have problems such as greater difficulty in modeling systems with higher complexity and lower universality of the established models on other systems. Because it is difficult to clearly and comprehensively describe the complex process of bearing degradation, data-driven can automatically infer hidden data. The causal relationship in the system directly extracts the degradation characteristics of the complex system, which can better process the massive monitoring data and provide accurate RUL prediction results. Generally speaking, the data-driven RUL prediction process includes three steps. First, different types of monitoring data, such as vibration or sound wave emission signals, are obtained through various sensors. Second, representative features are extracted from them using signal processing techniques, and then the selected features are input into Gaussian process regression, support vector machine (SVM), recurrent neural network (RNN), and other learning models to train the degradation behavior of the machine. Finally, these models are used to predict and estimate RUL.

With the continuous improvement of computer hardware level, deep learning technology has received more and more attention in data-driven RUL prediction. Scholars have also conducted some research on machine RUL prediction based on deep learning. Kang et al. used an improved sparse autoencoder to extract features, combined with a bi-directional LSTM (Bi-LSTM) to construct a prediction model, and estimate rolling elements Bearing RUL. Tian et al. developed a method based on artificial neural network (ANN) to achieve more accurate RUL prediction for state monitoring equipment, but the prediction results for small samples are not ideal. Sun et al. developed a RUL prediction model based on SVM and applied the model to bearing life prediction. The model is applied to the life prediction of bearings, and the results show that the model has high accuracy, but this method cannot extract the features needed to fully reflect the bearing degradation by using traditional feature extraction methods. Ren et al. proposed a deep learning method for predicting the RUL of multiple bearings by combining time domain and frequency-domain features. Cao et al. proposed a new deep prediction network-parallel multiple CNN with compression excitation mechanism and Bi-LSTM to integrate the entire network, using RUL prediction for rolling bearings. Han et al. tried to use the horizontal vibration acceleration signal and the vertical acceleration signal of the bearing to construct a one-dimensional CNN to realize automatic feature extraction. By setting the starting degradation point of the bearing sample, the RUL of the training sample is more accurate. Ding et al. proposed a stratified sampling which is different from the traditional time series data division method and applied to data division, in order to fully understand the data characteristics, use the fast Fourier transform to obtain the frequency characteristics from the original data, and use the mean square. The root is used as the tracking metric; finally, a deep convolutional neural network model with no pooling layer composed of three convolutional layers and two fully connected layers is constructed. J. Yoo et al. developed a time-frequency image to construct a health index (HI) and used CNN to predict RUL. In order to improve the RUL prediction value, RNN, AE, and bi-directional combined gated recurrent unit (Bi-GRU) have also been applied to the prediction of RUL bearings. CNN has also obtained excellent RUL prediction results due to its wide application and powerful feature extraction capabilities.

When rolling bearings are in different fault states, due to the difference in stiffness non-linearity, clearance, friction, and external load, different non-linear and non-stationary influencing factors are caused. These influences are reflected by vibration signals. How to extract effective fault features from the fault vibration signal is the key to fault diagnosis. The complexity of the geometric shape of the vibration signal of rolling bearings is related to the degree of non-linearity of the system. Fractal dimension provides an analysis method to characterize the complexity of the vibration signal. It can extract the state information of the equipment as a whole. It has been successfully applied in many fields. Fractal dimension is to characterize the irregularity and complexity of the fractal set at different scales. The fractal dimension obtained according to the fractal theory can describe the structural characteristics of the signal. Therefore, this paper takes fractal features into account in the extraction of the degradation characteristics of rolling bearings and combines the time-domain, frequency-domain, and time-frequency-domain features proposed in the previous literature to better reflect the degradation state of the bearing. The main tasks of this article are as follows:

1. Three commonly used data sets in the field of bearing failure health management, PHM 2012 challenge data set, IMS data set, and XJTU-SY data set were selected. Systematically analyze the bearing degradation information in these three data sets and deeply understand the degradation status of bearings under different environmental conditions.
2. Introduce the fractal dimension into the measurement to characterize the degradation of rolling bearings. Combine multiple feature domains such as time domain, frequency domain, wavelet packet domain, and entropy domain to extract the multi-dimensional fusion features of the original bearing. The percentage of remaining life is used as the tracking metric label of the training model.
3. The method of data stratified sampling is used to divide the training set and the test set, in which the length of the intercepted data is consistent with the dimension of the extracted features. The method of data stratified sampling can make the training model robust.
4. Construct a 6-layer CNN to fit the hidden mapping relationship between the degradation index and the multi-dimensional fusion features. CNN can automatically extract deep features. Due to its generalization and robustness, the prediction effect on other bearings is also better without changing the super parameters of the model.

The organization work of this article is as follows: The Preliminaries section introduces the basic knowledge of CNN structure and fractal dimension; the Experimental framework section explains the construction of the model and the evaluation indicators used in this article; the Experiment section presents the specific experimental results and analysis; The last section is the conclusion of this article and discusses future work.
Preliminaries

CNN structure

Convolutional neural network is a very popular deep learning framework model with powerful feature extraction capabilities and has been well applied in image recognition, natural language processing, and other fields. The weight-sharing network structure of CNN makes it more similar to a biological neural network, which reduces the complexity of the network model and reduces the number of weights. This advantage is more obvious when the input of the network is a multi-dimensional image, so that the image can be directly used as the input of the network, avoiding the complicated process of feature extraction and data reconstruction in traditional recognition algorithms.

Convolutional neural network mainly consists of three main components: convolutional layer, pooling layer, and fully connected layer. Generally, the convolutional layer and the pooling layer will appear in pairs. The neurons in the convolutional layer in CNN are only connected to certain neurons instead of connecting to all neurons. Local connection, weight sharing, and sub-sampling in space or time make the three structural characteristics of CNN. The more common CNN structure is shown in Figure 1.

The function of the convolutional layer is to perform feature extraction by performing convolution operations on the local area of the input data and the convolution kernel, and to make the local receptive field traverse the entire input data by sliding the convolution kernel window. Each element that composes the convolution kernel corresponds to a weight coefficient and a bias, which is similar to the neuron of a feedforward neural network. Each neuron in the convolutional layer is close to the area of the previous layer. Each neuron is connected, and the size of the region depends on the size of the convolution kernel, which is called "receptive field" in the literature.

The convolution formula is defined as follows:

$$X_l^{(t+1)} = f(W_l^{(t+1)} * X_l^{(t)} + B_l^{(t+1)})$$

(1)

where $X_l^{(t+1)}$ represents the $(l + 1)$th feature of the output value of the $l$th layer. $W_l^{(t+1)}$ represents the weight matrix of the $l$th convolution kernel of the $(l + 1)$ layer. $*$ represents the convolution operation, $X_l^{(t)}$ represents the output of the $l$-th layer, and $B_l^{(t+1)}$ represents the bias term. The function $f$ represents the output activation function. CNN solves the non-linear problems in real life through the non-linear activation function. The commonly used activation functions are: Sigmoid, Tanh, and ReLU activation function.

The convolution process of CNN enables the convolution kernel to act on the feature area through a specified step size, thereby extracting new features. The weight of the same convolution kernel is the same, and different convolution kernels will get different outputs, and all outputs form the feature input of the next layer of pooling.

After feature extraction in the convolutional layer, the input features will be passed to the pooling layer for feature selection and information filtering. The pooling layer contains a preset pooling function, and its function is to replace the result of a single point of the feature map with the feature statistics of the adjacent area. The pooling layer selects the pooling area in the same steps as the scanning feature of the convolution kernel, which is controlled by the pooling size, step size, and filling. The role of the pooling layer is spatial merging or down sampling, which can reduce the dimension of the feature map while maintaining the most important signals. Generally, there are average pooling and maximum pooling; the expression for maximum pooling is as follows:

$$y_l^{(t+1)} = \max(x_l^{(t)}(k), k \in D_j)$$

(2)

$y_l^{(t+1)}$ represents the element in the $i$th feature map of the $(l + 1)$th layer after pooling; $D_j$ represents the j-th pooling area, and $x_l^{(t)}(k)$ represents the element of the $i$th feature map of the $l$th layer in the range of the pooling core. The operation steps of maximum pooling is shown in the Figure 2.

In the CNN structure, after multiple convolutional layers and pooling layers, one or more fully connected layers are finally connected. Each neuron in the fully connected layer is fully connected with all neurons in the previous layer. The fully connected layer can integrate the local information of the convolutional layer or the pooling layer.
Fractal dimension

Because the fractal dimension uses numerical methods to describe the fault characteristics, compared with the frequency spectrum method, the fault description is cleaner and more intuitive, and the amount of calculation is smaller.

Fractal dimension is a parameter used to quantitatively describe the complexity and irregularity of an object. It can describe the object as a whole without relying on mathematical models. Therefore, the fractal dimension is an important feature of describing an object. At present, the more common fractal dimensions are box dimension, similar dimension, information dimension, correlation dimension, and generalized dimension. The box dimension is currently the most commonly used method for calculating the fractal dimension. Its calculation method is simple. The specific method is to cover the entire graph with small squares, and calculates the dimension by counting the number of all small squares covering the graph. The box dimension is specifically calculated by the following formula

\[ D_f = \lim_{{r \to 0}} \frac{\log(N(r))}{\log(1/r)} \]  

(3)

where \( N(r) \) represents the smallest grid count with the maximum diameter of \( r \) and that can cover the entire signal;

The size of the fractal dimension indicates the complexity and irregularity of the fractal object, and the larger the fractal dimension, the more complex the object and the more detailed components. Multifractal is a further extension of fractal theory. It uses a spectral function to describe the deeper and more detailed characteristics of the signal. With more in-depth research on fractal theory, scholars have found that only a single fractal cannot be described well. The more detailed characteristics of complex signals, and the generation of multifractals solves this problem.

Multifractal is not like a single fractal, which is a dimensional value. It is a collection of singularity measures of multiple scale indexes. Multifractal spectrum is usually used to express the relationship between the part and the whole.

The vibration signal of a rolling bearing is a non-linear and non-stationary signal. The fractal method is used to study it. Single fractal and multifractal are used to extract the characteristics of the rolling bearing vibration signal. Multifractal spectrum can more finely reflect the detail information contained in the bearing signal and the degree of degradation information. The remaining life prediction is carried out by selecting the fractal box dimension and the multifractal spectrum corresponding to the original vibration signal as part of the degradation characteristics of the rolling bearing. The feature quantities extracted from the time domain, frequency domain, time-frequency domain, entropy domain, and fractal domain can complement each other’s advantages. Combining the features of multiple feature domains can reveal the degradation state of the bearing more comprehensively and achieve a more accurate remaining life prediction.

Experimental framework

In this section, all steps of the proposed RUL prediction method will be explained in detail. Our paper provide the whole sketch map of the proposed method as shown in Figure 3. First, extract 19-dimensional features from the fractal domain, time domain, frequency domain, wavelet packet domain, and entropy domain from the original vibration signal of the rolling bearing. Each feature can reflect the degradation process of the bearing throughout its life. Normalization processing can effectively avoid the impact of differences in the data itself. Then, the extracted 19-dimensional features are expanded into a one-dimensional vector in chronological order, where the first set of features represent the features extracted at the first sampling time, and the last set of features are the comprehensive features extracted at the last sampling time. CNN secondary feature extraction is performed on the unfolded one-dimensional vector, and the remaining percentage of the bearing is used as the tracking metric. The size of the convolution kernel of the first convolution layer is 19*1, and the step size is also 19, which means that each step The operation calculates 19 comprehensive features at the same time. The activation function uses the ReLU activation function, which can effectively avoid the insufficient gradient of the Sigmoid activation function or the insufficient gradient explosion. The maximum pooling layer is used to reduce the dimensionality of features, while not losing useful features as much as possible, and can greatly speed up training. Finally, two fully connected layers are connected to integrate the features after convolution and predict the final output result.

Bearing feature extraction

Time domain features and frequency-domain features can roughly reflect the decay process and fault information of rolling bearings. However, due to the non-linear characteristics of the bearing signal, the time-domain or frequency-domain features are very insensitive to non-linear components. When there are subtle changes, they cannot truly reflect their characteristics.

Wavelet packet decomposition is based on wavelet transform (WT), which can locally characterize the bearing signal characteristics. According to the research in Ref. 26, morlet is selected as the mother wavelet to extract part of the impact characteristics in the bearing, the number of decomposition layers is 4, and the first five normalized energies are used as time-frequency domain characteristic parameters.

In the entropy domain, the physical meaning of sample entropy and approximate entropy is roughly similar. They both measure the complexity of the time series by measuring the probability of generating new patterns in the signal. The greater the probability of a new pattern, the greater the complexity of the sequence. Compared with approximate entropy, sample entropy has two advantages: the calculation of sample entropy does not depend on the length of the data and sample entropy has better consistency, that is, changes in parameters \( m \) and \( r \) have the same effect on sample entropy. The lower the value of sample entropy, the higher the self-similarity of the sequence; the larger the value of sample entropy, the more complex the sample sequence. At present,
sample entropy has applications in evaluating the complexity of physiological time series (EEG, sEMG, etc.) and diagnosing pathological conditions.

In the fractal domain, single fractal and multifractal are used to extract the characteristics of rolling bearing vibration signals. The multifractal spectrum can more finely reflect the detailed information contained in the bearing signal and the degree of degradation information. The remaining life prediction is carried out by selecting the fractal box dimension and the multifractal spectrum corresponding to the original vibration signal as part of the degradation characteristics of the rolling bearing.

Table 1, $X_i$ represents the original vibration signal, $F_i$ represents the frequency spectrum after FFT, and $S$ represents the standard deviation.

It can be seen from the above Figure 4 that the feature quantities extracted from the time domain, frequency domain, time-frequency domain, entropy domain, and fractal domain can complement each other. Starting from the characteristics of multiple feature domains, it can reveal the degradation state of the bearing more comprehensively, to achieve a more accurate remaining life prediction.

### Percentage of Remaining Useful Life

Under normal operating conditions, the life of rolling bearings can reach thousands or tens of thousands of hours. However, running such a long time in a laboratory environment obviously does not conform to the actual conditions of the experiment. For the feasibility of the experiment, artificial methods are used to change the acceleration degradation process of the bearing. The main methods are to increase the speed of the rolling bearing, apply the extra load pressure of the rolling bearing, and reduce the use of lubricating oil. Therefore, the acquisition of bearing vibration signals is generally equal interval sampling.

For an experimental dataset, assuming that a total of $n$ samples are taken during the experiment, the sampling frequency of the sensor is $f$, and the duration of each collection is $s$, then $k (k = f^s)$ data are collected each time. If the sampling interval is $m$, the life of the bearing is $n^m$. Set the number of

| Feature | Expression |
|---------|------------|
| RMS | $\text{RMS}(X) = \sqrt{\frac{1}{N} \sum_{i=1}^{N} X_i^2}$ |
| Pulse | $\text{Pulse}(X) = \frac{\text{max}(X)}{\text{RMS}(X)}$ |
| Margin | $\text{Margin}(X) = \frac{\text{max}(X)}{\text{mean}(X)}$ |
| Kurtosis | $\text{Kurt}(X) = \frac{\sum_{i=1}^{M} F_i^4}{C_0^4}$ |
| Skewness | $\text{Skew}(X) = \frac{\sum_{i=1}^{M} F_i^3}{C_0^3}$ |
| Peak_Peak | $\text{Peak_Peak}(X) = \frac{\text{max}(X) - \text{min}(X)}{\text{RMS}(X)}$ |
| Variance | $\text{Var}(X) = \frac{1}{N} \sum_{i=1}^{N} (X_i - \overline{X})^2$ |
| Spectrum mean | $\text{Spec_mean}(X) = \frac{1}{M} \sum_{i=1}^{M} F_i$ |
| Spectrum RMS | $\text{Spec_RMS}(X) = \sqrt{\frac{1}{M} \sum_{i=1}^{M} F_i^2}$ |
| Spectrum std | $\text{Spec_std}(X) = \frac{1}{M} \sum_{i=1}^{M} (F_i - \overline{F})^2$ |
| Spectrum kurtosis | $\text{Spec_Kurt}(X) = \sum_{i=1}^{M} (F_i - \overline{F})^3$ |
| Spectrum skewness | $\text{Spec_Skew}(X) = \sum_{i=1}^{M} (F_i - \overline{F})^4$ |
| WPD Energy | $\text{WPD_Energy}(X) = \sum_{i=1}^{N} \frac{C_0^2}{2}$ |
| Fractal dimension | Ref. 23 |
| Sample entropy | Ref. 26 |
rows as the number of acquisitions, and the number of columns as the length of the data collected once.

For example, for the IMS data set, each bearing in the second set of experiments was collected 984 times, 20,480 data were collected each time, its life was 984*10 = 9840 min, a total of 164 h, the number of rows was set to be the number of times of collection, and the number of columns was the length of the data collected once.

Set the life label of the \( i \)-th row of data, which represents the ratio of the time from the time corresponding to the \( i \)-th row to the time when the bearing fails, and the ratio of the time from the start time of the bearing to the time when the bearing fails. Calculated as follows

\[
y_i = \frac{n - i}{n - 1}
\]

In formula (4), \( n \) is the number of rows currently in, and \( i \) is the total number of rows. The normalization of the life label can reduce the difference between different working conditions and different life values of the bearing, which is beneficial to improve the prediction accuracy of the RUL. Obviously, when the signal is first collected, the percentage of remaining life of the bearing is 1, and when the signal is finally collected, the percentage of remaining life of the bearing is 0.

**Dataset division**

Refer to the idea in the Ref. 18, using a stratified sampling method to divide the training set and the test set in order to better understand the entire degradation state of the data. The specific division steps are shown in the Figure 5. First, the extracted multi-domain features are expanded in a time series, and then extracted at a ratio of 19:1. The first 19 points of each group of data are taken as the training set, and the last point of each group of data is taken as test set. The advantage of dividing the training set and the test set in this way is to understand the degraded state of the bearing during the entire operation process and improve the accuracy of model training.

**Evaluation index**

The following six indicators are used to measure the prediction effect of the proposed prediction model: mean absolute error (MAE), root mean squared error (RMSE), mean absolute percentage error (Mean), absolute percentage error (MAPE), and related indexes \( R^2 \), Adjusted \( R^2 \). The calculation methods are

\[
\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \quad (5)
\]

\[
\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \quad (6)
\]

\[
\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (7)
\]

\[
R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \quad (8)
\]

\[
\text{Adjusted}_R^2 = 1 - (1 - R^2) \frac{n - 1}{n - p - 1} \quad (9)
\]
The error between the length of the sampled data. Mean absolute error represents the effectiveness of the method proposed in this paper is verified. For rolling bearing and reduce the use of lubricants. The effectiveness of the model is increased by applying the additional load pressure of the bearing to be that the amplitude of the vibration signal is greater than 20 g. The specific working condition information of the experimental platform is shown in the Table 2.

The experiment set up three different working conditions, of which the first and second working conditions have seven bearings, and the third working condition has three bearings. A total of the lateral vibration signals and longitudinal vibration signals of these 17 bearings during the entire service life are collected. Here, the horizontal vibration signals are selected for the experiment. The sampling frequency of the data is 25.6 kHz, and the data is recorded every 10 s, the length of each recording is 0.1 s, and the length of data collected each time is 2560.

For example, bearing1_2 has collected 871 times, 2560 data each time, its life span is 871*10 = 8710 s, set the number of rows as the number of collections, and the number of columns as the length of the data collected once.

The original vibration signals of rolling bearings under three different working conditions and the corresponding root mean square values are shown in the Figure 6. It can be seen that under the three working conditions, the degradation degree and degradation process of each bearing are different.

2. IMS dataset

The second set of rolling bearing life cycle data comes from the University of Cincinnati, Intelligent Maintenance System (IMS). There are a total of four bearings, the acceleration vibration signal is measured, the sampling frequency is 20 kHz, the sampling period is 10 min and each sampling time is 1 s, so the number of collection points each time is 20,480. The specific working condition information of the experimental platform is shown in the Table 3. Based on the principle of the amount of debris absorbed by the magnetic screw plug in the tubing, when the preset threshold is reached, the bearing can be considered to be degraded and the collection work is stopped.

This paper selects the vibration acceleration signals of the four rolling bearings in the second set of experiments in the database as the signal data of the full life of the bearing. Each bearing in the second set of experiments is sampled 984 times, and the data is sampled every 10 min, and the sampling time is 1 s each time. The sampling frequency is 20 kHz, so the number of acquisition points each time is 20,480, and the outer ring fault frequency is 236.4 Hz. The original vibration data is 0.1 s, and the length of data collected each time is 2560.

In order to avoid the excessive influence caused by the error caused by the excessive number of data samples, the adjusted $R^2$ is used to estimate the correlation index, where $n$ represents the number of sample data, and $p$ represents the dimension of the data feature. The relative accuracy (RA) definition is used to measure the difference between the predicted value and the true value. The result is between 0 and 1. The closer to one, the better the result.

$$RA = 1 - \frac{\|y_i - \hat{y}_i\|}{\|y_i\|}$$

(10)

In formulas (5)–(10), $\hat{y}_i$ represents the predicted data, $y_i$ represents the original acquired data, and $n$ represents the length of the sampled data. Mean absolute error represents the error between $y_i$ and $\hat{y}_i$. RMSE reflects the error distribution, that is, the deviation between $y_i$ and $\hat{y}_i$. Mean absolute error is used to measure the quality of a model’s prediction results. The smaller the three kinds of errors, the better the model. The correlation index $R^2$ is used to describe the effect of fitting. The larger the $R^2$, the better the fitting effect of the model.

Experiment

Experimental dataset

Under normal operating conditions, the life of rolling under normal operating conditions, the life of rolling bearings can reach thousands or tens of thousands of hours. However, running such a long time in a laboratory environment obviously does not conform to the actual conditions of the experiment. For the feasibility of the experiment, the accelerated degradation process of the bearing is changed by manual methods. The main method is to increase the speed of rotation such that under the three working conditions, the degradation state assessment and remaining service life prediction. The acceleration sensor collects vibration information in the horizontal and vertical directions. The experiment set the condition for the complete degradation of the bearing to be that the amplitude of the vibration signal is greater than 20 g. The specific working condition information of the experimental platform is shown in the Table 2.

The experiment set up three different working conditions, of which the first and second working conditions have seven bearings, and the third working condition has three bearings. A total of the lateral vibration signals and longitudinal vibration signals of these 17 bearings during the entire service life are collected. Here, the horizontal vibration signals are selected for the experiment. The sampling frequency of the data is 25.6 kHz, and the data is recorded every 10 s, the length of each recording is 0.1 s, and the length of data collected each time is 2560.

For example, bearing1_2 has collected 871 times, 2560 data each time, its life span is 871*10 = 8710 s, set the number of rows as the number of collections, and the number of columns as the length of the data collected once.

The original vibration signals of rolling bearings under three different working conditions and the corresponding root mean square values are shown in the Figure 6. It can be seen that under the three working conditions, the degradation degree and degradation process of each bearing are different.

1. FEMTO-ST bearing dataset(IEEE PHM 2012)

The first data set selected for the experiment comes from the PHM data challenge held by the Institute of Electrical and Electronics Engineers (IEEE) in 2012. It is the vibration acceleration data collected from the rolling bearing accelerated life test bench (PRONOSTIA). The experimental platform is dedicated to It is an experimental platform for testing and verifying data-driven rolling bearing failure, degradation state assessment and remaining service life prediction. The acceleration sensor collects vibration information in the horizontal and vertical directions. The experiment set the condition for the complete degradation of the bearing to be that the amplitude of the vibration signal is greater than 20 g. The specific working condition information of the experimental platform is shown in the Table 2.

The experiment set up three different working conditions, of which the first and second working conditions have seven bearings, and the third working condition has three bearings. A total of the lateral vibration signals and longitudinal vibration signals of these 17 bearings during the entire service life are collected. Here, the horizontal vibration signals are selected for the experiment. The sampling frequency of the data is 25.6 kHz, and the data is recorded every 10 s, the length of each recording is 0.1 s, and the length of data collected each time is 2560.

For example, bearing1_2 has collected 871 times, 2560 data each time, its life span is 871*10 = 8710 s, set the number of rows as the number of collections, and the number of columns as the length of the data collected once.

The original vibration signals of rolling bearings under three different working conditions and the corresponding root mean square values are shown in the Figure 6. It can be seen that under the three working conditions, the degradation degree and degradation process of each bearing are different.

2. IMS dataset

The second set of rolling bearing life cycle data comes from the University of Cincinnati, Intelligent Maintenance System (IMS). There are a total of four bearings, the acceleration vibration signal is measured, the sampling frequency is 20 kHz, the sampling period is 10 min and each sampling time is 1 s, so the number of collection points each time is 20,480. The specific working condition information of the experimental platform is shown in the Table 3. Based on the principle of the amount of debris absorbed by the magnetic screw plug in the tubing, when the preset threshold is reached, the bearing can be considered to be degraded and the collection work is stopped.

This paper selects the vibration acceleration signals of the four rolling bearings in the second set of experiments in the database as the signal data of the full life of the bearing. Each bearing in the second set of experiments is sampled 984 times, and the data is sampled every 10 min, and the sampling time is 1 s each time. The sampling frequency is 20 kHz, so the number of acquisition points each time is 20,480, and the outer ring fault frequency is 236.4 Hz. The original vibration data is 0.1 s, and the length of data collected each time is 2560.

In order to avoid the excessive influence caused by the error caused by the excessive number of data samples, the adjusted $R^2$ is used to estimate the correlation index, where $n$ represents the number of sample data, and $p$ represents the dimension of the data feature. The relative accuracy (RA) definition is used to measure the difference between the predicted value and the true value. The result is between 0 and 1. The closer to one, the better the result.

$$RA = 1 - \frac{\|y_i - \hat{y}_i\|}{\|y_i\|}$$

(10)

In formulas (5)–(10), $\hat{y}_i$ represents the predicted data, $y_i$ represents the original acquired data, and $n$ represents the length of the sampled data. Mean absolute error represents the error between $y_i$ and $\hat{y}_i$. RMSE reflects the error distribution, that is, the deviation between $y_i$ and $\hat{y}_i$. Mean absolute error is used to measure the quality of a model’s prediction results. The smaller the three kinds of errors, the better the model. The correlation index $R^2$ is used to describe the effect of fitting. The larger the $R^2$, the better the fitting effect of the model.

In order to avoid the excessive influence caused by the error caused by the excessive number of data samples, the adjusted $R^2$ is used to estimate the correlation index, where $n$ represents the number of sample data, and $p$ represents the dimension of the data feature. The relative accuracy (RA) definition is used to measure the difference between the predicted value and the true value. The result is between 0 and 1. The closer to one, the better the result.

$$RA = 1 - \frac{\|y_i - \hat{y}_i\|}{\|y_i\|}$$

(10)
signals of rolling bearings under three different working conditions and the corresponding root mean square values are shown in the Figure 7.

3. XJTU-SY dataset

The third data set selected for the experiment comes from XJTU-SY data set. The dataset contains the full-life cycle vibration signals of 15 rolling bearings under three working conditions. This experimental platform is shown in the figure below, which is mainly composed of AC motor, motor speed controller, rotating shaft, support bearing, hydraulic loading.

---

**Table 2. PHM bearing test condition information.**

| Condition | Rotating Speed (rpm) | Load (N) | Data file |
|-----------|----------------------|----------|-----------|
| Condition 1 | 1800                 | 4000     | 1_1, 2_1, 3_1, 4_1, 5_1, 6_1, 7 |
| Condition 2 | 1650                 | 4200     | 2_1, 2_2, 3_2, 4_2, 5_2, 6_2, 7 |
| Condition 3 | 1500                 | 4500     | 3_1, 3_2, 3_3 |

**Figure 6.** Prognostics and health management bearing data information. (a) Raw vibration signal of the prognostics and health management bearing 1_1, 2_1, 3_1. (b) RMS of the prognostics and health management bearing 1_1, 2_1, 3_1.

**Table 3. IMS bearing test bed condition information.**

| Information | First group | Second group | Third group |
|-------------|-------------|--------------|-------------|
| Bearing number | B1, B2, B3, B4 | B5, B6, B7, B8 | B9, B10, B11, B12 |
| File description | 2156 files in total | 984 files in total | 4448 files in total |
system and test bearing. The specific working condition information of the experimental platform is shown in the Table 4. In the experiment, the sampling frequency is set to 25.6 kHz, the sampling interval is 1 min, and the duration of each sampling is 1.28 s, that is, the number of sampling points each time is 32,768. The original vibration signals of rolling bearings under three different working conditions and the corresponding root mean square values are shown in the Figure 8.

The abscissa of the bearing in the three working conditions of the PHM data set represents the time point of the bearing operation; the interval between each time point is 10 s; the ordinate is the value corresponding to the RMS at different time points. The analysis shows that the degradation trend of each bearing varies with the running time, which indicates that the degradation trend of the four bearings is different. The sampling interval of the IMS data set is 10 min. The sampling interval of the XJTU-SY data set is 1 min. The sensor sampling frequency and sampling time interval of the three experimental data sets are not the same, so that the complex situation in the actual working conditions can be better reflected.

**Table 4.** XJTU-SY bearing test bed condition information.

| Condition | Rotating Speed (r/min) | Load (kN) | Data file |
|-----------|------------------------|-----------|-----------|
| Condition 1 | 2100                   | 12        | 1_1.1_2.1_3.1_4.1_5 |
| Condition 2 | 1650                   | 11        | 2_1.2_2.2_3.2_4.2_5 |
| Condition 3 | 1500                   | 10        | 3_1.3_2.3_3.3_43.5 |

**Figure 7.** Intelligent Maintenance System bearing data information. (a) Raw vibration signal of the Intelligent Maintenance System bearing 1,2,3,4. (b) RMS of the Intelligent Maintenance System bearing 1,2,3,4.

**Experiment procedure**

A CNN model is constructed to estimate the RUL. The hidden hyperparameters of CNN are robust. This CNN structure has six layers, including two convolutional layers, two maximum pooling layers, and two fully connected
layers. The specific structure of the network is shown in Figure 9. The experiment was carried out using Windows 10 system, and the central processing unit (CPU) used the 1.80 GHz i5 processor internal storage to be 8.0 GB, and the real test software was used. The version of MATLAB R2019a was used for the experiment.

In the experimental process, the data of each incoming CNN network structure is normalized to a 19-dimensional feature. In the process of training CNN, MSE is used as the loss function of the training CNN, Adam optimizer is used, and the number of iterations is 200.\textsuperscript{18,19} The experimental verification is carried out by the method of leaving one,
and the results of six evaluation indicators of the three data sets obtained by the test are shown in the Tables 5–7.

1. PHM dataset experiment results

For the PHM 2012 data set test, it can be seen that the prediction effect under the three working conditions is still relatively good, and the trend of the remaining life of the bearing can be predicted relatively well. The load of the first working condition is 4 kN, which is relatively low in the comparative experimental load, so the running time of the bearing is relatively long. The corresponding acquisition time is relatively long, and the collected training data is relatively long, so the training effect of the model is better. The load of the second working condition is 4.2 kN. As the load increases, the rotation speed decreases accordingly. The running time of the bearing decreases correspondingly, and the collected data also decreases, but the CNN model after

| Bearing number | MAE   | RMSE   | MAPE   | $R^2$   | Adjusted $R^2$ | RA    |
|----------------|-------|--------|--------|---------|----------------|-------|
| Bearing 1_1    | 0.0342| 0.0490 | 0.2870 | 0.9715  | 0.9712         | 0.9371|
| Bearing 1_2    | 0.1166| 0.1552 | 0.7263 | 0.6844  | 0.6767         | 0.8342|
| Bearing 1_3    | 0.1185| 0.1513 | 0.7245 | 0.7292  | 0.7226         | 0.8357|
| Bearing 1_4    | 0.0700| 0.0820 | 0.6247 | 0.9143  | 0.9131         | 0.8839|
| Bearing 1_5    | 0.0579| 0.0769 | 0.9348 | 0.9307  | 0.9301         | 0.9176|
| Bearing 1_6    | 0.0752| 0.0959 | 0.4596 | 0.8926  | 0.8917         | 0.8762|
| Bearing 1_7    | 0.0741| 0.0911 | 0.3243 | 0.9009  | 0.9000         | 0.8788|
| Bearing 2_1    | 0.0916| 0.1155 | 0.5263 | 0.8346  | 0.8307         | 0.8825|
| Bearing 2_2    | 0.0445| 0.0531 | 0.2053 | 0.9616  | 0.9606         | 0.9295|
| Bearing 2_3    | 0.0621| 0.0849 | 0.5510 | 0.9043  | 0.9033         | 0.9052|
| Bearing 2_4    | 0.0841| 0.1086 | 0.9404 | 0.7676  | 0.7610         | 0.9069|
| Bearing 2_5    | 0.0574| 0.0719 | 0.5187 | 0.9278  | 0.9272         | 0.9207|
| Bearing 2_6    | 0.0763| 0.1028 | 0.7806 | 0.8208  | 0.8153         | 0.8929|
| Bearing 2_7    | 0.0568| 0.0733 | 0.4893 | 0.7538  | 0.7006         | 0.8739|
| Bearing 3_1    | 0.1190| 0.1483 | 0.6758 | 0.6992  | 0.6862         | 0.8529|
| Bearing 3_2    | 0.0894| 0.1206 | 0.5213 | 0.8168  | 0.8145         | 0.8816|
| Bearing 3_3    | 0.0659| 0.0873 | 0.3272 | 0.9090  | 0.9043         | 0.8979|

| Bearing number | MAE   | RMSE   | MAPE   | $R^2$   | Adjusted $R^2$ | RA    |
|----------------|-------|--------|--------|---------|----------------|-------|
| Bearing 1      | 0.0462| 0.0691 | 0.9311 | 0.9305  | 0.9290         | 0.9316|
| Bearing 2      | 0.0980| 0.1180 | 1.3092 | 0.8561  | 0.8530         | 0.8612|
| Bearing 3      | 0.0868| 0.1071 | 0.3476 | 0.8640  | 0.8611         | 0.8609|
| Bearing 4      | 0.0677| 0.0822 | 0.6050 | 0.9141  | 0.9123         | 0.8812|

| Bearing number | MAE   | RMSE   | MAPE   | $R^2$   | Adjusted $R^2$ | RA    |
|----------------|-------|--------|--------|---------|----------------|-------|
| Bearing 1_1    | 0.0344| 0.0448 | 0.1481 | 0.9690  | 0.9676         | 0.9405|
| Bearing 1_2    | 0.0288| 0.0367 | 0.1712 | 0.9849  | 0.9844         | 0.9501|
| Bearing 1_3    | 0.0229| 0.0294 | 0.0961 | 0.9888  | 0.9884         | 0.9596|
| Bearing 1_4    | 0.0546| 0.0705 | 0.2656 | 0.9417  | 0.9391         | 0.9148|
| Bearing 1_5    | 0.0832| 0.0429 | 0.7653 | 0.9639  | 0.9630         | 0.9219|
| Bearing 2_1    | 0.0791| 0.1209 | 2.3234 | 0.8711  | 0.8697         | 0.8791|
| Bearing 2_2    | 0.0271| 0.0334 | 0.3010 | 0.9873  | 0.9869         | 0.9623|
| Bearing 2_3    | 0.0456| 0.0551 | 0.3632 | 0.9658  | 0.9655         | 0.9364|
| Bearing 2_4    | 0.0461| 0.0555 | 0.4944 | 0.9549  | 0.9542         | 0.9640|
| Bearing 2_5    | 0.0322| 0.0405 | 0.1984 | 0.9752  | 0.9748         | 0.9494|
| Bearing 3_1    | 0.0393| 0.0505 | 0.1862 | 0.9689  | 0.9688         | 0.9391|
| Bearing 3_2    | 0.0520| 0.0700 | 0.7948 | 0.9427  | 0.9426         | 0.9236|
| Bearing 3_3    | 0.0633| 0.0818 | 1.1367 | 0.8984  | 0.8970         | 0.8883|
| Bearing 3_4    | 0.0665| 0.0850 | 2.1873 | 0.8814  | 0.8810         | 0.8821|
| Bearing 3_5    | 0.0411| 0.0534 | 0.1816 | 0.9639  | 0.9621         | 0.9359|
training can still predict it well. The load of the third working condition is 5 kN. In the same period experiment, the load of the third working condition is the largest, the running time of the corresponding bearing is shorter, and the collected data is less. The number of CNN model training samples is greatly reduced, and the error of the corresponding model is also relatively large. However, due to the use of stratified sampling training ideas, the prediction effect of the model is still relatively good, and the RUL of the bearing can still be predicted relatively well (Figure 10 and 11).

2. IMS dataset experiment results

For the IMS dataset, the vibration acceleration signals of the four rolling bearings in the second set of experiments of the database selected in this paper are used as the signal data of the full life of the bearing. The fault type is manifested as an outer ring fault with a total duration of 164 h. It has a radial load of 26.69 kN and is installed on the same shaft. It maintains a constant speed of 2000 r/min. It can be seen that the test results of the IMS data set show that the effects of bearing 1, bearing 2 and bearing 4 are better. The result of bearing 3 is slightly worse, but the remaining life of the bearing can still be predicted comparatively.

3. XJTU-SY dataset experiment results

For the XJTU-SY dataset, the radial load of the five bearings in working condition one is 12 kN, which is the largest in the same period of experiment, and the corresponding speed is the lowest. Therefore, the running time of the bearing is also relatively short. The data is relatively small, but the prediction effect of our proposed method is still relatively good. The radial load of working condition 2 is 11 kN, and the speed increases as the radial load decreases. Therefore, the amount of collected data is also relatively large, and the prediction results of the model are also relatively good. The radial load of working condition 3 is 10 kN, the maximum speed is 2400 r/min, and the collected data is also the most. Therefore, the trained CNN model is better, and the final prediction result is also better.

Evaluation indicators

As a comparison, LSTM and ELM are selected for comparative experiments. Among them, the number of nerve cells in each hidden layer of LSTM is 200, and the number of hidden layers is 5. The weights of the ELM input layer and the offset of the hidden layer are randomly selected numbers. The results show that the experimental results in this article
are due to the above two methods. Two bearings are selected for each data set for comparison experiment.

In the two datasets selected for PHM, the MAE of Bearing 1_1 is 0.0342, RMSE is 0.0490, MAPE is 0.2870, $R^2$ is 0.9715, Adjusted $R^2$ is 0.9712, and RA is 0.9371. Among them, MAE, RMSE, and MAPE are all smaller than the prediction effects obtained by ELM and LSTM, and $R^2$, Adjusted $R^2$, and RA are all higher than the prediction accuracy of the above two models, indicating that the model proposed in this paper is superior to the above two models.

In the two datasets selected for IMS, the MAE of Bearing 1 is 0.0462, RMSE is 0.0691, MAPE is 0.9311, $R^2$ is 0.9305, Adjusted $R^2$ is 0.9290, and RA is 0.9316. Among them, MAE, RMSE, and MAPE are all smaller than the prediction effects obtained by ELM and LSTM, and $R^2$, Adjusted $R^2$, and RA are all higher than the prediction accuracy of the above two models.

However, the MAPE of the IMS dataset is generally a little larger, mainly because the sampling frequency of the IMS is relatively high, and the number of sampling data points is relatively large. MAPE is a cumulative sum index, so compared to other datasets, the MAPE result of the IMS dataset is a bit larger.

In the two datasets selected for XJTU-SY, the MAE of Bearing 1_1 is 0.0344, RMSE is 0.0448, MAPE is 0.1481, $R^2$ is 0.9690, Adjusted $R^2$ is 0.9676, and RA is 0.9405. Among them, MAE, RMSE, and MAPE are all smaller than the prediction effects obtained by ELM and LSTM, and $R^2$, Adjusted $R^2$, and RA are all higher than the prediction accuracy of the above two models, indicating that the model proposed in our paper is superior to the above two models.

Among the three RUL prediction methods, it is obvious that proposed method has the smallest error and best fitting effect compared with other methods.

**Conclusion**

In this paper, a new RUL prediction method based on fractal dimension and convolution neural network is proposed. First, this method introduces fractal dimension into the characterization of rolling bearing fault characteristics, and combines time domain, frequency domain, wavelet packet domain and entropy domain to construct multi-dimensional fusion features, which can reveal the degradation state of rolling bearings more comprehensively. Secondly, the percentage of remaining life of the bearing is used as a tracking indicator.
metric label, and training samples and test samples are di-
vided based on the idea of hierarchical data sampling. Finally, a
6-layer one-dimensional convolutional neural network is
constructed to fit the hidden non-linear mapping relationship
between the multi-dimensional fusion features and the per-
centage of remaining life, so as to perform RUL prediction.
Experimental verification is carried out on three common
full-life rolling bearing data sets. The results show that
compared with LSTM and ELM, the prediction effect of
rolling bearing remaining life based on fractal dimension and
CNN is improved.

In the future, we will consider the remaining life pre-
diction of different types of bearings, and combine the
currently popular methods of constructing health indexes(HI)
to predict RUL. Moreover, some methods of adaptively
extracting and extracting features should also be applicable.
Other potential degradation indicators will be attempted to be
combined with the CNN model for even higher health
prognosis accuracy.

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to
the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for
the research, authorship, and/or publication of this article: This work
was supported in part by the National Natural Science Foundation of
China under Grant 61671338 and Grant 51877161.

ORCID iDs
Guorong Ding  http://orcid.org/0000-0001-9320-0084
Wenbo Wang  http://orcid.org/0000-0002-2130-7672

References
1. Wang Y, Li S, Kang S, et al. Method of Predicting Remaining
Useful Life of Rolling Bearing Combining CNN and LSTM.
J Vibration, Measurement Diagn, 2021.
2. Pan Y, Hong R, Chen J, et al. Performance degradation as-
sessment of a wind turbine gearbox based on multi-sensor data
fusion. Mechanism Machine Theor 2019; 137: 509–526.
3. Dai Q, Chen Y and Tao R. Research on PHM model of point
machine based on dynamic PSO to optimize HSMM. Railway
Stand Des 2018; 62(09): 174–178.
4. Jun Pi, Gao S, Huang J, et al. Prediction of fatigue life of
tribe blades based on QAR data[J]. J Syst Simulation 2019;
31(06): 1165–1171.
5. Zeng D, Yang J, Zou Y, et al. Bearing life prediction method
based on parallel multi-channel convolutional long and short-
term memory network. China Mech Eng 2020; 31(20):
2454–2462+2471.
6. Xia T, Song Y, Zheng Y, et al. An ensemble framework based
on convolutional bi-directional LSTM with multiple time
windows for remaining useful life estimation. Comput Industry
2020; 115: 103182.
7. Aye SA and Heyns PS. An integrated Gaussian process re-
gression for prediction of remaining useful life of slow speed
bearings based on acoustic emission. Mech Syst Signal Process
2017; 84: 485–498.
8. Loutas TH, Roulias D and Georgoulas G. Remaining useful life
estimation in rolling bearings utilizing data-driven probabilistic
E-support vectors regression. IEEE Trans Reliability 2013;
62(4): 821–832.
9. Zhang C. Bearing life prediction of subway fan based on EMD
noise reduction and LSTM network. Fan Technol 2020; 62(3):
77–82.
10. LeCun Y, Bengio Y and Hinton G. Deep learning. Nature 2015;
521(7553): 436–444.
11. Jun Z, Nan C and Weiwen P. Estimation of bearing remaining
useful life based on multiscale convolutional neural network.
IEEE Transactions on Industrial Electronics 2018; 66: 1–11.
12. Kang S, Zhou Y, Wang Y, et al. Rolling bearing RUL prediction
method based on improved SAE and bidirectional LSTM. Acta
Automatica Sinica 2020; 11: 1–11.
13. Tian Z. An artificial neural network method for remaining
useful life prediction of equipment subject to condition
monitoring. J Intell Manufacturing 2012; 23: 227–237.
14. Sun C, Zhang Z and He Z. Research on bearing life prediction
based on support vector machine and its application. J Phys
Conf Ser 2011; 305(1): 012028.
15. Ren L, Cui J, Sun Y, et al. Sun Yaqiang, etMulti-bearing
remaining useful life collaborative prediction: A deep learning
approach. J Manufacturing Syst 2017; 43: 248–256.
16. Cao Z and Ye C. Bearing remaining service life prediction
based on parallel CNN-SE-Bi-LSTM. Computer Application
Research 2021; 99: 1–6. DOI: 10.19734/j.issn.1001-3695.
2020.08.0224.
17. Han L, Shi C and Zhang J. Bearing remaining life prediction
based on one-dimensional convolutional neural network.
Manufacturing Automation 2020; 42(03): 10–13.
18. Ding H, Yang L, Cheng Z, et al. A remaining useful life
prediction method based on deep neural networks. Meas
2021; 172: 108878.
19. Yoo Y and Baek J-G. A novel image feature for the remaining
useful lifetime prediction of bearings based on continuous
wavelet transform and convolutional neural network. Appl Sci
2018; 8: 1102.
20. Guo L, Li N, Jia F, et al. A recurrent neural network based
health indicator for remaining useful life prediction of bearings.
Neurocomputing 2017; 248: 98–109.
21. Ding H, Yang L and Yang Z. A predictive maintenance method
for shearer key parts based on qualitative and quantitative
analysis of monitoring data. IEEE Access 2019; 7:
108684–108702.
22. Li H and Chen T. The Local Demagnetization Fault Diagnosis
of PMSM Based on Fractal Dimension TRANSACTIONS OF
CHINA ELECTECHNICAL SOCIETY, 2017.
23. Qi Y, Fan J, Liu L, et al. Fault diagnosis of wind turbine bearing
based on morphological fractal and extreme learning machine.
J Solar Energy 2020; 41(06): 102–112.
24. Wojciech M, Ricardo B, Mateusz K, et al. Fractal dimension
bending–torsion fatigue fracture characterization. Meas 2021;
184: 109910.
25. Hu Y, Shen J and Qi Y. Estimation of rice biomass at different
growth stages by using fractal dimension in image processing.
Appl Sci 2021; 11(15): 7151.
26. Pan G, Li S, An Z and Zeng Y. S-transform method of
synchronous extrusion for extracting impact features of
rolling bearings[J]. Chinese Journal of Vibration
27. Gouriveau NZR, Javed K, Medjaher K, Nectoux P, Ramasso E, et al. IEEE PHM 2012 Data challenge The PRONOSTIA platform, 2012. https://github.com/wkzs111/phm-ieee-2012-data-challenge-dataset

28. IMS bearings data set (2014). http://ti.arc.nasa.gov/tech/dash/pcoe/prognostic-data-repository/.

29. Wang B, Lei Y, Li N, et al. A hybrid prognostics approach for estimating remaining useful life of rolling element bearings. *IEEE Trans Rel* 2020; 69(1): 401–412.