Research Article

Xianbiao Zhan, Zixuan Liu, Hao Yan, Zhenghao Wu, Chiming Guo, and Xisheng Jia*

A two-stage framework for predicting the remaining useful life of bearings

https://doi.org/10.1515/phys-2023-0187
received November 07, 2023; accepted January 09, 2024

Abstract: The traditional prediction of remaining useful life (RUL) for bearings cannot be calculated in parallel and requires manual feature extraction and artificial label construction. Therefore, this article proposes a two-stage framework for predicting the RUL of bearings. In the first stage, an unsupervised approach using a temporal convolutional network (TCN) is employed to construct a health indicator (HI). This helps reduce human interference and the reliance on expert knowledge. In the second stage, a prediction framework based on a convolutional neural network (CNN)–transformer is developed to address the limitations of traditional neural networks, specifically their inability to perform parallel calculations and their low prediction accuracy. The life prediction framework primarily maps the complete life data of bearings onto the HI vector. Based on the HI constructed through TCN, the known HI is input into the CNN–transformer network, which sequentially predicts the remaining unknown HI. Finally, the effectiveness and superiority of the proposed method are verified using two bearing datasets, providing validation of its capabilities.

Keywords: temporal convolutional network, transformer, health indicators, remaining useful life

1 Introduction

Bearings are used in various industrial fields and are an important component of mechanical equipment. Often works under complex and harsh environmental conditions. Often, before the replacement time is reached, bearings have already been malfunctioned and damaged, causing equipment downtime and, in severe cases, personal injury or death. So predicting the remaining useful life (RUL) of bearings can help improve the safe and reliable operation of bearings [1].

The existing RUL prediction methods are mainly divided into three parts: data collection, construction of health indicators (HIs), and life prediction [2]. A good HI can intuitively and accurately reflect the degradation status of mechanical equipment and contribute to subsequent life prediction [3]. The common HI construction methods mainly extract the time-domain, frequency-domain, or time–frequency domain features of the original signal as HI, or fuse the extracted multi-dimensional features to construct HI. For example, Malhi *et al.* [4] used the continuous wavelet transform to extract root mean square (RMS) and peak values from the original signal as HI to predict the remaining life of bearings. Widodo *et al.* [5] constructed HI by dimensionality reduction of multidimensional time-domain features using the principal component analysis (PCA) technique. The above HI construction method can better reflect the degradation law of bearings. However, the HI constructed using the above method requires manual feature extraction, which can easily introduce human error. In recent years, the rise of deep learning has provided effective technical means for data mining. By training neural network models, hidden features within the original data can be effectively extracted, and the mapping relationship between features and lifespan can be well established. Wu *et al.* [6] proposed a multi-scale convolutional neural network (CNN) method to construct HI and compared it with other existing HI construction methods. It was found that the method proposed in this article can better characterize the degradation state of bearings. Zhang *et al.* [7] proposed an HI construction method based on a deep multilayer perceptron (MLP) CNN model to address the need for manual feature extraction in traditional HI construction methods. The effectiveness and reliability of the proposed method were verified on different datasets of bearings. Although using traditional deep learning techniques to construct HI for life prediction has achieved good

*** Corresponding author: Xisheng Jia,** Shijiazhuang Campus, Army Engineering University of PLA, Shijiazhuang, China,
e-mail: xs_jia2022@163.com
Xianbiao Zhan: Shijiazhuang Campus, Army Engineering University of PLA, Shijiazhuang, China; Hebei Key Laboratory of Condition Monitoring and Assessment of Mechanical Equipment, Shijiazhuang, China
Zixuan Liu: North China Institute of Aerospace Engineering, Langfang, China
Hao Yan, Zhenghao Wu, Chiming Guo: Shijiazhuang Campus, Army Engineering University of PLA, Shijiazhuang, China

results, it overcomes the drawbacks of manually extracting features. However, the above methods require manual labeling of training data, which is time-consuming and laborious, and cannot effectively reflect the degradation trend of mechanical equipment. Therefore, it is necessary to conduct research on unsupervised HI construction methods.

The current RUL methods are mainly divided into physical model-based and data-driven model-based prediction methods. The prediction method based on physical models requires expert experience and is difficult to establish accurate models for complex devices [8,9]. Therefore, existing prediction methods mainly adopt data-driven methods. The data-driven prediction method does not require the establishment of a bearing failure model, so its applicability is wider. Wang *et al.* [10] combined deep convolutional autoencoder (DCAE) and self-organizing map methods to construct HI and then used a CNN to predict the degradation trend of bearings. Zhou *et al.* [11] proposed a set of Gaussian mixture models and Kullback Leibler divergence methods to construct HI for the difficulty in obtaining life prediction labels. An improved gated recurrent unit (GRU) network is used to predict the lifespan of bearings. The existing deep learning-based bearing prediction methods mainly utilize recurrent neural network (RNN) and improved RNN (long short term memory [LSTM] and GRU) methods for RUL prediction [12]. However, RNN and improved methods are unable to perform parallel calculations and have drawbacks such as long-term dependence, resulting in low prediction accuracy. Therefore, existing scholars have proposed transformer models for lifespan prediction in response to the problems existing in RNN and its variants. The transformer model has achieved good prediction accuracy in natural language processing, financial time series prediction, and traffic flow prediction. Zhang *et al.* [13] used the transformer model to predict RUL for hydraulic systems, bearings, and gearboxes. Jiang *et al.* [14] proposed a dual-channel transformer prediction model to address the shortcomings of small receptive fields and long-term dependence in current mainstream deep learning frameworks. Compared with existing bearing prediction methods, it was found that the proposed method has superiority in bearing RUL prediction.

Although the above deep learning models have achieved good prediction results in HI construction and RUL prediction, there are still the following problems:

1) In HI construction and RUL prediction, there is still a need for manual feature extraction. It is necessary to manually label the training data, which may have human errors and be time-consuming and laborious;
2) The traditional RNN and its variants have defects such as the inability to perform parallel computation and long-term dependence, resulting in low prediction accuracy. Moreover, a single deep learning model may find it difficult to extract deep temporal features effectively when dealing with large datasets.

To address the aforementioned issues, this article presents an integrated network model designed to enhance the accuracy of RUL prediction for bearings. The model primarily utilizes a temporal convolutional network (TCN) for unsupervised HI construction, coupled with a CNN–transformer architecture for bearing life prediction.

The contributions and innovations of this article are summarized as follows:

1) It proposes an integrated network model for predicting RUL, which overcomes the limitations of traditional manual feature extraction and subjective labeling processes.
2) It combines CNN with the transformer architecture for life prediction, enhancing the accuracy of bearing RUL prediction. While existing transformer models are mainly applied in image classification, studies on life prediction are scarce.
3) The proposed method has been verified on multiple datasets. Initially, the approach was validated on the PHM2012 dataset and compared with other existing methods, establishing its superior performance. To further explore the degradation pattern of bearings, a complete lifecycle dataset of 957 h was collected in the laboratory to study the degradation trend of bearings.

The CNN–transformer prediction model proposed in this article primarily utilizes CNN to extract local features from the input sequences, which are then fed into the transformer model for global modeling and long-term relationship patterning. In the convolutional layers, the adjustment of kernel size and stride within the convolutional filters enables the capture of local sequence features. In the transformer model, the self-attention mechanism allows for capturing global dependencies between sequences. Therefore, the prediction method proposed in this article can effectively improve the accuracy of RUL predictions.

2 Theoretical analysis

2.1 Unsupervised HI construction method based on TCN

TCN was first proposed by Lea *et al.* in 2016 [15] and has been widely used in time series [16]. Due to its dilated causal convolution, TCN has a large receptive field and can effectively ensure that information is not "leaked." So TCN is more suitable for processing long sequence data than regular CNN. Compared with traditional deep temporal networks such as RNN and LSTM, TCN is more effective due to its dilated causal convolution and the use of residual connections between various network layers. Therefore, it can effectively extract sequence features and avoid the occurrence of gradient vanishing or explosion phenomena. The dilated causal convolution structure is shown in Figure 1(a), and the residual block is shown in Figure 1(b).

TCN is mainly composed of dilation causal convolution, residual module, and a one-dimensional full convolutional network. Expansive causal convolution is a one-dimensional sequence with input $x \in R^n$ and a convolution kernel $f: \{0,\ldots, k-1\} \rightarrow R$. The expansion convolution operation is applied to the element s in the sequence, which is defined as follows:

$$F(s) = (x \cdot_d f)(s) = \sum_{i=0}^{k-1} f(i) \cdot x_{s-d \cdot i}, \tag{1}$$

where d is the expansion factor. d will increase exponentially with the depth of the network layer i, ensuring that the receptive field expands while covering all valid inputs of the input time series. k is the size of the convolutional kernel.

The residual block replaces the simple links between layers in traditional network structures, making the network stable even at deep depths and improving the generalization ability of the network model. The expression for the residual block is

$$o = \text{Activation}(x + F(x)). \tag{2}$$

The one-dimensional fully convolutional network module ensures that the input and output time of each layer are the same, so that each time step's input has a corresponding output.

TCN has good temporal processing ability, which can effectively extract degraded features and largely avoid gradient vanishing and exploding behaviors. Therefore, TCN is selected to extract the deep temporal features of each degraded sequence as HI.

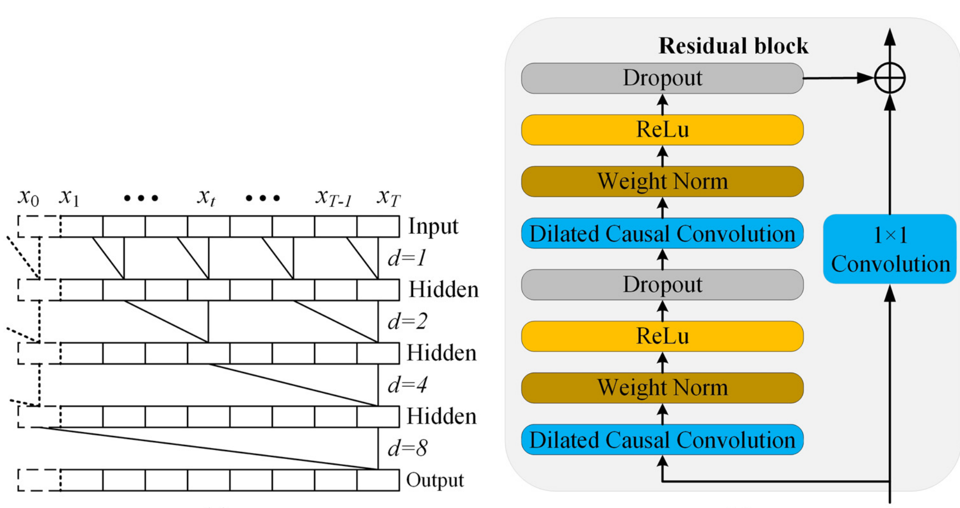
The original vibration signal of the bearing is input into the TCN network through a fast Fourier transform (FFT).

Figure 2 is the network structure diagram of the TCN autoencoder (TCNAE), which uses three residual blocks and one fully connected layer during the encoding process. The decoding process uses one residual block and an MLP network structure. The error backpropagation process was optimized using the Adam optimizer.

Therefore, the steps for constructing HI are as follows:

Step 1: Collect the original vibration signal and extract frequency domain features using the FFT transform. The frequency domain feature of the signal at the i moment represented by X_i is $X_i = (x_i^1, x_i^2,\ldots,x_i^n)$, where n is the number of features of each sample at each time. The compiled dataset is $(X_1, X_2,\ldots,X_m)^T$, where m represents the number of samples.

Step 2: Using the TCNAE network structure, perform deep feature extraction on the obtained frequency domain signal data and extract deep features with good representation ability and robustness as HI.



(a) (b)

Figure 1: TCN network structure: (a) structures of causal and dilated convolutions; (b) residual block.

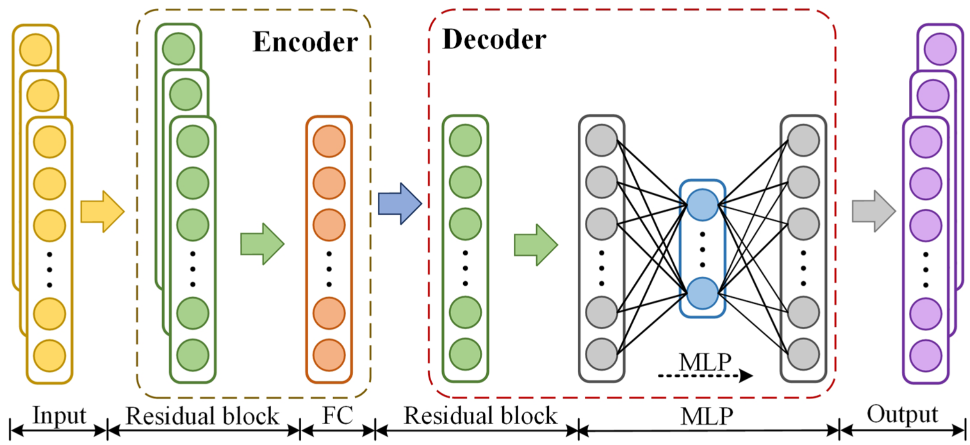


Figure 2: Construction of HIs in the TCN network.

2.2 Prediction model

2.2.1 CNN model

The CNN model was proposed by Le and Bottou in 1998 [17]. CNN is composed of multiple convolutional, pooling, and fully connected layers [18]. Each convolutional layer contains multiple convolutional kernels, which are calculated as shown in Eq. (3). CNN mainly uses convolutional and pooling layers to extract features and reduce feature dimensions. The convolutional layer formula is as follows:

$$l_t = \tanh(x_t \times k_t + b_t), \tag{3}$$

where l_t and x_t are the output and input vectors, k_t and b_t are the weights and biases of the convolutional kernel, and tanh is the activation function.

2.2.2 Transformer model

Transformer is a deep learning architecture that relies on attention mechanism [19], which solves the problem of traditional RNN not being able to perform parallel computation and accelerates the training speed of the network. Compared to CNN and other networks, they can handle longer sequence data and adopt a self-attention mechanism model to capture relationships at different positions in the sequence. The network structure of the transformer is shown in Figure 3. The transformer mainly adopts an encoder–decoder architecture, which is composed of a stack of M-layer networks with the same structure. Each layer includes two sub-layers: multi-head attention layer and full link layer. Use residual linking and normalization in each sub-layer to improve performance. A decoder is similar to an encoder, except that it contains two multi-head self-attention layers.

2.2.3 HI evaluation indicators

This article selects monotonicity, correlation, and robustness as the evaluation indicators for HI, which are used to quantitatively evaluate the performance of an HI. To this end, the polynomial fitting method is first used to decompose HI into smooth trends and random errors [20]:

$$H(t_n) = H_T(t_n) + H_R(t_n), \tag{4}$$

where $H(t_n)$ represents the value of HI at time t_n, $H_T(t_n)$ represents its smoothing trend, and $H_R(t_n)$ represents random error.

2.2.3.1 Monotonicity (Mon)

The monotonicity index is used to evaluate the trend of changes in an HI, including monotonic upward and downward trends. The formula is as follows [20]:

$$\mathrm{Mon(HI)} = \left| \frac{\mathrm{d}H_T(t_n) > 0}{m-1} - \frac{\mathrm{d}H_T(t_n) < 0}{m-1} \right|, \tag{5}$$

where Mon(HI) is the monotonicity of HI; $\mathrm{d}H_T(t_n)$ is the derivative of $H_T(t_n)$, and m is the vector length of HI.

2.2.3.2 Correlation (Corr)

The linear correlation between HI and the corresponding time is measured through correlation, and the calculation formula is [20]

$$\mathrm{Corr}(\mathrm{HI},T) = \frac{|\sum_{n=1}^{N}(H(t_n) - \overline{H})(t_n - \overline{T})|}{\sqrt{\sum_{n=1}^{N}(H(t_n) - \overline{H})^2 \sum_{n=1}^{N}(t_n - \overline{T})^2}}, \tag{6}$$

where $\mathrm{Corr}(\mathrm{HI},T)$ represents the correlation between HI and time, $\overline{H} = (1/N\sum_{n=1}^{N}H(t_n))$, and $\overline{T} = (1/N\sum_{n=1}^{N}t_n)$.

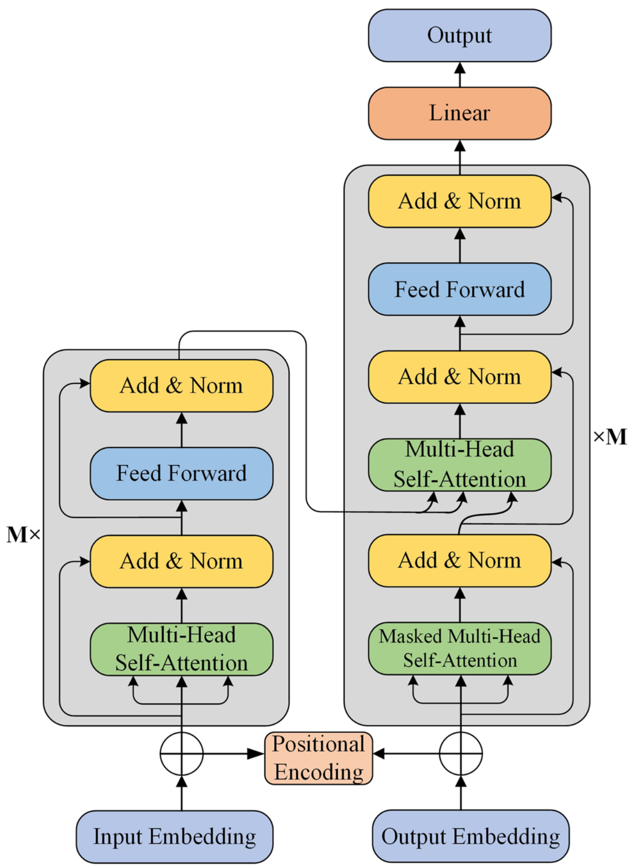


Figure 3: The structure of transformer memory cell.

2.2.3.3 Robustness (Rob)

Robustness is the evaluation of HI's tolerance for outliers, calculated using the following formula [20]:

$$\mathrm{Rob(HI)} = \frac{1}{N}\sum_{n=1}^{N} \exp\left(-\left|\frac{H_R(t_n)}{H(t_n)}\right|\right), \tag{7}$$

where Rob(HI) represents the robustness of HI.

In order to evaluate the overall performance of HI, a comprehensive indicator (CI) containing all three indicators is defined as follows [20,21]:

$$\mathrm{CI} = 0.3 \times \mathrm{Corr} + 0.4 \times \mathrm{Mon} + 0.3 \times \mathrm{Rob}. \tag{8}$$

2.2.4 Life prediction evaluation indicators

This article uses mean absolute error (MAE) and root mean square error (RMSE) as evaluation indicators for prediction performance, with the following formula [21]:

$$\mathrm{MAE} = \frac{1}{m}\sum_{i=1}^{m} |y_i - \widehat{y_i}|, \tag{9}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{m}\sum_{i=1}^{m} (y_i - \widehat{y_i})^2}, \tag{10}$$

where m is the length of the experimental data, y_i represents a vector composed of m actual labels, and $\widehat{y_i}$ represents a vector composed of m predicted labels.

2.2.5 Prediction model structure

The traditional RUL prediction model requires manual feature extraction, manual labeling of training data, and low prediction accuracy. This article proposes a new HI construction and RUL prediction framework. This framework mainly consists of two parts: first, the original vibration signal of the bearing is input into the TCNAE model through FFT transformation for deep feature extraction to obtain HI. Then, the HI is divided into training and testing sets and input into the CNN–transformer network for life prediction. This method not only overcomes the drawbacks of traditional manual feature extraction but also does not require manual labeling of training data. Combining the advantages of CNN and transformer, it can effectively improve the prediction accuracy of bearing remaining life. The overall flowchart of the method proposed in this article is shown in Figure 4, and the specific steps are as follows:

Step 1: Data acquisition and processing. Perform FFT transformation and data normalization on the original data.

Step 2: HI Build. Input the FFT-transformed data into the TCNAE model for deep feature extraction to obtain HI.

Step 3: Life prediction. Divide the constructed HI into a training set and a testing set. Input the training set data into the CNN–transformer network for training. Then, input the test set data into the trained prediction model and output the prediction results.

3 Instance verification

3.1 Case 1

In order to verify the effectiveness of the method proposed in this article, the research team conducted accelerated degradation experiments on the bearing model LYC6220E in the PHM Key Laboratory. This experiment collected a total of 957 h of full-life data. The acquisition equipment selected is the DH5981 acquisition system manufactured by Donghua Company, and the sensor model is CT1005LC. The

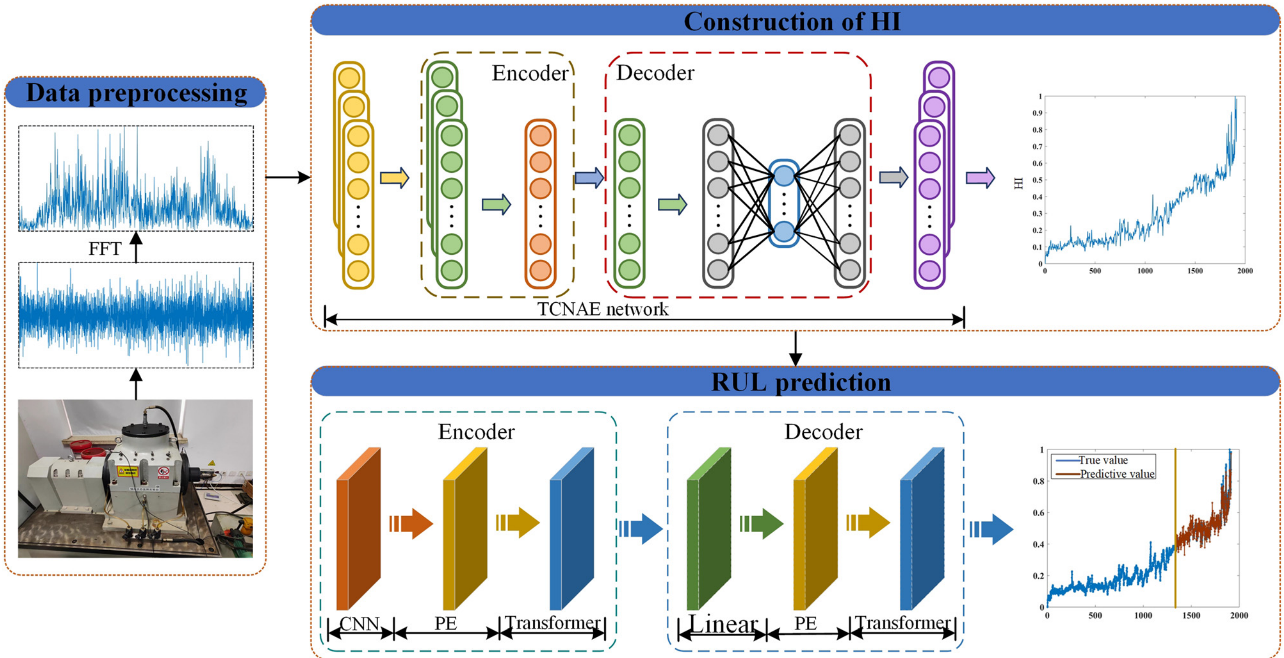


Figure 4: Prediction flowchart.

sampling frequency is 20 kHz, collected every 15 min for 12 s each time. The specific sensor layout and data acquisition system are shown in Figure 5.

The original vibration signals are acquired, and the original vibration signals are first transformed by FFT to transform the time-domain signals into frequency-domain signals. Then, input the frequency-domain signal into the TCNAE network. TCNAE mainly consists of two parts: encoding and decoding. The encoding part mainly consists of three residual blocks with 128, 64, and 32 neurons and a fully connected layer with 16 neurons, where the dilation factors of the residual module are 1, 2, and 4, respectively. The decoder part is composed of a residual module with a number of neurons of 128 and an MLP network structure, where the dilation factor of the residual block is 1. The network is trained using Adam optimizer with an initial learning rate of 0.001, batch size of 128, training epoch of 200, activation function of ReLu, loss function of mean squared error (MSE), and dropout of 0.25. The convolutional kernel size of residual blocks in TCN is all 5. In order to further eliminate the fluctuation caused by noise and obtain more intuitive health information about the device, locally weighted scatterplot smoothing (LOESS) [22] is used to smooth the HIs. LOESS can not only eliminate noise but also obtain smoother degradation characteristics of bearings in order to better grasp the current and future

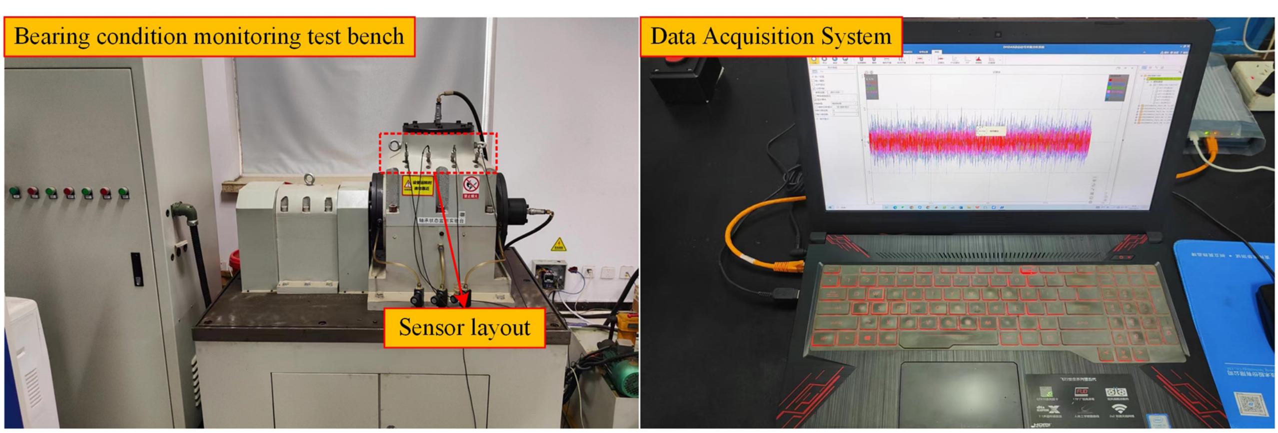


Figure 5: Bearing status monitoring experimental platform.

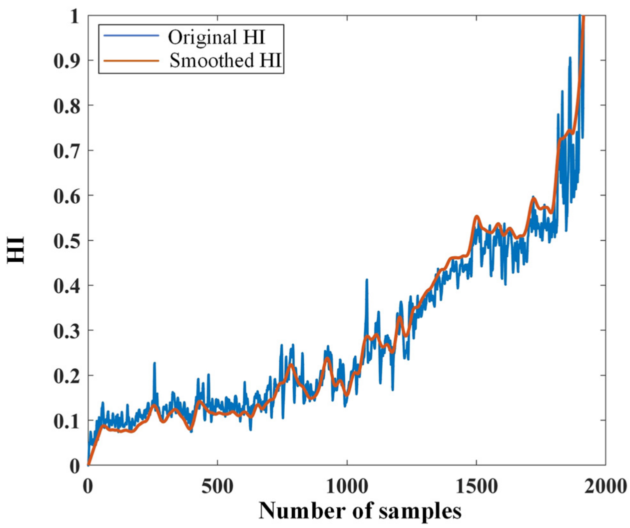


Figure 6: Health indicators constructed by TCNAE.

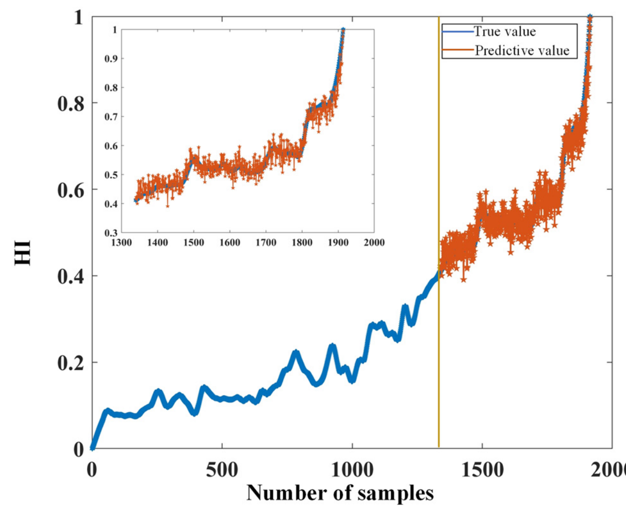


Figure 7: Prediction results of the proposed method in this article.

degradation situation of bearings. The HIs constructed using TCNAE and the results of the smoothing process are shown in Figure 6.

This paper conducted a comparative analysis with other methods to further verify the effectiveness of the method proposed herein. This article compares stacked sparse autoencoder (SSAE), convolutional autoencoder (CAE), and direct extraction of multi-domain features using PCA dimensionality reduction to construct HI methods. In order to better see the comparison results, the evaluation indicators are limited to the range of [0, 1]. The comparison results are shown in Table 1.

It can be concluded from the comparison results in Table 1 that the PCA-HI method is the least effective. This is because the direct fusion of multi-domain features is likely to fuse some inapplicable features together, which rather reduces the performance of HI. While the SSAE-HI and CAE-HI methods outperform the proposed method on individual indicators, their overall indicator (CI) is lower than the proposed method. Therefore, the HI construction method proposed in this article can better characterize the degradation trend of bearings and improve the accuracy of life prediction.

Using the HI construction method proposed in this article for life prediction, the HI is divided into a training set and a testing set at 7:3. First, input the training set into the CNN–transformer network for training. The basic parameters of the CNN–transformer network are as follows: batch size of 46, basic learning rate of 0.001, weight attenuation coefficient of 0.0001, loss function selection of MSE function, forward propagation dimension of encoder and decoder of 512, head count of multi-head attention mechanism of encoder and decoder of 4, and training batch of 200. The training results are shown in Figure 7, and the prediction results are localized and enlarged in order to better see the prediction curves of the proposed method in this article.

To verify the effectiveness of the method proposed in this article, RMSE and MAE were selected for quantitative analysis. We compared bidirectional long short term memory (BiLSTM), support vector regression (SVR), and transformer methods, and the predicted results are shown in Table 2.

From Table 2, it can be concluded that the method proposed in this article has better predictive performance and is superior to other methods. The prediction performance of a single transformer network is poor, possibly due to the inability to extract local features well and only the ability to extract global features. Therefore, utilizing

Table 1: Evaluation indicators for different HIs

Method	Mon	Corr	Rob	CI
TCNAE-HI	0.646	0.931	0.927	0.816
SSAE-HI	0.499	0.910	0.715	0.687
CAE-HI	0.668	0.906	0.764	0.768
PCA-HI	0.359	0.752	0.841	0.622

Table 2: Comparison results of different life-prediction methods

Method	MAE	RMSE
CNN–Transformer	0.021	0.129
BiLSTM	0.037	0.134
SVR	0.102	0.337
Transformer	0.059	0.298

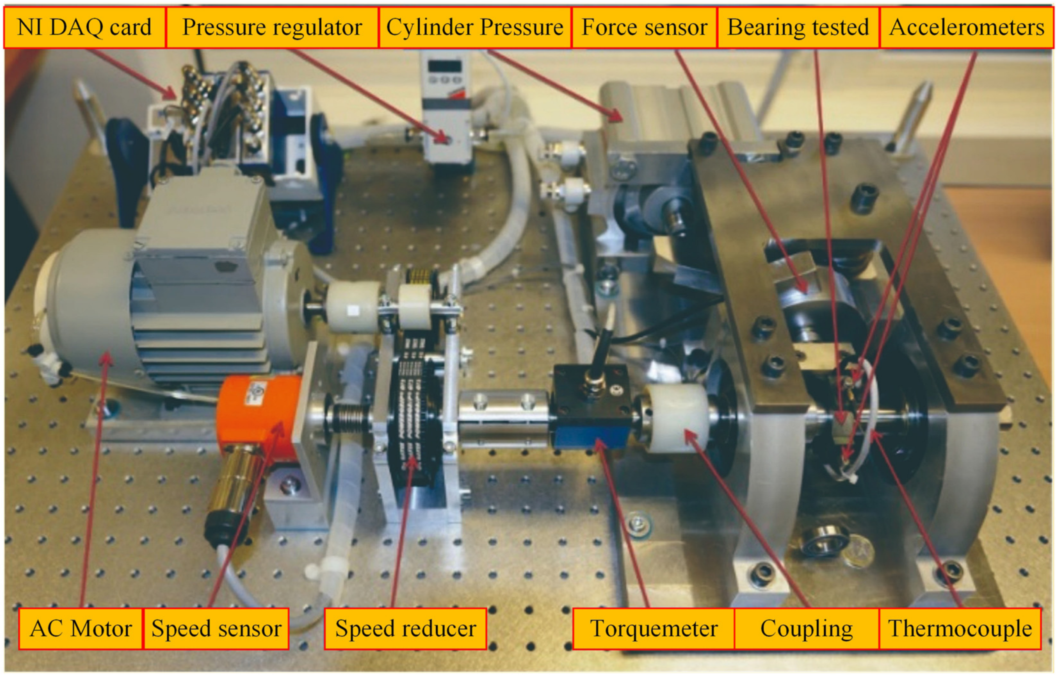


Figure 8: The PRONOSTIA platform.

the respective advantages of CNN and transformer can effectively extract local and global features and improve the accuracy of life prediction.

3.2 Case 2

The experimental data used in this article come from the IEEE 2012 PHM Data Challenge. As shown in Figure 8, select an acceleration sensor with a sampling frequency of 25.6 kHz for data collection. Each sample contains 2,560 points, recorded every 10 s. This article selects bear1-3 bearings under operating conditions of 1,800 rpm and 4,000 N load as examples for experimental analysis. Since bearings 1–3 have collected abundant experimental data with a longer full lifecycle, it is thus chosen to take bearings 1–3 as examples for experimental analysis, which is conducive to examining the predictive performance of the method proposed in this paper [23].

The network parameter method used in this experiment is the same as in case 1. First, collect the original vibration signal, perform FFT transformation on the original vibration signal, and transform the time-domain signal into a frequency-domain signal. Figure 9 shows the time-domain signal and the frequency-domain signal after FFT transformation at the beginning of the experiment at 0.1 s. Then, input the frequency domain signal into the TCNAE network. The HIs and smoothing results constructed using TCNAE are shown in Figure 10.

In order to better validate the effectiveness of the method proposed in this article, we compared the DCAE method in reference [24] with the regularized sparse autoencoder (SAEwR) method in reference [25]. The comparison results are shown in Table 3.

From the comparison results in Table 3, it can be concluded that the correlation between the HI method constructed in references [14,19] is higher than that of the method proposed in this article. However, its performance indicators of monotonicity and robustness are both lower than the method proposed in this article, and its comprehensive indicator (CI) is lower than the method proposed in this article. Therefore, the HI construction method proposed in this article can better characterize the degradation trend of bearings and improve the accuracy of life prediction.

Based on the construction of HI in this article, life prediction is carried out, and HI is divided into a training set and a testing set in a 7:3 ratio. The training results are shown in Figure 11. From Figure 8, it can be seen that using the CNN–transformer network can achieve good prediction results. However, the prediction results were poor after 2,300 samples because the degradation trend in the early stage of HI construction was relatively gentle, and bearing failure occurred suddenly after 2,300 samples. The divided training set does not contain sudden failure data, so the predicted HI curve still slowly rises.

In order to verify the effectiveness of the method proposed in this article, existing predictive methods were

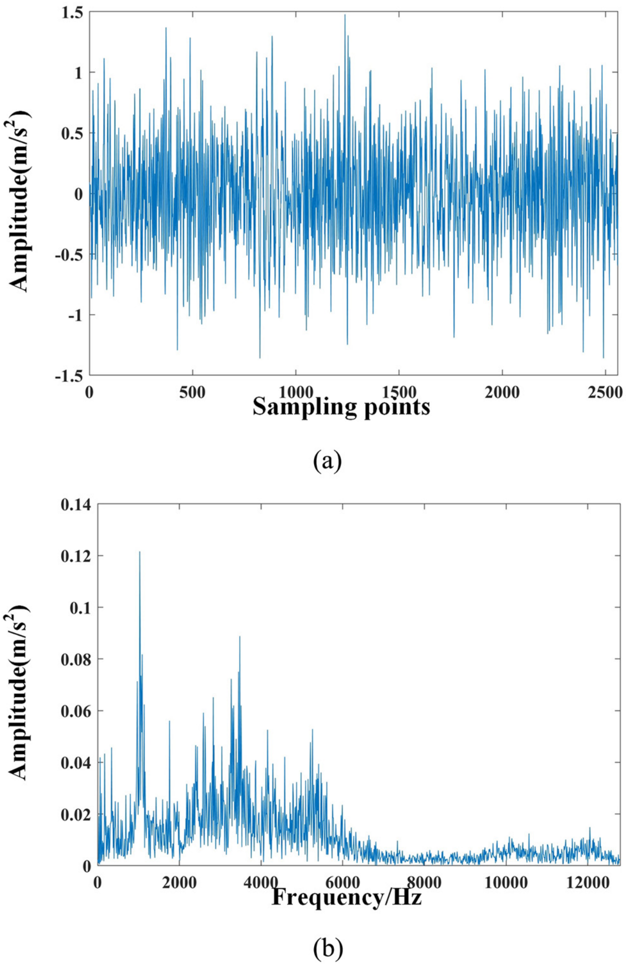


Figure 9: Time- (a) and frequency-domain (b) signals of the bearing at the beginning of the 1 s.

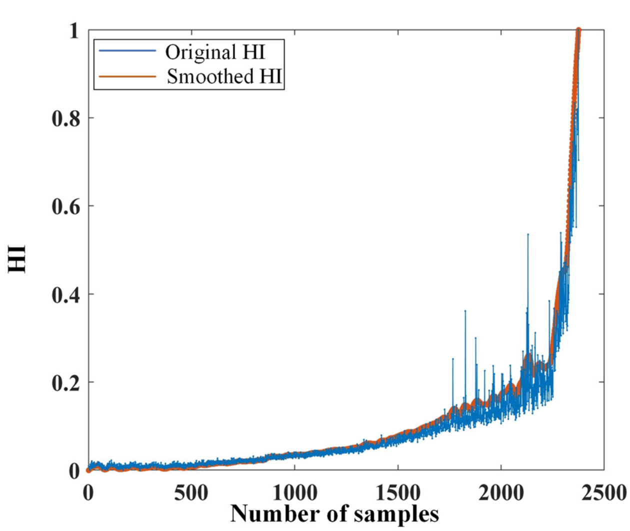


Figure 10: Health indicators constructed by TCNAE.

Table 3: Evaluation indicators for different HIs

Method	Mon	Corr	Rob	CI
TCNAE-HI	0.546	0.931	0.971	0.789
DCAE-HI [24]	0.350	0.970	#	#
SAEwR-HI [25]	0.283	0.994	0.927	0.690

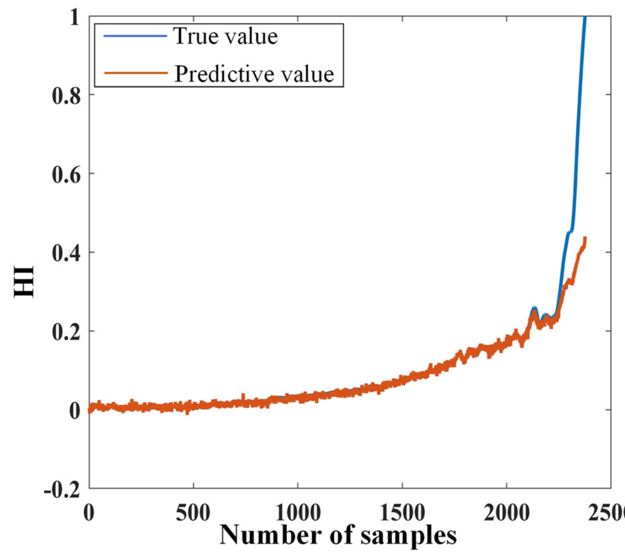


Figure 11: Prediction results of the proposed method in this article.

Table 4: Comparison results of different prediction methods

Method	MAE	RMSE
CNN–Transformer	0.0174	0.1091
BiLSTM [26]	0.0447	0.1207
GDAU [27]	#	0.1870

compared, such as the BiLSTM method proposed in reference [26] and the gated dual attention unit (GDAU) method proposed in reference [27]. The comparison results are shown in Table 4.

From Table 4, it can be concluded that although references [26,27] can achieve good predictive performance indicators, the predictive evaluation indicators are both greater than the prediction methods proposed in this article. The possible reason is that the unsupervised TCNAE construction HI method used in this article has good results in monotonicity, correlation, and robustness, which is beneficial for improving the prediction accuracy of subsequent lifespans. Li *et al.* [26] used the kernel principal component analysis dimensionality reduction method to construct HI, while Qin *et al.* [27] directly used RMS as HI. Although the HI construction method mentioned in the above literature is simple, it cannot comprehensively reflect the degradation status of bearings. This article uses the first 70% of HI (1,663 sets of data) as the training set to predict the remaining 30% of HI values. Its training set is smaller than the 1,900 sets of data in the study by Li *et al.* [26], and on the contrary, it achieved good prediction results, verifying the superiority of the CNN–transformer network.

4 Conclusion

This article proposes a two-stage RUL prediction framework. The first stage is to use TCNAE for unsupervised HI construction. The second stage is to use the constructed HI to train the CNN–transformer model for life prediction. The proposed methods have been validated on both public and experimental datasets. The experimental results show that the unsupervised HI method proposed in this article breaks away from the prior knowledge of traditional HI, and its performance evaluation indicators such as monotonicity, correlation, and robustness are better than existing methods. The proposed CNN–transformer prediction method effectively improves the problems of low accuracy in traditional life prediction methods, and through comparison with existing methods, it is concluded that the method proposed in this article has superior performance in prediction. Therefore, the method proposed in this article has important engineering value and promotional value in RUL prediction.

Funding information: This work is supported by the National Natural Science Foundation of China (grant no. 71871220) and the Hebei Province Key Research and Development Program Project (grant no. 22351702D). The support is gratefully acknowledged. The authors would also like to thank the reviewers for their valuable suggestions and comments.

Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Conflict of interest: The authors state no conflict of interest.

Data availability statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

[1] Euldji R, Bouamhdi M, Rebhi R, Bachene M, Ikumapayi OM, Al-Dujaili AQ, et al. Optimizing condition monitoring of ball bearings: An integrated approach using decision tree and extreme learning machine for effective decision-making. Open Phys. 2023;21(1):20220239.
[2] Lu B, Chen B, Wang D, Li C, Gao R, Liu Y, et al. A bidirectional recursive gated dual attention unit based RUL prediction approach. Eng Appl Artif Intell. 2023;120:105885.
[3] Qin Y, Yang J, Zhou J, Pu H, Mao Y. A new supervised multi-head self-attention autoencoder for health indicator construction and similarity-based machinery RUL prediction. Adv Eng Inform. 2023;56:101973.
[4] Malhi A, Yan R, Gao RX. Prognosis of defect propagation based on recurrent neural networks. IEEE Trans Instrum Meas. 2011;60(3):703–11.
[5] Widodo A, Yang BS. Application of relevance vector machine and survival probability to machine degradation assessment. Expert Syst Appl. 2011;38(3):2592–9.
[6] Wu C, Feng F, Wu S, Jiang P, Wang J. A method for constructing rolling bearing lifetime health indicator based on multi-scale convolutional neural networks. J Braz Soc Mech Sci Eng. 2019;41(11):526.
[7] Zhang D, Stewart E, Ye J, Entezami M, Roberts C. Roller bearing degradation assessment based on a deep MLP convolution neural network considering outlier regions. IEEE Trans Instrum Meas. 2019;69(6):2996–3004.
[8] Lotfy K, El-Bary AA. Magneto-photo-thermo-microstretch semiconductor elastic medium due to photothermal transport process. Silicon. 2022;14:4809–21.
[9] Mahdy AMS, Lotfy K, El-Bary A, Sarhan HH. Effect of rotation and magnetic field on a numerical-refined heat conduction in a semiconductor medium during photo-excitation processes. Eur Phys J Plus. 2021;136:553.
[10] Wang C, Jiang W, Yang X, Zhang S. RUL prediction of rolling bearings based on a DCAE and CNN. Appl Sci. 2021;11(23):11516.
[11] Zhou J, Qin Y, Chen D, Liu F, Qian Q. Remaining useful life prediction of bearings by a new reinforced memory GRU network. Adv Eng Inform. 2022;53:101682.
[12] Que Z, Jin X, Xu Z. Remaining useful life prediction for bearings based on a gated recurrent unit. IEEE Trans Instrum Meas. 2021;70:1–11.
[13] Zhang Z, Farnsworth M, Song B, Tiwari D, Tiwari A. Deep transfer learning with self-attention for industry sensor fusion tasks. IEEE Sens J. 2022;22(15):15235–47.
[14] Jiang L, Zhang T, Lei W, Zhuang K, Li Y. A new convolutional dual-channel Transformer network with time window concatenation for remaining useful life prediction of rolling bearings. Adv Eng Inform. 2023;56:101966.
[15] Lea C, Vidal R, Reiter A, Hager GD. Temporal convolutional networks: A unified approach to action segmentation//Computer Vision–ECCV 2016 Workshops: Amsterdam, The Netherlands, October 8–10 and 15–16, 2016. Proceedings, Part III 14. Springer International Publishing; 2016. p. 47–54.
[16] Qiu H, Niu Y, Shang J, Gao L, Xu D. A piecewise method for bearing remaining useful life estimation using temporal convolutional networks. J Manuf Syst. 2023;68:227–41.
[17] Le CY, Bottou L. Gradient-based learning applied to document recognition. Proc IEEE. 1998;86(11):2278–324.

[18] Zhou Y, Wang Z, Zuo X, Zhao H. Identification of wear mechanisms of main bearings of marine diesel engine using recurrence plot based on CNN model. Wear. 2023;520:204656.
[19] Wen Y, Xu P, Li Z, Xu W, Wang X. RPConvformer: A novel Transformer-based deep neural networks for traffic flow prediction. Expert Syst Appl. 2023;218:119587.
[20] Chen D, Qin Y, Wang Y, Zhou J. Health indicator construction by quadratic function-based deep convolutional auto-encoder and its application into bearing RUL prediction. ISA Trans. 2021;114:44–56.
[21] Zhan X, Liu Z, Yan H, Wu Z, Guo C, Jia X. A novel method of health indicator construction and remaining useful life prediction based on deep learning. Eksploat Niezawodn – Maint Reliab. 2023;25(4):171374. doi: 10.17531/ein/171374.
[22] Javed K, Gouriveau R, Zerhouni N, Nectoux P. Enabling health monitoring approach based on vibration data for accurate prognostics. IEEE Trans Ind Electron. 2014;62(1):647–56.
[23] Zhou J, Qin Y, Chen D, Liu F, Qian Q. Remaining useful life prediction of bearings by a new reinforced memory GRU network. Adv Eng Inform. 2022;53:101682.
[24] Wang C, Jiang W, Yang X, Zhang S. RUL prediction of rolling bearings based on a DCAE and CNN. Appl Sci. 2021;11(23):11516.
[25] She D, Jia M, Pecht MG. Sparse auto-encoder with regularization method for health indicator construction and remaining useful life prediction of rolling bearing. Meas Sci Technol. 2020;31(10):105005.
[26] Li ZE, Zhang XY, Hu W, Xin LR. Evaluation and prediction of high-speed shaft bearings in wind turbines based on health index. Acta Energ Sol Sin. 2021;42(10):290–7.
[27] Qin Y, Chen D, Xiang S, Zhu C. Gated dual attention unit neural networks for remaining useful life prediction of rolling bearings. IEEE Trans Ind Inform. 2020;17(9):6438–47.