International Journal of Electrical, Electronic and Communication Sciences
ISSN: 2517-9438
Vol:13, No:7, 2019

Remaining Useful Life Estimation of Bearings Based on Nonlinear Dimensional Reduction Combined with Timing Signals

Zhongmin Wang, Wudong Fan, Hengshan Zhang, Yimin Zhou

Abstract—In data-driven prognostic methods, the prediction accuracy of the estimation for remaining useful life of bearings mainly depends on the performance of health indicators, which are usually fused some statistical features extracted from vibrating signals. However, the existing health indicators have the following two drawbacks: (1) The differnet ranges of the statistical features have the different contributions to construct the health indicators, the expert knowledge is required to extract the features. (2) When convolutional neural networks are utilized to tackle time-frequency features of signals, the time-series of signals are not considered. To overcome these drawbacks, in this study, the method combining convolutional neural network with gated recurrent unit is proposed to extract the time-frequency image features. The extracted features are utilized to construct health indicator and predict remaining useful life of bearings. First, original signals are converted into time-frequency images by using continuous wavelet transform so as to form the original feature sets. Second, with convolutional and pooling layers of convolutional neural networks, the most sensitive features of time-frequency images are selected from the original feature sets. Finally, these selected features are fed into the gated recurrent unit to construct the health indicator. The results state that the proposed method shows the enhance performance than the related studies which have used the same bearing dataset provided by PRONOSTIA.

Keywords—continuous wavelet transform, convolution neural network, gated recurrent unit, health indicators, remaining useful life.

I. Introduction

IN the complex and automated modern industry, rolling element bearings are the ones of the most critical components in rotating machinery. Any unexpected failure of bearings may result in undesired consequences, such as downtime increasing, productivity reduction or safety risks [1], [2], [3], [4], [5]. To solve these problems, the Remaining Useful Life (RUL) prediction is required to schedule the future action to avoid the catastrophic occurrence [6].

Recently, many RUL prediction methods have been proposed for Prognostic and Health Management (PHM). These methods can be classified as model-based [7] and data-driven [8]. The methods of model-based tend to be more accurate if the complex system degradation is modeled precisely [9], however, these methods require extensive expert knowledge about physical systems, which are usually unavailable. The data-driven methods are able to model the degradation characteristics based on the historical sensor data. The underlying correlations and causalities in the collected sensor data can be revealed, and the corresponding system information such as RUL can be inferred. Many data-driven methods have been proposed recently with satisfied prognostic results achievement. Usually, the data-driven methods are mostly dependent on the signal processing and feature extraction.

The framework of the data-driven methods generally comprises three steps: (1) data acquisition, (2) construction of Health Indicator (HI), and (3) prognostics. The main operation of the data acquisition is to use sensors to record the original signal at a certain sampling rate. The performance of the HI has a crucial effect on the accuracy of RUL [10]. Hong et al. [11] extract time-frequency features using wavelet packet and empirical mode decomposition techniques. Based on the extracted features, self-organizing maps is used to construct the HI confidence value.

The Weighted Minimum Quantization Error (WMQE) is used as a HI, which is a fusion of 10 time domain features and 16 time-frequency domain features, where the two features are extracted trigonometric functions [12]. In all, most of these studies are relied on feature extraction, selection, and fusion techniques to construct the HI, however the processes are quite complex where are labor-intensive and rely on the expert knowledge of the failure mechanism. The method of manually extracting features based on expert knowledge is not suitable for most applied scenarios.

In order to solve these drawbacks, Guo et al. [13] extract 11 time domain features, 5 frequency domain features, and 8 time-frequency domain features. where the most sensitive features have been selected using monotonicity and correlation metrics. The selected features are fed into Recurrent Neural Network (RNN) architecture to construct HI. Although, the proposed method can improve the predict on performance by integrating the deep learning, expert knowledge or feature engineering is still required.

The original signal features of the time-frequency images are extracted and RUL is predicted by transforming high-dimensional features into low-dimension using Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) [14]. The time-frequency domain features of the original signals are extracted directly by using S transform, in order to preserve the main time-frequency features of

Zhongmin Wang, Wudong Fan and Hengshan Zhang are with the Shanxi Key Laboratory of Network Data Analysis and Intelligent Processing, School of Computer Science and Technology, Xian University of Posts and Telecommunications, Xian, China (e-mail: zmwang@xupt.edu.cn, zhanghs@xupt.edu.cn).

Yimin Zhou is with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China (e-mail: ym.zhou@siat.ac.cn).

the original signals. This method does not require expert knowledge to select different vibrating signal features. However, the PCA is used to extract features of the time-frequency images, where the extracted time-frequency features may contain a large number of ones that are not related to HI.

In order to solve the defects of the above mentioned methods, the method integrating Convolutional Neural Networks (CNN) and Gated Recurrent Unit (GRU) networks based on the time-frequency image features is proposed in this study, it adopts the Continuous Wavelet Transform (CWT) for HI construction and RUL prediction. The time-frequency images can visually show that the varied frequency components. In addition, it requires no experts for fault diagnosis or signal processing [15]. In the proposed method, CWT is used to extracts the time-frequency features from one-dimensional vibration signals, and then the time-frequency features are converted to time-frequency images. CNN requires a large amount of training data and hyper parameter adjustments, and the state of art studies have shown that it performs well in image analysis [16]. The bearing vibration signals are time-series data, so the CNN does not have a mechanism to process them. GRU is a the more effective model involving time-series data. The outputs of the CNN are fed into the GRU network in order to fully consider the timing-related features for the time-frequency signal. The two networks are integrated to process images with time-series and time-frequency, because it can improve the accuracy of HI prediction. An exponential model is used to predict the RUL of bearing by nonlinear regression.

The remainder of this paper is organized as follows. Section II discusses the related work. Section III presents the proposed method. Section IV conducts the tests of the proposed method on the PRONOSTIA dataset. The conclusions and future work are given in Section V.

II. The Related Works

This section summarizes the methods of feature extraction and selection as well as RUL prediction methods. The time-frequency feature extraction of CWT method and the feature selection of CNN and GRU networks are introduced in detail.

A. Continuous Wavelet Transform (CWT)

Wavelet analysis has been developed rapidly due to the basic works of Y. Meyer et al [17]. A wavelet discussed generally refers to a substrate of ψ or other space produced by a function of a mother wavelet and a base wavelet that is stretched and translated. The defined function $\psi(t)$ satisfies, $\int_{-\infty}^{\infty}\psi(t)dt = 0$ where $\psi(t) \in L^2$ is a mother wavelet or is called a wavelet function.

The family of time-scale waveforms is obtained by shifting and scaling the mother wavelet,

$$\psi_{a,b}(t) = |a|^{-\frac{1}{2}}\psi\left(\frac{t-b}{a}\right) \tag{1}$$

where a is the scale parameter for dilating or contracting the wavelet, b is the shifting parameter for transitioning the wavelet along the time axis, $\psi_{a,b}(t)$ is a continuous wavelet basis function.

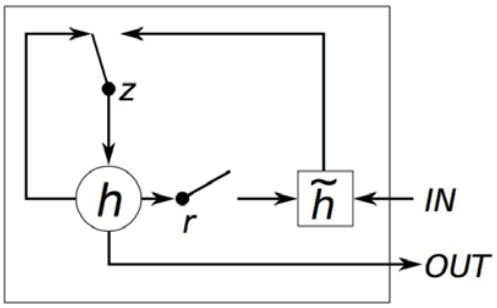


Fig. 1 Gated Recurrent Unit

For a given signal, $x(t)$, wavelet coefficient $wt(a,b)$ can be expressed,

$$wt(a,b) = |a|^{-\frac{1}{2}}\int_{-\infty}^{\infty} x(t)\,\psi^*\left(\frac{t-b}{a}\right)dt \tag{2}$$

where ψ^* represents the complex conjugate of ψ.

B. Basic Theory of CNN

CNN is a deep learning model inspired by the visual cortex of the brain. It is widely used in many applications, especially in image feature extraction such as computer vision because it can analyze the original image directly without complicated preprocessing. CNN can extract features of the image while performing dimensionality reduction. The CNN model alternately stacks the convolutional layer and the pooling layer for feature selection and dimensionality reduction. The fully connected layer is used to output the extracted image features. The detailed procedure of CNN can be referred [18].

C. The Basic Theory of GRU

As shown in Fig. 1, GRU [19] is a kind of RNN. Traditional neural networks such as CNN is not good at processing time-series information, and GRU, as a variant of RNN, can combine historical and the current time information predict the further information. The GRU network integrates the forgetting gate and input gate of the Long Short Term Memory (LSTM) network into a single update gate, and it can simplify the connection structure between neuron, reduces the training parameters and improve the training efficiency [20].

The input of the GRU is IN , z represent the update gate and r represent the reset gate. z and r are used to control the direction of the data stream at time t.

D. Double Exponential Model

The double exponential model has proven to be an effective model for curve fitting and prediction [21], described as,

$$Y = ae^{bt} + ce^{dt} \tag{3}$$

where Y is the degraded state value of the bearing, t is the current time moment, and a, b, c, and d are the parameters of the model.

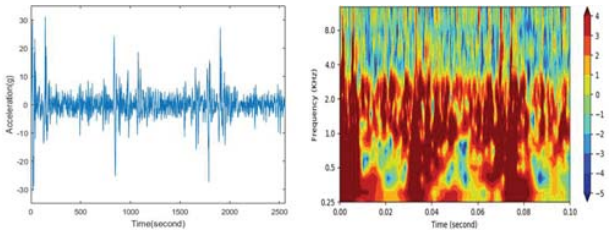
III. The Proposed CCG-RUL Method for Prediction RUL

This section describes the procedure for construction of the proposed CCG-RUL method. It is mainly composed of three steps. Step 1, the original vibration signals are converted to time-frequency images. Step 2, the CNN convolution and pool layers are used to reduce the dimension of the time-frequency images, and GRU layers are used to select the output features of the CNN to construct HI. Step 3 is to predict the RUL of the bearing's RUL using a double exponential model.

A. Signal-to-Image Conversion Method

In the traditional data-driven prediction of RUL, the data preprocessing method is crucial. The main function of the data preprocessing is to extract the features of the original signals from a large volume of the historical data. However, extracting the correct features is an energy-intensive task and has a huge impact on the final prediction [16].Here CWT is used to extract the time-frequency domain features of the original signals.

The complex morlet wavelet is a cosine signal with a squared exponential decay, which is similar to the most of the engineering machineries as transient shock signals. According to the maximum matching principle of the wavelet analysis, when the wavelet is more similar to the analyzed signal in geometry, the signal features extracted by the wavelet are more accurate [22]. As a result, this paper uses complex morlet wavelet to extract the time-frequency features of the original signal. Cmor1-1.5 is chosen as the mother wavelet function. As shown in Fig. 2, the original vibration signals of the Bearing1_1 at 28000 s is converted into a time-frequency image, where Fig. 2(a) presents the original vibration signals and Fig. 2(b) shows the transformed time-frequency image.



(a) original vibration signal (b) time-frequency image

Fig. 2 Signal-to-image of Bearing1_1 at 28000 s

B. The Proposed Model with CNN and GRU for HI

After constructing the time-frequency images of the original vibration signals, the network combined CNN with GRU is used to construct the HI of bearing. The proposed CCG-HI model is shown in Fig. 3. In this model, the input data is prepared by converting the CWT coefficients into the time-frequency images. The five convolution and pooling layers are alternately used to extract the features and reduce the dimension of the time-frequency images. Next, reshape the feature vectors outputting from the last pooling layer into two-dimensional data and feed it into the GRU layer. Three GRU and one-dimensional pool layers are used to extract the features. After the GRU layers, the time-frequency image features are flattened and connected to the fully connected layer. Finally, an output neuron is used to connect the last fully connected layer to build CCG-HI. In this paper, in order to limit the output to $[0, 1]$, a sigmoid function is chosen as the activation function of the output layers. The Rectified Linear unit (ReLu) function is used as the activation function in the convolution layer and the fully connected layer, in order to prevent network overfitting and gradient disappearance or explosion. The mean square error is chosen as the loss function.

C. RUL Prediction

The HI predicted by the CCG-HI model are fed into the double exponential model for training the related parameters. The time interval is sent to the trained double exponential model to obtain the HIs value. When HI is equal to 1, time t is the end of the bearing life, and calculated as,

$$RUL = t_{final} - t_{current} \tag{4}$$

where t_{final} is the time for bearing to reach the failure threshold, $t_{current}$ is the time of the last vibration signal collected in the bearing dataset.

IV. Experiments and Discussions

In this section, the experiments are conducted to show the working and efficacies of proposed CCG-RUL. The famous bearing degradation dataset PRONOSTIA is utilized.

A. Description of The Experimental Platform PRONOSTIA

The experimental dataset was collected on the PRONOSTIA platform, provided by the FEMTO-ST Institute [23]. PRONOSTIA is an experimental platform dedicated to test, verify, and validate methods related to bearing health assessment, diagnostics, and prognostics. A general overview of the platform is shown in Fig. 4. The main purpose of PRONOSTIA is to provide real data related to bearing degradations [24]. This dataset was used in the IEEE PHM 2012 Data Challenge for predicting the RUL of bearings.

The experimental dataset from PRONOSTIA was collected by conducting accelerated degradation tests of bearings. Two accelerometers are horizontally and vertically mounted on the bearing to monitor its vibration. Vibration signals collected by the two accelerometers are sampled each 10 s, and the duration of the sampling lasts 0.1 s with a sampling frequency 25.6 kHz. To avoid unnecessary damage to the test rig (for security as well), experiments finished once the amplitude of the monitoring data passes 20 g.

In the experiments, the dataset involves three different operating conditions. In condition I, there are seven bearings tested at a rotating speed of 1800 rmp under a radial load of 4000 N radial load. In condition II, seven bearings are operated at a rotating speed of 1650 rmp under a load of 4200 N. In condition III, three bearings are operated at a rotating speed of 1500 rmp under the load of 5000 N. These conditions are

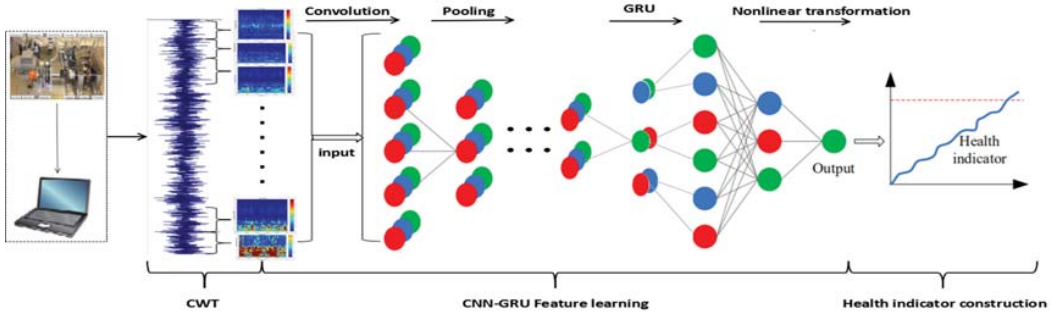


Fig. 3 The proposed CCG-HI model

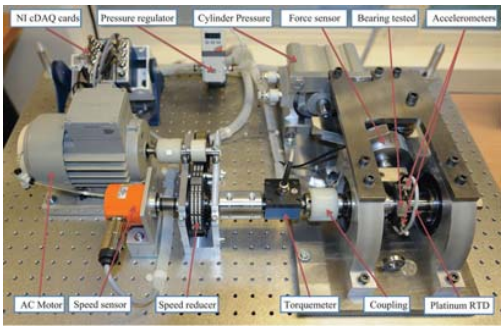


Fig. 4 The experimental platform

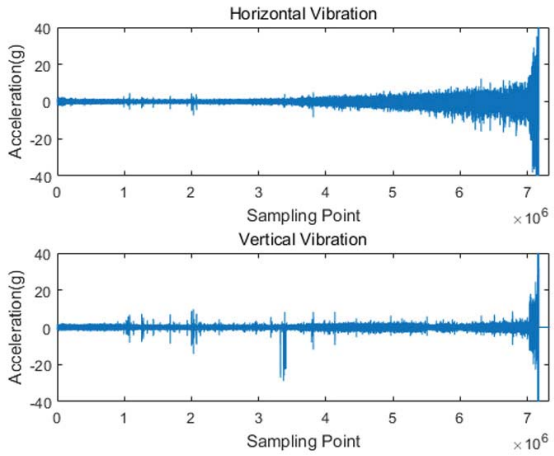


Fig. 5 The temporal vibration signal of bearing1_1

shown in Table I. The first two bearings in each group are regarded as a training set and the others are used as the testing set.

TABLE I
BRIEF INTRODUCTION ABOUT THE EXPERIMENTAL DATA

Condition	Load(N)	Speed(rmp)	Experimental Data
1	4000	1800	Bearing1_1 Bearing1_2 Bearing1_3 Bearing1_4 Bearing1_5 Bearing1_6 Bearing1_7
2	4200	1650	Bearing2_1 Bearing2_2 Bearing2_3 Bearing2_4 Bearing2_5 Bearing2_6 Bearing2_7
3	5000	1500	Bearing3_1 Bearing3_2 Bearing3_3

B. The Conversion of The Vibration Signal to The Images

Fig. 5 shows the horizontal and vertical vibration signals of the Bearing1_1 during its whole life cycle.

The complex wavelet (cmor1.0-1.5) based on CWT, is used to extract the time-frequency features of the original vibration signals to construct the time-frequency image. The data is collected in each 0.1 seconds to draw time-frequency images. For example, there are 2803 samples in the bearing1_1 training dataset, and it can draw 2803 time-frequency images.

Fig. 6 demonstrates, the time-frequency images of bearing1_1 during the run-to-failure experiment of the training bearing, which are the time-frequency distributions of the vibration signals. In the time-frequency images, the horizontal axis and vertical axis represent time and frequency, respectively. The color of each point indicates the magnitude of the wavelet coefficients on the time-frequency grid. The red color represents that the energy level is high. When the bearing is operating in normal condition, the most of the energy is concentrated around 4 kHz. On the other hand, energy bursts in the frequency range from 0.25 kHz to 8 kHz. The impacts occur in the frequency between 0.25 kHz and 1 kHz at regular interval. When high energy is observed in the low frequency band, it indicates that the bearing is heading to the defection [25]. From Fig. 6, it is easily obtained that the time-frequency images of different time periods are significantly different, so the time-frequency images can be used to predict the RUL of the bearing. It can be seen from Fig. 5 and Fig. 6 that the horizontal vibration signal is better than the vertical vibration signal to illustrate the wear state of the bearing. Thus, the time-frequency image of the horizontal vibration signal is selected to represent the bearing degradation process.

Fig. 7 depicts image features of bearing1_1 drawn by Zhao et al [15]. It can be seen from Fig. 6 and Fig. 7 that the time frequency image generated by CWT method is superior to the time frequency image generated by S transform in expressing the original signal features.

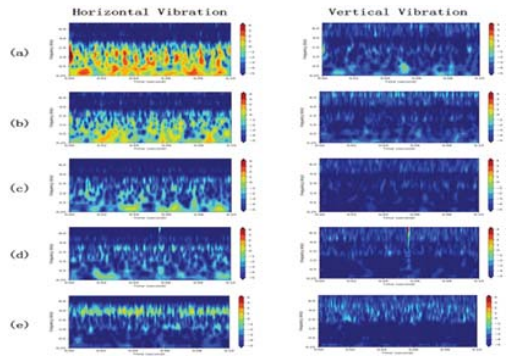


Fig. 6 CWT of of bearing1_1, where (a)-(e) refer to the 2000^{th}, 1500^{th}, 1000^{th}, 500^{th} and 1^{th}

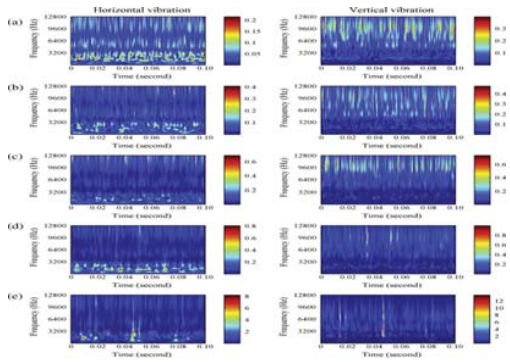


Fig. 7 TFRs of bearing1_1, where (a)-(e) refer to the 2000^{th}, 1500^{th}, 1000^{th}, 500^{th} and 1^{th}

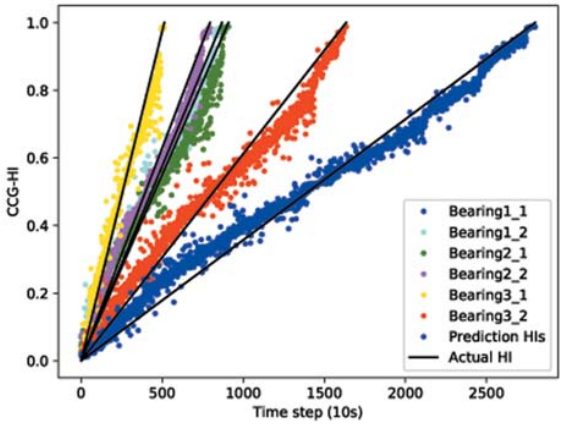


Fig. 8 CCG-HI of the six training datasets

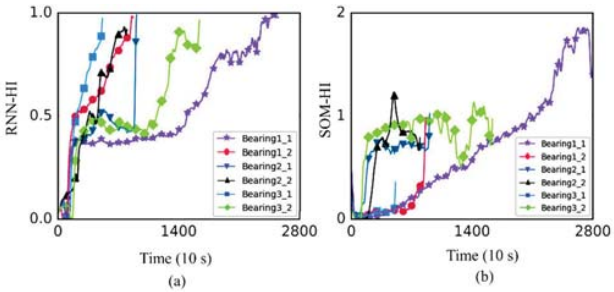


Fig. 9 HIs of the six training bearings: (a) RNN-HI (b) SOM-HI

C. CCG-HI Construction

The construction of CCG-HI consists of two steps: training and testing.

- During the training step, six bearings are used as a training set to construct the model for CCG-HI estimation. The lifetime samples of the bearings are used to form a training set, $Datatrain = \{x_t, y_t\}_{t=1}^{T}$, where $x_t \in R^{N\times N}$ is $N \times N$ image features extracted at time t, and $y_t \in [0, 1]$ is its associated label which indicates the degradation percentage of bearings at time t. In this study, the size of the input image feature is 128×128. The degradation percentage is calculated as $y_t = \frac{t}{T}$, where t is the operation time, and T is the failure time. For example, suppose that the failure time of a bearing is 2800 s, and the current inspection point is 1400 s, and the label y_t is 0.5.
- In the testing step, The testing dataseta are denoted as $Datatest = \{x_t, y_t\}_{t=1}^{T}$ where $x_t \in R^{N\times N}$ is $N \times N$ image features extracted from the testing datasets. The selected features of the testing data are directly input to the trained CNN-GRU to obtain the CCG-HI.

In the training step, 7534 feature of images are extracted from the six training bearings to construct the CCG-HI model. The training results are shown in Fig. 8. The estimated HI is marked as the colored points and the black line represents the actual HI of the training bearing.

From Fig. 8, it can be easily obtained that the estimated HI values of the six bearings are quite close to the actual HI. Some training errors are acceptable because most estimations and actual values tend to match each other. It implies that the CWT based image features can extract high-quality information about the state of the bearing. Thus, CNN-GRU is an effective model to estimate the degradation of bearings by images analyzation.

The RNN-HI curves of the six training bearings are shown in Fig. 9(a). A self-organizing map based HI (SOM-HI) [26] is shown in Fig. 9(b).

The HI curves predicted by both methods have the same disadvantages. In the RNN-HI method, the HI curves of Bearing1_1, Bearing2_1 and Bearing3_2 have a gradual trend in some stages, and even the HI curve of Bearing2_1 and Bearing3_2 appear to decrease at a certain time. The SOM-HI method also has similar problems. The SOM-HI ranges from 0.54 to 1.75 for the six training bearings at the failure condition, which could causes that the FT is difficult to determine. In the proposed CCG-HI model, the HI curves are approximately proportional to t. The bearing FT is controlled in $[0, 1]$. As a result, the proposed method is superior to the RNN-HI method and SOM-HI method.

D. RUL Estimation

In this paper, the double exponential model is used to predict the RUL. The estimated and predicted CCG-HI of bearing1_3, one of the testing datasets, are shown in Fig. 10. The obtained CCG-HI up to the current time is indicated by dot. Purple

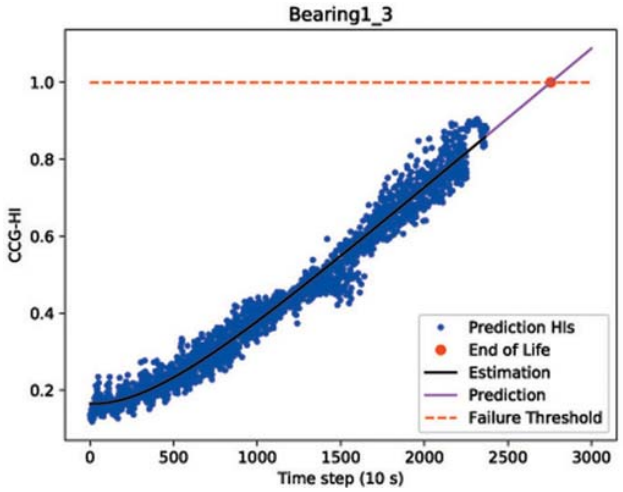


Fig. 10 RUL prediction result of bearing1_3

solid line represents the RUL of bearing. The black solid line represents the HI curve predicted using the double exponential model. It identifies that the RUL of the bearing gradually degrades over time with the increase of HI value.

E. Performance of The Proposed Method

A percent error of prediction results that for the IEEE PHM 2012 Prognostic Challenge is applied to evaluate the performance of the prediction methods. It is defined as follows:

$$Er_i = \frac{ActRUL_i - \widehat{RUL}_i}{ActRUL_i} \times 100\% \quad (5)$$

where $ActRUL_i$ and $\widehat{RUL}_i$ are the actual RUL and predicted one of the i^{th} testing data, respectively.

To verify the effectiveness of the proposed method, the prediction results are compared with the other five similar studies using the same dataset. The prediction results for the rest life of the bearings are shown in Table II.

For the method proposed in [28], the disadvantage is that the definition for the anomaly detection time is based on the subjective criteria which is used to calculate the bearing survival time ratios. Furthermore, the calculated percentage error of this method is quite different from the actual result. The method constructed in [11] shows the errors reduction compared with the previous studies, but it requires the extraction of the approximately one hundred features to estimate the bearing performance. The study published in [12] develops a new HI (i.e, WMQE) to predict the RUL of the bearings, which is constructed by fusing a few selected weighted features based on correlation clustering among the 28 features which is extracted from the bearings. The study in [12] shows the best performance among existing studies. The method of RNN-HI presented in [13] is also constructed by selecting and fusing multiple features which are extracted from the time, frequency, and time-frequency domains. The method presented in [13] demonstrates its superiority over SOM-based HIs. The methods of RNN-HI and SOM-HI require the expert's experiences to manually select a large number of features. The study published in [27] proposes an end-to-end deep framework for RUL estimation based on the convolutional and long-short-term memory recurrent units, which uses the convolutional layer of the neural network to extract the local features directly from the sensor data, next the LSTM layer is introduced to capture the degradation process, finally the RUL is estimated using the LSTM outputs and the prediction time value. The methods developed in [12] and [13] require the expert's knowledge to extract the large number of features of the raw data. The features of the raw signal are extracted from the convolution layer in the method constructed in [27], its accuracy is lower than our method.

The errors of the proposed method is 18.51, which is the lowest among the experiments which indicates that the model can work accurately and reliably on each tested bearing.

V. Conclusion and Future Work

RUL prediction accuracy highly depends on the performance of the HI. In this paper, the CCG-HI is proposed to enhance RUL prediction accuracy of the bearings. During the construction procedure of the HI, the vibration signals are converted into the time-frequency images instead of the original vibration signals or extracting multiple statistical features. CNN-GRU network can extracted features from the images and construct the HI without expert prior knowledge and signal processing, considering the timing correlation between vibration signals. The CNN-GRU model is used as a regression model to estimate the HI between 0 and 1. The results demonstrate that the proposed method outperforms the other methods, which is accurate and effective in the RUL prediction. However, there are still some drawbacks in this method. Here, we only consider a vibration signal and it is possible to ignore some important information when predicting RUL, which would reduce the accuracy of RUL predictions. In the future, the HI based on multisensory signals will be studied by considering temperature and two vibration sensors data.

Acknowledgment

This work is sponsored by National Science Foundations of China under Grant No. 61702414 and No. 61373116, Shaanxi Science and Technology Co-ordination and Innovation Project (Program No.2016KTZDGY04-01), Postgraduate Innovation Foundation of Xi'an University of posts & Telecommunications Under Grant No. CXJJ2017066.

TABLE II
RUL PREDICTION RESULTS

Testing Dataset	Current Time	Actually RUL	Predict RUL	Hinchi et al.[27]	Guo et al.[13]	Lei et al.[12]	Hong et al.[11]	Sutrisno et al. [28]	CCG-HI
Bearing1_3	18010 s	5730 s	4293 s	54.73%	43.28%	−0.35%	−31.76%	37%	25.09%
Bearing1_4	11380 s	339 s	394 s	38.69%	67.55%	5.60%	62.76%	80%	−16.22%
Bearing1_5	23010 s	1610 s	1363 s	−99.4%	−22.98%	100.00%	−136.03	9%	15.34%
Bearing1_6	23010 s	1460 s	1076 s	−120.07%	21.23%	28.08%	−32.88%	−5%	26.30%
Bearing1_7	15010 s	7570 s	8076 s	70.65%	17.83%	−19.55%	−11.09%	−2%	−6.68%
Bearing2_3	12010 s	7530 s	5178 s	75.53%	37.84%	−20.19%	44.22%	64%	31.23%
Bearing2_4	6110 s	1390 s	1590	19.81%	−19.42%	8.63%	−55.40%	10%	−14.39%
Bearing2_5	20010s	3090s	1803 s	8.2%	54.37%	23.30%	68.61%	−440%	41.65%
Bearing2_6	5710s	1290s	1435 s	17.87%	−13.95%	58.91%	−51.94%	49%	−11.24%
Bearing2_7	1710s	580s	508 s	1.69%	−55.17%	5.17%	−68.97%	−317%	12.41%
Bearing3_1	3510s	820s	795 s	2.93%	3.66%	40.24%	−21.96%	90%	3.05%
Mean of Er				45.87%	32.48%	28.18%	53.24%	100.27%	18.51%

REFERENCES

[1] F. Jia, Y. Lei, J. Lin, X. Zhou, and N. Lu, "Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data," *Mechanical Systems and Signal Processing*, vol. 72, pp. 303–315, 2016.

[2] L. Jie, W. Wang, and F. Golnaraghi, "An enhanced diagnostic scheme for bearing condition monitoring," *IEEE Transactions on Instrumentation & Measurement*, vol. 59, no. 2, pp. 309–321, 2010.

[3] A. Ghods and H. H. Lee, "Probabilistic frequency-domain discrete wavelet transform for better detection of bearing faults in induction motors," *Neurocomputing*, vol. 188, pp. 206–216, 2016.

[4] Y. Lei, L. Jing, M. J. Zuo, and Z. He, "Condition monitoring and fault diagnosis of planetary gearboxes: A review," *Measurement*, vol. 48, no. 1, pp. 292–305, 2014.

[5] W. Yi, G. Xu, Q. Zhang, L. Dan, and K. Jiang, "Rotating speed isolation and its application to rolling element bearing fault diagnosis under large speed variation conditions," *Journal of Sound & Vibration*, vol. 348, no. Switzerland, pp. 381–396, 2015.

[6] D. Wang and C. Shen, "An equivalent cyclic energy indicator for bearing performance degradation assessment," *Journal of Vibration and Control*, vol. 22, no. 10, pp. 2380–2388, 2016.

[7] M. Pecht and J. Gu, "Physics-of-failure-based prognostics for electronic products," *Transactions of the Institute of Measurement and Control*, vol. 31, no. 3-4, pp. 309–322, 2009.

[8] F. O. Heimes, "Recurrent neural networks for remaining useful life estimation," in *Prognostics and Health Management, 2008. PHM 2008. International Conference on*. IEEE, 2008, pp. 1–6.

[9] Y. Qian, R. Yan, and R. X. Gao, "A multi-time scale approach to remaining useful life prediction in rolling bearing," *Mechanical Systems & Signal Processing*, vol. 83, pp. 549–567, 2017.

[10] K. Javed, "A robust & reliable data-driven prognostics approach based on extreme learning machine and fuzzy clustering." Ph.D. dissertation, Université de Franche-Comté, 2014.

[11] S. Hong, Z. Zhou, E. Zio, and K. Hong, "Condition assessment for the performance degradation of bearing based on a combinatorial feature extraction method," *Digital Signal Processing*, vol. 27, no. 1, pp. 159–166, 2014.

[12] Y. Lei, N. Li, S. Gontarz, L. Jing, S. Radkowski, and J. Dybala, "A model-based method for remaining useful life prediction of machinery," *IEEE Transactions on Reliability*, vol. 65, no. 3, pp. 1314–1326, 2016.

[13] G. Liang, N. Li, J. Feng, Y. Lei, and L. Jing, "A recurrent neural network based health indicator for remaining useful life prediction of bearings," *Neurocomputing*, vol. 240, no. C, pp. 98–109, 2017.

[14] G. S. Babu, P. Zhao, and X. L. Li, "Deep convolutional neural network based regression approach for estimation of remaining useful life," in *International Conference on Database Systems for Advanced Applications*, 2016.

[15] M. Zhao, B. Tang, and T. Qian, "Bearing remaining useful life estimation based on time frequency representation and supervised dimensionality reduction," *Measurement*, vol. 86, pp. 41–55, 2016.

[16] W. Long, X. Li, G. Liang, and Y. Zhang, "A new convolutional neural network-based data-driven fault diagnosis method," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 7, pp. 5990–5998, 2018.

[17] Y. Meyer, "L es ondelettes: Algorithmes et applications. armand c olin, paris, 1992," *Engl. Transl. Wavelets, Algorithms and Applications, SIAM, Philadelphia, PA*, 1993.

[18] J. Bouvrie, "Notes on convolutional neural networks," 2006.

[19] R. Jozefowicz, W. Zaremba, and I. Sutskever, "An empirical exploration of recurrent network architectures," in *International Conference on International Conference on Machine Learning*, 2015.

[20] K. Cho, B. V. Merrienboer, C. Gulcehre, F. Bougares, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," *Computer Science*, 2014.

[21] X. Jin, S. Yi, Z. Que, W. Yu, and T. W. S. Chow, "Anomaly detection and fault prognosis for bearings," *IEEE Transactions on Instrumentation & Measurement*, vol. 65, no. 9, pp. 2046–2054, 2016.

[22] D. Eberhard and E. Voges, "Digital single sideband detection for interferometric sensors," in *International Conference on Optical Fiber Sensors*, 1984.

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

[24] T. Benkedjouh, K. Medjaher, N. Zerhouni, and S. Rechak, "Remaining useful life estimation based on nonlinear feature reduction and support vector regression," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 7, pp. 1751–1760, 2013.

[25] Z. K. Peng and F. L. Chu, "Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography," *Mechanical Systems & Signal Processing*, vol. 18, no. 2, pp. 199–221, 2004.

[26] P. Sarlin, "Self-organizing time map: An abstraction of temporal multivariate patterns," *Neurocomputing*, vol. 99, no. 1, pp. 496–508, 2013.

[27] A. Z. Hinchi and M. Tkiouat, "Rolling element bearing remaining useful life estimation based on a convolutional long-short-term memory network," *Procedia Computer Science*, vol. 127, pp. 123–132, 2018.

[28] E. Sutrisno, H. Oh, A. S. S. Vasan, and M. Pecht, "Estimation of remaining useful life of ball bearings using data driven methodologies," in *IEEE Conference on Prognostics & Health Management*, 2012.

Zhongmin Wang Zhongmin Wang was born in 1967, PhD. He is currently a professor with the School of Computer Science, Xian University of Posts & Telecommunications, and senior member of China Computer Federation. His main research interests include embedded intelligent perception, big data technology and application, intelligent information processing, etc.

Wudong Fan Wudog Fan was born in 1993, master student. His main research interests include deep learning and remaining useful life prediction, etc.

Hengshan Zhang Hengshan Zhang was born in 1969, Ph.D. He is currently an Lecturer with the School of Computer Science, Xian University of Posts & Telecommunications. His research interests include wisdom of crowds for decision making, machine learning and data mining, and information aggregation.