A Novel Feature-Transferred Prediction Network For Remaining Useful Life of Rotating Machinery

Tianjiao Lin (✉ lintj@mail.buct.edu.cn)

Beijing University of Chemical Technology

Huaqing Wang

Beijing University of Chemical Technology

Liuyang Song

Beijing University of Chemical Technology

Lingli Cui

Beijing University of Chemical Technology

Research Article

Keywords: Bearings, Remaining Useful Life(RUL), Feature Transferred Learning, Convolutional Network (CNN), Gate Recurrent Unit (GRU)

Posted Date: October 1st, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-917030/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

A Novel Feature-transferred Prediction Network for Remaining Useful Life of Rotating Machinery

Tianjiao Lin¹, Huaqing Wang¹, Liuyang Song^{1*} and Lingli Cui²

^{1*}College of Mechanical and Electrical Engineering, Beijing University of Chemical Technology, Chao Yang District, Beijing, 100029, China.

²Key Laboratory of Advanced Manufacturing Technology, Beijing University of Technology, Chao Yang District, Beijing, 100124, China.

*Corresponding author(s). E-mail(s): xq_0703@163.com;
Contributing authors: lintj@mail.buct.edu.cn;
hqwang@mail.buct.edu.cn; acuilingli@163.com;

Abstract

With the increasing complexity of CNC machine tools and other rotating machinery, it becomes more and more important to improve the reliability of such machines. It is necessary to estimate the remaining Useful Life (RUL) of the important parts such as bearings of these devices. However, the operating conditions of such parts are often very complicated, and there is a great difference between different devices. Therefore, it is difficult to use the traditional mechanism analysis method, which is not only very hard, but also generally has low prediction accuracy. In order to solve the above problems, this paper proposes a Feature-transferred Prediction Network (FTPN), which can adapt to various working conditions, combining with the neural network method in the field of artificial intelligence, and effectively realizes RUL prediction. Since the existing neural network methods were originally proposed to solve the problems in other fields, such as semantic recognition, this paper uses the feature transfer method to improve them. Firstly, the source network based on Convolutional Neural Network (CNN) is pre-trained for fault recognition, and the fault feature information extracted after training is stored

in the fault feature layer. Second, CNN and Gate Recurrent Unit (GRU) are used to build target networks and fit the relationship between time series and remaining longevity. Finally, a special loss function is designed to transfer the features extracted from the source network to the target network to help the target network learn fault features and better predict the mechanical RUL. In order to verify the effectiveness of the proposed method, experiments are carried out using the public data set of accelerated life of bearings, and high prediction accuracy is obtained, which proves that the proposed method has certain generalization. The comparison with the existing methods on the same data set shows that the proposed method has a broad industrial application prospect.

Keywords: Bearings, Remaining Useful Life(RUL), Feature Transferred Learning, Convolutional Network (CNN), Gate Recurrent Unit (GRU)

1 Introduction

With the rapid development of modern industry, CNC machine tools, as an important equipment to achieve intelligence, are constantly developing towards the direction of high precision and high integration. Spindle is the core component of CNC machine tools, it bears the main movement of parts processing, and often run under bad conditions, so its performance directly affects the machining accuracy and product quality of CNC machine tools. According to statistics, the main shaft faults are mostly caused by bearing failure[1],[2]. When the bearing degrades or fails, it will affect the operation of the whole CNC machine tool and even cause safety accidents. In applications, methods such as post-maintenance and regular maintenance are often adopted, but the above-mentioned methods will seriously reduce the efficiency of the equipment and cause economic losses[3],[4]. Therefore, prognostics and health management (PHM) technology has been widely discussed recently. Among them, remaining useful life (RUL) prediction, as the main core of PHM technology research, has become a focus in the current PHM field research[5],[6].

How to improve the accuracy and generalization of the RUL prediction method is the focus and difficulty of the research. At present, there are mainly two methods of RUL prediction: failure mechanism analysis-based prediction [7],[8],[9] and data-driven prediction [10],[11],[12]. Traditional methods often rely on modeling of failure mechanisms. This method not only requires a lot of professional knowledge, and established model can only be applied to a specific equipment under a specific working condition. For example, Marble and Morton developed a physics-based bearing RUL prediction method by constructing a finite element model [13]. This model needs to describe a series of elements such as the geometry of the bearing, the load and speed of the rolling elements, and cannot solve a wider range of problems.

With the rapid development of big data and artificial intelligence technology, the data-driven method is the most widely used RUL prediction

method[8]. For example, using machine learning techniques to achieve data-driven RUL prediction has received great attention, e.g., artificial neural networks (ANNs) [14], [15], support vector machine (SVM) [16], and Gaussian process regression (GPR) [17], [18] have been successfully practiced in mechanical RUL prediction. Zhang et al. proposed a RUL prediction method based on Naive Bayes (NB) and Weibull distribution (WD), using an improved WD algorithm to fit the fluctuation characteristics in different decay stages, and as the NB training stage enter [19]. Ali et al. developed an Ann-based method that can achieve accurate pump bearing prediction [20]. Aye et al. used a combination of principal component analysis and Gaussian mixture model to construct a health assessment index, and established an improved GPR model for predicting slow-speed bearing failures [21]. Soualhi et al. used Hilbert-Huang transform to extract features from the vibration signals and input them into SVM for prediction [22].

However, due to the limited training depth and high requirements on the amount of training data, the above methods are not competent for a wider range of machine RUL prediction [23], [24], [25]. Therefore, some neural network methods with deep network structure become the focus of attention in the field of RUL prediction, such as Gate Recurrent Unit (GRU), long and short-term memory network (LSTM), recurrent convolutional neural network (RCNN), etc [26], [27], [28]. Zheng et al. proposed a method for estimating RUL based on LSTM [29], which uses multi-layer LSTM units combined with standard feed-forward layers to extract hidden fault degradation from sensors and operating data of multiple operating conditions, faults and degradation models.

These methods were originally proposed to solve other problems such as semantic analysis, so how to use the existing neural network framework to improve and innovate the machine RUL prediction is very important. Considering that one of the advantages of neural network lies in the extraction of fault features, and the change of fault features over time is the core problem of RUL prediction, if the extracted fault features can be deliberately migrated to the RUL prediction process, it may effectively help it to achieve prediction. Based on this, this paper establishes a RUL prediction method based on fault feature migration, which effectively improves the prediction accuracy, and has been successfully applied to the RUL prediction experiment of bearings. Therefore, this method has important significance and application value for the reliability monitoring of machine tools and similar equipment.

2 Methods

In this section, the structure and principle of FTPN are introduced in detail. The method has a source network and a target network. The Convolutional Neural Network (CNN) and Gate Recurrent Unit (GRU) is used in this method to build the basic structure. Moreover, a core feature transfer layer is constructed based on them. After the parameters of the feature transport layer are fixed, the target network and the parameters of the feature receiving layer

are combined with a special loss function to assist its training. The migrated fault features in the source network are pretrained by simple fault identification tasks. The diagram of the network structure is shown in Fig.1.

Fig. 1 The computational framework of the proposed FTPN.

2.1 Source Network

The source network is actually a fault feature extraction network. It trains convolutional neural networks with simple fault recognition tasks. Prior to this, many scholars have successfully applied CNNs to machine fault recognition[30],[31],[32]. In the experiment, they concluded that when the CNN completed the fault identification task, the fault features would be saved in the parameters of the network. Therefore, we can use this important function of CNN to build a source network to extract features in advance. The input to the source network is a sample in the form of a one-dimensional vector labeled

"fault" and "good". CNN contains convolution layers and linear layers, the former are used for downsampling and extracting the features of high dimensional space, and the latter are used to select and fit the most suitable feature from multidimensional features.

2.1.1 Convolution Layer

Assuming that the input is one-dimensional signal $x(i)$, the output is $y(i)$, the length of the convolution layer is k , the interval distance between the sampling source data in the convolution operation is d , and the convolution layer is expressed as $w(k)$. In this case, the working principle of the convolution layer can be expressed as follows.

$$y(i) = \sum_{k=1}^k x(i + kd \cdot k) \cdot w(k) \quad (1)$$

2.1.2 Linear Layer

The linear layer includes a linear part and a nonlinear part. The operation method of linear part is linear weighted sum operation. That is, for an input vector x containing n terms, the network has a corresponding parameter matrix W and bias term b with a size of $m \times n$. The relationship between the above parameters and the output vector Z is shown in the following formula.

$$W * x + b = Z \quad (2)$$

The network constantly updates parameters W and b through training, and finally makes the network fit the functional relationship between the input and the target output well. However, only the linear part mentioned above is not enough to represent all relations, so nonlinear factors, namely activation functions, need to be introduced. It introduces non-linear factors so that it can approximate any non-linear function. Commonly used activation functions are Sigmoid, Tanh, ReLU, etc. The activation function used by CNN in this paper is the PReLU function. The method proposed by He et al. is an innovation and promotion of the ReLU function [33]. They add a small number of parameters to the ReLU function to make it adaptively adjust according to the existing data. The definition of PReLU is shown in formula :

$$f(x_i) = \begin{cases} x_i, & x_i > 0, \\ a_i x_i, & x_i \leq 0 \end{cases} \quad (3)$$

The subscript i represents different channels, and the parameter a_i is dynamically updated using the momentum method, which is shown in formula:

$$\Delta a_i = \mu \Delta a_i + \phi \frac{\partial \epsilon}{\partial a_i} \epsilon \quad (4)$$

In the above formula, μ the momentum, is the objective function, ϕ is the learning rate, and a_i is the slope of the negative axis. After experimental verification by He et al., the initial value of a_i can be set to 0.25 [33]. When $a_i = 0$, the PReLU function at this time is converted to the traditional ReLU function. The difference between the two can be clearly seen in the Fig.2. It can be seen from the definition that the PReLU function does not add many parameters, so the calculation amount of the network will not increase significantly. The a_i value is constantly updated during the training process, which can better adapt to the data characteristics and improve the network performance. The parameters of the linear layers after training update can represent the selected fault features to a certain extent, which is more accurate than the fault feature information contained in the convolutional layer. Therefore, in this method, the linear layers is regarded as the fault feature layer, or the source domain of feature migration.

Fig. 2 The difference between ReLU and PReLU function.

2.2 Target Network

The target network is the main network to realize RUL prediction. RUL prediction needs to deal with time series tasks, but networks such as CNN can only analyze each sample independently without other improvements, and cannot well learn the relationship between the inputs before and after. Therefore, we choose GRU which is more suitable for dealing with time series as the main body of the prediction model. In addition, in order to enable the fault features extracted from the source network to be successfully migrated to the target network, a feature receiving layer with the same homologous network structure should be set up before GRU. Since the structure of the feature receiving layer is completely consistent with the source network in A, only the GRU layer behind the feature receiving layer is introduced in this section. GRU was proposed by Cho, et al. in 2014 [34]. It is a kind of recurrent neural network, which consists of linear layers and also has weight W and bias b . But in the linear layers of CNN, weight links only exist between layers, while in the hidden layer of GRU, weight links also exist between neurons within layers. Its principle schematic diagram is shown in Fig.3. The GRU has two gates: reset

Fig. 3 Schematic diagram of GRU.

gate and Update gate. They determine how new input information is combined with previously remembered information. In GRU, for the input x_t at the current time t , GRU has a corresponding hidden state, h_t , which contains information about the previous node and is the output of the unit. Based on these parameters, the working principle of GRU is as follows,

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]), \quad (5)$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]), \quad (6)$$

$$\tilde{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t]), \quad (7)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t. \quad (8)$$

It can be seen from the above expressions that h_t is not only related to the input at the current moment, but also to the h_{t-1} at the previous moment.

2.3 Training Method based on Lock-unlock Feature Transfer

2.3.1 Pre-training of Source Network

The source network should be pre-trained before the formal training begins. Because the source network performs the task of fault identification, the cross-entropy function was chosen because it is suitable for classification, which is written as,

$$loss(x, class) = -x[class] + \log \left(\sum_j \exp(x[j]) \right) \quad (9)$$

The output of the last layer of the network is y_{pred} , m represents the total number of samples and y_{true} represents the fault type corresponding to the samples.

2.3.2 Feature Transfer and Target Network Training

In the course of target network training, there are two stages of locking and unlocking. The locking and unlocking here refer to the feature transfer process, that is to say, feature transfer does not occur in the whole training process. Feature migration is introduced at the early stage of training to help the target network learn fault features quickly, but the influence of migrated features on the network is unlocked at the later stage of training. The above process is achieved by the following loss function. At this stage, the loss function of the network consists of two parts, one is fault characteristic loss, the other is RUL loss. Indicates that the characteristic migration layer parameter of source network is f_{sour} , the characteristic receiving layer parameter of target network is f_{tar} , the output of target network is r , and the true value of RUL is R . The loss function can be written as,

$$MSELoss(\vec{x}, \vec{y}) = \frac{1}{N} \sum_i^N (x_i - y_i)^2, \quad (10)$$

$$\begin{aligned} Loss_{main} = & a_1 * MSELoss(\vec{f}_{sour}, \vec{f}_{tar}) \\ & + a_2 * MSELoss(r, R), \end{aligned} \quad (11)$$

Parameter a_1 is larger than a_2 at the beginning of training, and it is locked at this time. As the training progresses, if the main loss function no longer attenuates, the leader a_1 also attenuates to 0, and the training continues in the unlocked state until the network converges completely.

3 Experiment and Analysis

3.1 Data Description

Fig.4 is a test bed for accelerated life of rolling bearings [35]. The test bench is used for bearing acceleration and attenuation experiments under different working conditions. It is composed of ac motor, motor governor, support shaft, two support bearings and hydraulic loading system. The two accelerometers of PCB 352C33 are placed at 90° position of the bearing under test, that is, one is placed on the horizontal axis and the other on the vertical axis. The radial force is generated by the hydraulic loading system and applied to the bearing bush under test. The ac induction motor speed controller sets and maintains the speed, sets the sampling frequency as 25.6 kHz, records 32,768 data points every 1 minute, and finally outputs the complete bearing operation to failure data.

As shown in Table 1, 2, the test carried out tests on 15 bearings of the above types under three different working conditions. There were many reasons for the failure of the bearings under test, including inner ring wear, cage fracture, outer ring wear and outer ring fracture. As shown in Fig.5, the bearing degradation process obviously includes two stages, namely normal operation stage

Fig. 4 Bearing testbed.

Table 1 Bearing dataset information

Condition	Dataset	Sample Number	Failure Location
1	1-1	123	Outer
1	1-2	161	Outer
1	1-3	158	Outer
1	1-4	122	Cage
1	1-5	52	Outer
2	2-1	491	Inner
2	2-2	161	Inner, Outer
2	2-3	533	Cage
2	2-4	42	Outer
2	2-5	339	Outer
3	3-1	2538	Outer
3	3-2	2496	Outer, Inner, Cage, Rolling body
3	3-3	371	Inner
3	3-4	1515	Inner
3	3-5	114	Outer

and degradation stage. In the normal operation stage, the vibration signal only has a low level of random fluctuation. In the degradation stage, the amplitude of vibration signal increases with the extension of working time, which contains abundant bearing degradation information. Therefore, this paper only discusses the experiment of RUL prediction when the bearing starts to degrade. According to the adaptive degradation detection method proposed by Li et al., the first Predicting Time (FPT) can be determined [36]. In addition, for safety reasons, the accelerated degradation test of bearing was stopped when the amplitude of vibration signal was higher than 20g. Accordingly, the moment

Table 2 Bearing life acceleration test condition information.

Condition	Speed (r/min)	Radial Force/kN
1	2100	12
2	2250	11
3	2400	10

when the amplitude of vibration signal exceeds $20g$ is regarded as the failure time of the bearing under test, and the RUL label at this moment is set to 0.

Fig. 5 Typical horizontal vibration signals.

3.2 Prognostics metrics

Take use of both the score function and the root mean-square error (RMSE) to evaluate the performance of this algorithm. Given the value of the label r_{label} and our predicted value r_{pre} , denote the difference between these two values as $d_i = r_{label}(i) - r_{pre}(i)$, the score function S and the RMSE is calculated to be,

$$S = \sum_i^M s_i,$$

$$s_i = \begin{cases} e^{-\frac{d}{13}-1}, & d_i < 0, \\ e^{\frac{d}{10}-1}, & d_i > 0, \end{cases} \quad (12)$$

$$RMSE = \sqrt{\frac{1}{M} \sum_{i=1}^M d_i} \quad (13)$$

Fig. 6 shows the difference of Scoring function and RMSE. $h = 0$ means estimated RUL is the same as true RUL. For both Scoring function and RMSE, the smaller value, the better the result is.

Fig. 6 Comparison of Scoring Function and RMSE.

3.3 Data Preprocessing

Since the load is applied in the horizontal direction in the experiment, the accelerometer placed in this direction can capture more degradation information of the bearing under test. Therefore, in this experiment, only horizontal vibration signals were selected to test the performance of FTPN. For the convenience of comparison, the data collected by each sensor in the 4 sub-data sets are normalized to the range of -1 to 1, using the min-max normalization method,

$$X_{norm}(i) = \frac{X(i) - X_{min}}{X_{max} - X_{min}} \quad (14)$$

After normalization, samples were made by sliding time window. Assume that the time window size is N . The number of input features is $N+1$. The intercepted samples are labeled according to the time series, which is the input of the target network. The data before FPT and the data after the amplitude reaches 20g can well represent the characteristics of the normal state and the fault state respectively. The samples are also made according to the above method and can be used as the input of the source network for pre-training.

3.3.1 RUL Prediction for the Tested Bearings

Bearing1-1, Bearing2-1 and Bearing3-1 were selected as test sets and other data sets as training sets. Samples were made according to the above data preprocessing method, and FTPN was input to carry out the learning process of

pre-training and main training in two stages. The predicted results are output after the loss function converges. Fig.7, 8, 9 shows the prediction results of FTPN and the comparison results with Multilayer Perceptron (MLP), support vector regression (SVR), and Deep Neural Networks (DNN).

Fig. 7 RUL prediction results using 4 different prognostics approaches for Bearing 1-1.

Fig. 8 RUL prediction results using 4 different prognostics approaches for Bearing 2-1.

The results of fractional function and RMSE are shown in Table 3. The first observation was that FTPN was superior to the other three machine learning methods in score and RMSE for each of the three tested bearings. Firstly, FTPN is a model based on deep learning method, which has advantages in

Fig. 9 RUL prediction results using 4 different prognostics approaches for Bearing 3-1.

learning degradation process. Secondly, this method innovatively adopts the method of feature transfer, which can help FTPN learn fault features quickly at the initial stage of training. Therefore, FTPN's learning effect is better than other methods in the same number of iterations. In addition, it can be seen from Figure that FTPN of the three bearings under test has a good prediction effect, which proves that FTPN can adapt to the prediction under different working conditions and has generalization.

Table 3 Bearing dataset information

		MLP	SVR	DNN	FTPN
Bearing1-1	RMSE	2.88	1.57	0.87	0.82
	Score	7.0090	6.2711	5.8666	5.8443
Bearing1-1	RMSE	4.21	3.25	2.93	1.84
	Score	17.0509	15.6947	14.9392	13.5202
Bearing1-1	RMSE	14.00	9.54	12.26	8.17
	Score	195.1906	117.4580	164.1902	98.5489

Compared with the other three methods, the superiority of FTPN in accuracy has been verified. The next step is to verify whether feature migration optimizes the prediction process. Therefore, FTPN without feature transfer is specially verified in this paper. The weight of the parameter representing feature migration in the main loss function was set to 0, and other structures and parameters remained unchanged, and the results were observed with the same number of iterations. The test loss function comparison of one training is shown in the Fig 10.

It can be seen from the figure that the loss of FTPN has staged difference when it is around 50 epochs. This is because in the experiment shown in the

Fig. 10 Test loss function of FTPN and FTPN without feature transfer.

figure, 50 epochs were set as "unlocked" time points. According to the experiment, we can clearly see that loss decreases rapidly in the first stage, because in this stage, the target network makes full use of the features transferred from the source network and quickly grasps the fault features from numerous information. As the training continued, loss gradually slowed down around 40 epochs and a platform emerged. But the network is not completely converging at the moment. Considering that the positive effect of feature transfer on training is not obvious at this time, or even hindering training, "unlocking" network is selected, that is, the influence of feature transfer is no longer considered, and the training continues directly to the predicted target. In the second stage after unlocking, Loss no longer keeps the previous platform falling again, and finally converges. However, in the network without feature transfer, loss keeps a slow decline due to the lack of existing feature knowledge to assist its training at the beginning of training. At 200 epochs at the end of this set of experiments, the network without feature transfer still did not converge. Through this group of experiments, it can be proved that the feature transfer in FTPN is not only effective in improving the training accuracy, but also can effectively reduce the convergence time and improve the training speed.

4 CONCLUSION

In this paper, a novel feature-transferred prediction network for RUL of Rotating Machinery is proposed. The application of this method in RUL prediction experiment of bearing, an important part of CNC machine tool, verifies its practical application value. The innovation of the proposed method lies in the

application of feature transferred learning and the establishment of the corresponding special loss function. CNN and GRU are used to construct the source network of fault feature extraction and the target network of RUL prediction. Then the special Loss fuctions is to make these two tasks interact and achieve the feature transfer. It is worth mentioning that the training process of FTPN is divided into two stages: locking and unlocking. The difference and purpose of the two stages is that the effect of feature transfer on training only plays a large role in the first half of training, and gradually weakens in the second half of training, which is more conducive to the training accuracy and convergence speed. In the bearings RUL prediction experiment, FTPN was compared with the other three PREDICTION methods of RUL, and it was found that the prediction accuracy of FTPN was significantly higher than the other three prediction methods in terms of score function and RMSE under the same number of iterations. In addition, the comparison between FTPN and the network with no feature transfer process and consistent other structures shows that the feature transfer process also has a very positive effect on the training accuracy, which proves the feasibility and superiority of this method. It can be seen that this method has the potential to be extended to the actual reliability monitoring process of machine tool components and has certain practical application value.

Declarations

- Funding: This work was supported by the National Natural Science Foundation of China (Grant NO.52075030 and 51805022), Joint Project of BRC-BC (XK2020-04) and the National Key R&D program of China (Grant No.2018YFB1503103).
- Competing interests: The authors declare no competing interests.
- Ethics approval: Not applicable.
- Consent to participate: Not applicable.
- Consent for publication: Not applicable.
- Availability of data and materials: Not applicable.
- Code availability: Not applicable.
- Authors' contributions: All authors took part in the work.

References

- [1] Martinova, e.a. L. I.: Creation of cnc system's components for monitoring machine tool health. The International Journal of Advanced Manufacturing Technology, 1–8 (2021)
- [2] Tobon-Mejia D, Z.N. Medjaher K: Cnc machine tool health assessment using dynamic bayesian networks. IFAC Proceedings Volumes **44**(1), 12910–12915 (2011)

- [3] Si, X.-S., Wang, W., Hu, C.H., , Zhou, D.-H.: Remaining useful life estimation—a review on the statistical data driven approaches. *European Journal of Operational Research* **213**(1), 1–14 (2011)
- [4] Azadeh, A., Asadzadeh, S.M., Salehi, N., Firoozi, M.: Condition-based maintenance effectiveness for series-parallel power generation system - a combined markovian simulation model. *Reliability Engineering and System Safety* **142**, 357–368 (2015)
- [5] Wang, D., Tsui, K.L., Miao, Q.: Prognostics and health management: A review of vibration based bearing and gear health indicators. *IEEE Access* (2017)
- [6] Y, L., N, L., L, G., et al.: Machinery health prognostics: A systematic review from data acquisition to rul prediction. *Mechanical Systems and Signal Processing* **104**, 799–834 (2018)
- [7] H, D., N, H., Z, W., et al.: Finite-horizon fault estimation under imperfect measurements and stochastic communication protocol: Dealing with finite-time boundedness. *International Journal of Robust & Nonlinear Control* (2018)
- [8] Pecht, M., Gu, J.: Physics-of-failure-based prognostics for electronic products. *Transactions of the Institute of Measurement and Control* **31**(3–4), 309–322 (2009)
- [9] Wu, B., Yufit, V., et al.: Differential thermal voltammetry for tracking of degradation in lithium-ion batteries. *Journal of Power Sources* **273**, 495–501 (2015)
- [10] H, Z., Q, M., X, Z., et al.: An improved unscented particle filter approach for lithium-ion battery remaining useful life prediction. *Microelectronics Reliability* **81**, 288–298 (2018)
- [11] B, W., V, Y., Y, M., et. al. : Differential thermal voltammetry for tracking of degradation in lithium-ion batteries. *Journal of Power Sources*, 495–501 (2015)
- [12] Q, M., L, X., H, C., et. al. : Remaining useful life prediction of lithium-ion battery with unscented particle filter technique. *Microelectronics Reliability*, 805–810 (2013)
- [13] Marble, S., Morton, B.P.: Predicting the remaining life of propulsion system bearings. *Aerospace Conference IEEE* (2006)
- [14] Chong, Z., et al.: Multiobjective deep belief networks ensemble for remaining useful life estimation in prognostics. *IEEE Trans Neural Netw Learn*

- Syst **28**(10), 2306–2318 (2017)
- [15] Guo, L., Li, N., Jia, F., Lei, Y., Lin, J.: A recurrent neural network based health indicator for remaining useful life prediction of bearings. *Neurocomputing* **240**, 98–109 (2017)
- [16] Loutas, T.H., Roulias, D., Georgoulas, G.: Remaining useful life estimation in rolling bearings utilizing data-driven probabilistic e-support vectors regression. *IEEE Trans. Reliab.* **62**, 821–932 (2013)
- [17] Sheng, H., Xiao, J., Cheng, Y., Ni, Q., Wang, S.: Short-term solar power forecasting based on weighted gaussian process regression. *IEEE Trans. Ind. Electron.* **65**(1), 300–308 (2018)
- [18] Liu, X., Song, P., Yang, C., Hao, C., Peng, W.: Prognostics and health management of bearings based on logarithmic linear recursive leastsquares and recursive maximum likelihood estimation. *IEEE Trans. Ind. Electron.* **65**(2), 1549–1558 (2018)
- [19] N, Z., L, W., Z, W., Y., G.: Remaining useful life prediction based on naive bayes and weibull distributions. *Entropy*, 944 (2018). <https://doi.org/10.3390/e20120944>
- [20] Ali, B., -Morello, B.C., L. Saidi, S.M., Fnaiech, F.: Accurate bearing remaining useful life prediction based on weibull distribution and artificial neural network. *Mech. Syst. Signal Process.*, 150–172 (2015)
- [21] Aye, S.A., Heyns, P.S.: An integrated gaussian process regression for prediction of remaining useful life of slow speed bearings based on acoustic emission. *Mechanical Systems and Signal Processing*, 485–498 (2017)
- [22] Soualhi, A., Medjaher, K., Zerhouni, N.: Bearing health monitoring based on hilbert-huang transform, support vector machine, and regression. *IEEE Transactions on Instrumentation and Measurement*, 485–498 (2017)
- [23] Liu, Q., Wang, H.: A case study on multi-sensor data fusion for imbalance diagnosis of rotating machinery. *Ai Edam*, 203–210 (2001)
- [24] Li, S., et al.: An adaptive data fusion strategy for fault diagnosis based on the convolutional neural network. *Measurement* (2020)
- [25] Hochreiter, S., Schmidhuber, J.: Long short-term memory. *Neural computation* **9**(8), 1735–1780 (1997)
- [26] Jiang, Y., et al.: Fusion network combined with bidirectional lstm network and multiscale cnn for useful life estimation. 2020 12th International Conference on Advanced Computational Intelligence (ICACI) IEEE

18 *Article Title*

(2020)

- [27] Li, F., Zhang, L., Chen, B., et al.: A light gradient boosting machine for remaining useful life estimation of aircraft engines. 2018 21st International Conference on Intelligent Transportation Systems (ITSC) IEEE (2018)
- [28] Li, X., Ding, Q., Sun, J.Q.: Remaining useful life estimation in prognostics using deep convolution neural networks. Reliability Engineering and System Safety **172**, 1–11 (2018)
- [29] Z, S., K, R., A, F., et al.: Long short-term memory network for remaining useful life estimation. IEEE International Conference on Prognostics and Health Management (ICPHM) (2017)
- [30] Wang, P., Song, L., et al.: A high-stability diagnosis model based on a multiscale feature fusion convolutional neural network. IEEE Transactions on Instrumentation and Measurement **70**, 1–9 (1997)
- [31] Song, L., Wang, H., Chen, P.: Automatic patrol and inspection method for machinery diagnosis robot—sound signal-based fuzzy search approach. IEEE Sensors **20**(15), 8276–8286 (2020)
- [32] Wang, P., Song, L., et al.: A light intelligent diagnosis model based on improved online dictionary learning sample-making and simplified convolutional neural network. Measurement **183** (2021)
- [33] K, H., X, Z., S, R., et al.: Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. 2015 IEEE International Conference on Computer Vision of the Proceeding (2015)
- [34] Cho, K., et al.: Learning phrase representations using rnn encoder-decoder for statistical machine translation. Computer Science (2014). <https://doi.org/10.3115/v1/D14-1179>
- [35] Wang, B., Lei, Y., Li, N., Li, N.: A hybrid prognostics approach for estimating remaining useful life of rolling element bearings. IEEE Trans. Reliab (2018)
- [36] Li, N., Lei, Y., Lin, J., Ding, S.X.: An improved exponential model for predicting remaining useful life of rolling element bearings. IEEE Trans. Ind. Electron **62**(12), 7762–7773 (2015)