machines

MDPI

Article

Prediction of Residual Life of Rolling Bearings Based on Multi-Scale Improved Temporal Convolutional Network (MITCN) Model

Keru Xia [1], Qi Li [2], Luyuan Han [3], Zhaohui Ren [2,*] and Hengfa Luo [2]

[1] School of Mechanical and Power Engineering, Shenyang University of Chemical Technology, Shenyang 110142, China; abc514863@126.com
[2] School of Mechanical Engineering and Automation, Northeastern University, Shenyang 110819, China; qli@mail.neu.edu.cn (Q.L.); lhf1456428108@163.com (H.L.)
[3] School of International Exchange, Guangzhou Maritime University, Guangzhou 510725, China; m17745196673@163.com
* Correspondence: zhhren_neu@126.com

Abstract: The method based on convolution neural networks (CNNs) has been widely developed and applied to residual life prediction, and many excellent results have been achieved. However, CNN models can only learn feature information relative to size, and it is difficult to extract complex time series features from data of long time series. In addition, the existing models still have some problems, such as capturing the correlation of each time series and generating a large amount of redundant information. In order to alleviate the above problems, this study proposes a residual life prediction method of rolling bearings based on a multi-scale improved temporal convolutional network (MITCN) model. It is used to solve problems such as the low accuracy of bearing life prediction and the difficulty of the temporal convolutional network (TCN) model to capture the correlation of each time series. The model adopts the framework of a time convolution network and has good ability to extract time series information. By introducing a multi-scale expanded causal convolution residual structure, improved temporal convolutional network (ITCN) modules with different expansion factors capture information on different time scales and combine soft threshold functions and channel attention mechanisms to adaptively generate thresholds and eliminate redundant information. Finally, the carbon border adjustment mechanism (CBAM) is an attention mechanism used to enhance useful features and suppress useless features, so as to realize the effective fusion of multi-scale features. The IEEE PHM 2012 challenge data set is hereby used to verify the proposed method, which can effectively solve the problem of the low prediction accuracy of the remaining life of bearings.

Keywords: convolution neural network; residual life prediction; MITCN model; attention mechanism; residual structure

check for updates

Academic Editor: Angelos P. Markopoulos

Received: 8 January 2025
Revised: 7 February 2025
Accepted: 10 February 2025
Published: 11 February 2025

Citation: Xia, K.; Li, Q.; Han, L.; Ren, Z.; Luo, H. Prediction of Residual Life of Rolling Bearings Based on Multi-Scale Improved Temporal Convolutional Network (MITCN) Model. *Machines* **2025**, *13*, 137. https://doi.org/10.3390/machines13020137

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Rolling bearing is an indispensable component of rotating machinery, and the estimation of its remaining service life plays a vital role in the stable operation and production safety of equipment [1]. Actual data show that about 50% of bearing damage is attributed to insufficient lubrication, which leads to a significant rise in internal bearing temperature and vibration [2,3]. In order to avoid danger, it is necessary to accurately predict the remaining service life of rolling bearings [4]. Traditional fault diagnosis methods, such as the physical

model [5], dynamic model prediction [6], and the Bayesian method [7], rely too much on manual experience and a theoretical basis and have a limited application scope and weak generalization abilities [8]. At present, remaining useful life (RUL) prediction technology is mainly divided into two methods: model-driven and data-driven methods [9].

The model-driven prediction method relies on the knowledge of a mechanical system's operation rule and fault mechanism and simulates the equipment's degradation process by mathematical and statistical models so as to achieve the effective estimation of bearing RUL [10–12]. The key premise is to select the mathematical model matching with the target system [13]. Yakout et al. [14] selected bearings with the same model and size and analyzed their dynamic characteristics. The experimental results show that the internal radial clearance has a significant impact on the dynamic characteristics of bearings. Cui et al. [15] proposed a comprehensive nonlinear dynamic model and simulated the bearing degradation process based on similarity theory. The degradation process was clearly divided into four stages: the steady-state stage, the defect formation stage, the defect diffusion stage, and the final damage stage. Wang et al. [16] proposed a nonlinear random coefficient regression (RCR) model, which updated the above information regarding random coefficients in real time under the Bayesian framework and calculated the RUL probability density function under specific failure threshold constraints, which effectively improved the RUL prediction accuracy compared to traditional Bayesian methods. Lei et al. [17] used the weighted minimum quantization error to fuse multi-feature information and accurately map it to the degradation process of bearings and then used the maximum likelihood estimation method to improve the accuracy of bearing RUL prediction. Although the model-driven method is very convincing in theory in predicting the RUL of mechanical equipment, the key to its success lies in whether it can build an accurate model reflecting the actual working state of the equipment. The limitation of this method is mainly reflected in its high sensitivity to the complex dynamic characteristics of rotating machinery, changeable operating environment, and external disturbances.

The prediction data-driven method provides an alternative scheme for predicting the RUL of rotating machinery and other equipment without relying on complex mathematical models. In RUL prediction, deep learning networks represented by CNNs and recurrent neural networks (RNNs) are most widely used [18–20]. Sloukia et al. [21] used the support vector machine (SVM) model to establish the correlation between the current and future state of bearings to predict their remaining life. Yan et al. [22] evaluated the bearing degradation stage through the dimensionless measurement index and trained the SVM classifier based on the measured values fitted by the generalized degradation model. Then, according to the classification results, a hybrid degradation tracking model was used to track the degradation process of bearings, thus optimizing RUL prediction. Rohani et al. [23] proposed a method to find a suitable parent wavelet, extract the required features from the decomposed wavelet coefficients, use the Weibull risk function and the exponential function for smooth fitting, and then feed this into the artificial neural network for RUL prediction. Zhang et al. [24] extracted spatial and temporal information from historical data with the help of 1D-CNN and the Bi-gated recurrent unit (BiGRU) hybrid model. The experimental results showed that the proposed parallel hybrid network could effectively predict the RUL. Guo et al. [25] proposed an RUL prediction method based on CNN-Bi-directional long short-term memory (BiLSTM)-Bootstrap model, which adaptively extracted vibration signal features and quantified the RUL prediction interval of rolling bearings by CNN and BiLSTM. Wang et al. [26] used a multi-scale convolution neural network to learn different receptive field features by expanding convolution blocks and improve the RUL prediction accuracy. The results showed that it was superior to the traditional one-dimensional convolution method. Dong et al. [27] proposed an RUL prediction method based on the

MSGCNN-TR model, which effectively extracted the spatio-temporal characteristics of rolling bearing life data through the depth convolution path with residual error and the gated cycle unit path. Jiang et al. [28] fused signal features using the Wasserstein distance, extracted weighted features through the double-attention mechanism and multi-scale convolution, and achieved RUL prediction, although the CNN was limited in terms of the insufficient capture of dynamic characteristics in time series data processing. To solve this problem, Bai et al. [29] proposed TCN, which used dilated causal convolution and residual learning to avoid future information leakage, and achieved remarkable results in the RUL prediction of rotating machinery through strict time constraints. He et al. [30] proposed a new RUL prediction framework based on TCN, which integrated a self-encoder and a self-calibration module, improved the ability of feature fusion and local information processing, and significantly improved the prediction accuracy and efficiency. Deng et al. [31] designed a multi-scale extended causal convolution residual block MsDCCRB and embedded SE units to adaptively calibrate feature responses, enhance feature learning ability, and directly use historical state monitoring data to achieve accurate RUL prediction.

With the development and progress of digital-twin technology, relevant enterprises and researchers began to consider the gradual integration of digital-twin technology and predictive maintenance [32]. Using virtual and real interactions and the real-time fusion characteristics of digital-twin technology, the prediction, early warning, and online diagnosis of the running state of bearings are of great significance for the efficient and reliable operation of equipment.

In summary, it is very common to predict the residual life of rolling bearings by introducing an improved CNN. Nevertheless, there are still some challenges in the existing research. Traditional CNNs have a fixed size of the convolution kernel. Convolution kernels with different sizes can only learn feature information relative to their size, so it is difficult to extract complex time series features from data derived from long time series. In addition, causal dilation convolution uses causality for filling, which produces a lot of redundant information.

In order to solve these challenges, this study proposes a residual life prediction method for rolling bearings based on the MITCN model. Firstly, time convolution, which has a better ability to extract time series information than other forms, is adopted as the network framework. Then, a new multi-scale extended causal convolution residual structure is proposed, which captures information on different time scales and adds a soft threshold function and a channel attention-adaptive threshold to eliminate redundant information. Finally, the CBAM form of the attention mechanism is introduced to enhance useful feature information and suppress useless feature information. The specific contributions of this study are as follows:

- A new multi-scale extended causal convolution residual structure is proposed, which uses several ITCN modules with different expansion factors to capture information on different time scales;
- We constructed an MITCN to predict the residual life of rotating machinery, adding a soft threshold function and a channel attention-adaptive generative threshold to eliminate redundant information, alongside the CBAM form of the attention mechanism to enhance useful feature information and suppress useless feature information;
- The effectiveness and superiority of the proposed method are hereby verified using bearing data sets.

The rest of this paper is structured as follows: The Section 2 introduces the specific details of the proposed method in detail. The Section 3 introduces the data processing and experimental setup. In the Section 4, the performance of the proposed method is verified using bearing data sets. The Section 5 draws the conclusion.

2. Methodology

This section describes the details of the proposed method. The overall residual life prediction scheme is shown in Figure 1. ITCN, CBAM, and MITCN were introduced. ICTN was the backbone of MITCN, and CBAM was the attention mechanism. Multiple ITCN modules with different size expansion factors were used to extract time series features of different scales in parallel. After feature fusion, the CBAM form of the attention mechanism was used to enhance the model's attention to key features and suppress invalid features to effectively predict the remaining life of bearings.

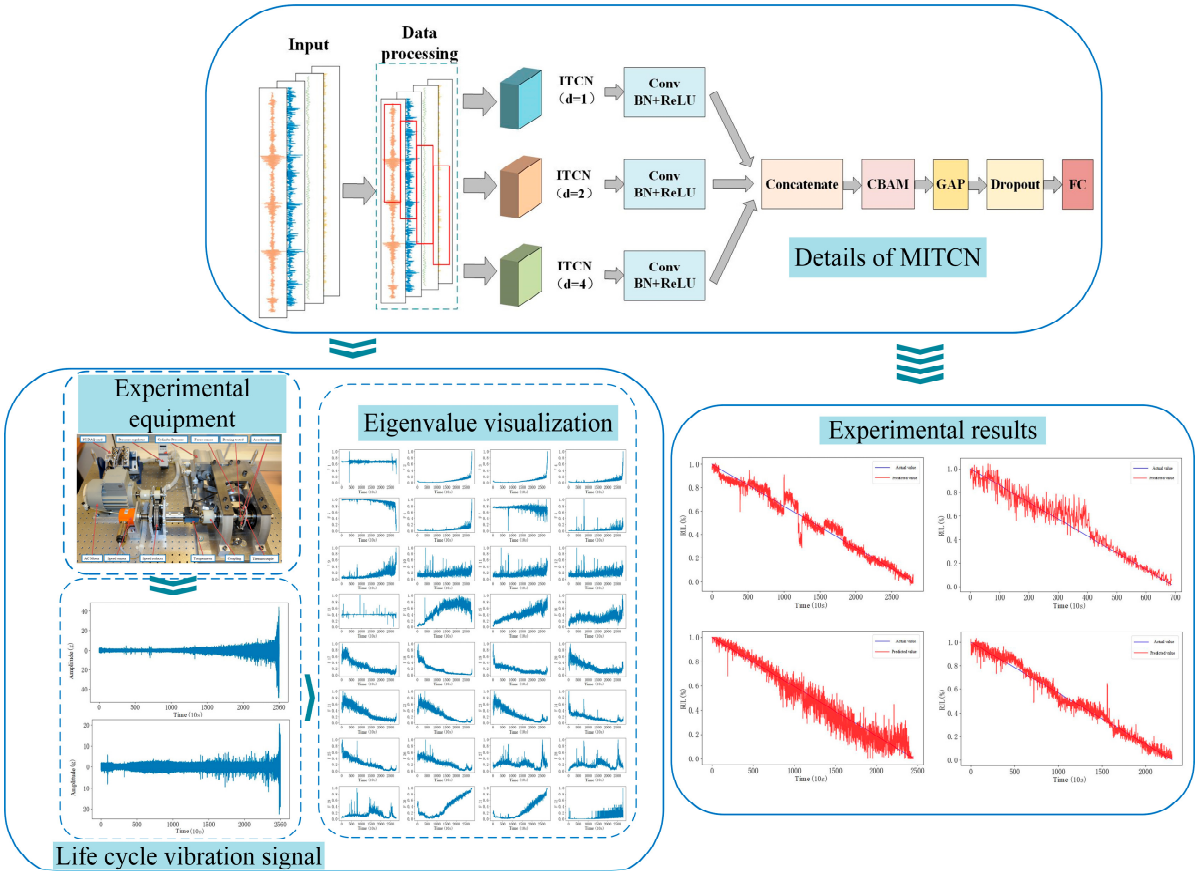


Figure 1. Frame of residual life prediction based on the MITCN model.

2.1. Details of ITCN

The ITCN module consisted of an extended causal convolution, a WeightNorm layer, an attention soft threshold layer, an ReLU activation function, and Dropout. Through residual connection and 1×1 convolution, the input and output dimensions were consistent, and the model's performance and feature learning ability were improved. The deep ITCN module was constructed by continuously stacking two layers of extended causal convolution. The network structure combination of extended causal convolution and residual connection could effectively improve the robustness and feature learning ability of the ITCN module. At the same time, a soft threshold was introduced into the network structure as a nonlinear transformation layer to effectively eliminate redundant information related to noise. Soft thresholds were dynamically generated by a specific subnetwork that adaptively calculated a set of thresholds for each vibration signal. In the ITCN module, the signal after two layers of extended causal convolution processing did not directly enter the next layer but first entered a specific subspace network. This subspace network was essentially a channel attention mechanism, which assigned different

weights to each feature channel to determine their respective noise reduction thresholds. The processed soft threshold signal was fused with the primary feature before it was sent to the following layer network. Adaptive noise reduction in the soft threshold layer and the channel attention mechanism were used to distribute weights in the subspace network. The WeightNorm layer and the ReLU activation function were placed before the convolution to stabilize the network and accelerate training, while the Dropout layer reduced the risk of over-fitting. This design was helpful to learn deep features, prevent over-fitting, and ensure training stability. Figure 2 shows the proposed ITCN module.

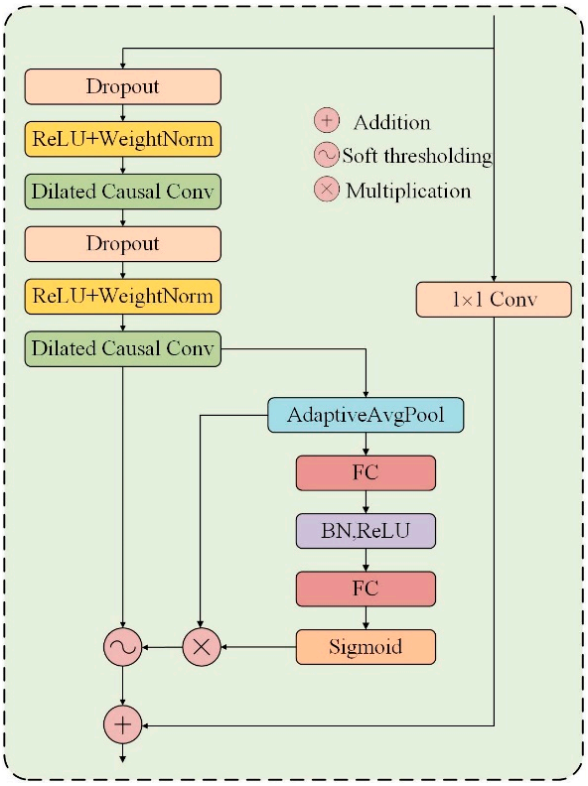


Figure 2. Details of ITCN.

2.2. Details of CBAM

CBAM was a kind of attention module combining channel and space, capable of enhancing useful features and suppressing useless features in the feature map, thus improving the ability of the network to extract implicit important features [33]. At the same time, CBAM was a plug-and-play simple attention block. Figure 3 is a schematic diagram of the CBAM type of attention modules. The two-channel attention module calculated the attention weight map through the channel and spatial attention modules and redefined the feature map. The output feature was activated by a sigmoid function, thus generating the final output result, via the following calculation process [34]:

$$\begin{aligned} M_c(F) &= \sigma(\mathrm{MLP}(\mathrm{AvgPool}(F)) + \mathrm{MLP}(\mathrm{MaxPool}(F))) \\ &= \sigma\left(W_1\left(W_0\left(F^c_{\mathrm{avg}}\right)\right) + W_1(W_0(F^c_{\mathrm{max}}))\right) \end{aligned} \tag{1}$$

where F represents the input feature, $M_c(F)$ represents the output, AvgPool and MaxPool indicate global average pooling and maximum pooling operations, MLP represents the multi-layer perceptron, and $\sigma(\cdot)$ represents the sigmoid activation function.

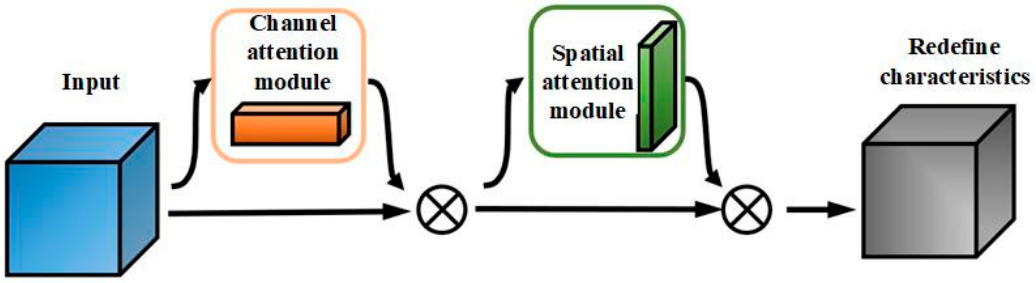


Figure 3. CBAM type of attention modules.

The spatial attention module took the output of the channel attention module as its own input and generated a maximum/average pooled feature using the maximum/average pooled operation. After this mosaic process, the spatial attention feature map was obtained by the convolution layer. The expression is as follows [34]:

$$\sigma\left(f^{7\times7}([\text{AvgPool}(F);\text{MaxPool}(F)])\right) = \sigma\left(f^{7\times7}\left(\left[F^s_{\text{avg}};F^s_{\text{max}}\right]\right)\right) \quad (2)$$

where $f^{7\times7}$ represents a convolution operation with a convolution kernel of 7 and [AvgPool(F); MaxPool(F)] represents the fact that the average and maximum pooling results are concatenated along the channel axis.

2.3. Details of MITCN

As shown in Figure 4, the MITCN module used several ITCN modules with different size expansion factors to extract time series features of different scales in parallel. After feature fusion, the CBAM type of attention mechanism was adopted to enhance the attention paid by the model to key features and suppress invalid features. The expansion coefficient d was usually set to an exponential power of 2, and the receptive field of the network expanded exponentially with the increase in the number of layers of the network. A branch of the ITCN module with multiple expansion factors (d = 1, 2, 4) was designed in the model. The expansion factor of d = 1 corresponded to the most basic time correlation, and the expansion rate of d = 4 enabled the model to explore time dependence over a longer range. Each ITCN module was followed by convolution layers of 3 and 1, and padding was set to 1 for further refinement of feature extraction. The activation function was ReLU. By fusing ITCN modules from different expansion factors and convolving the output, the model could integrate information across time scales. This cascading approach ensured that features at different time scales would be taken into account, providing a comprehensive representation of the features for subsequent attention mechanisms and decision levels. CBAM could assign different attention weights by learning the importance of each feature. Applying CBAM after feature fusion could effectively enhance the model's attention to key features and suppress unimportant information. The local average pooling layer (GAP) converted multi-dimensional features into a single-dimensional feature vector, which effectively reduced the dimension of data and prepared for the input of the full connection layer. Then, the Dropout layer randomly "discarded" (temporarily removed) some neurons in the neural network to prevent over-fitting of the network, and the full connection layer transformed the extracted features into the final prediction results. The parameters of MITCN model are shown in Table 1.

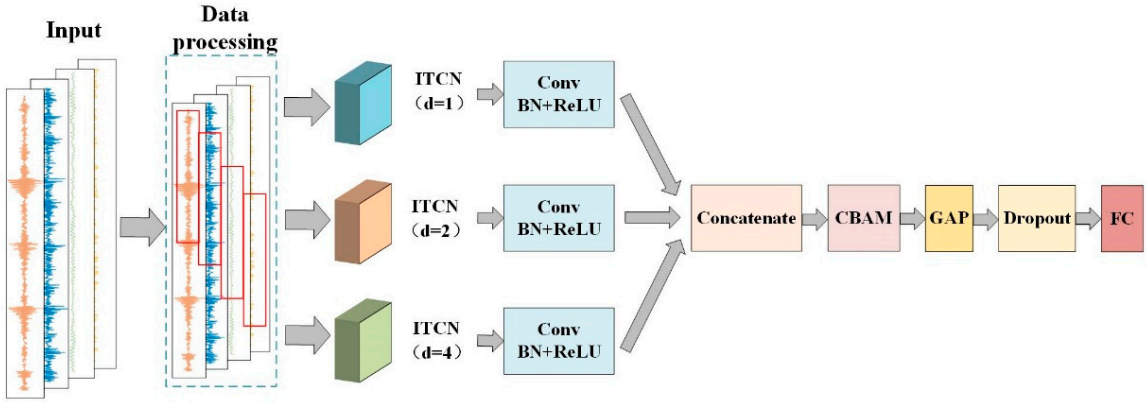


Figure 4. Details of MITCN.

Table 1. The parameters of the MITCN model.

Network Layer	Convolution Kernel	Step Size	Number of Channels	Output Layer
Convolution layer	3 × 3	1	16, 32, 64	(32, 16, 400), (32, 32, 400), (32, 64, 400)
Feature fusion layer	1 × 1	1	64	(32, 64, 400)
RELU	——	——	64	(32, 64, 400)
CBAM	——	——	64	(32, 64, 400)
Global mean pooling layer	1 × 1	1	64	(32, 64, 1)
Dropout layer	——	——	64	(32, 64)
Fully connected layer	——	——	1	(32, 1)
SoftMax classification layer	——	——	1	(32, 1)

3. Experimental Setting

This section describes the details for experimentation and data processing.

3.1. Description of Data Sets

The IEEE PHM 2012 Bearing Challenge data set [35] from the PRONOSTIA testbed was selected to verify the MITCN model-based prediction of rolling bearings' residual life proposed in this paper. PRONOSTIA is an experimental platform specifically designed to test, validate, and confirm methods related to bearing health assessment, diagnosis, and prediction [28]. The accelerated degradation residual life test platform is shown in Figure 5. Sensors were arranged in the vertical and horizontal directions of the bearing for collecting vibration signals. Temperature and acceleration sensors were installed in the vertical and horizontal directions of the bearing to record the relevant data of bearing degradation. The sampling frequencies of the acceleration sensor and temperature sensor were 25.6 kHz and 10 kHz, respectively. The bearing type used in the PRONOSTIA test bench (Besançon, France) was a single volleyball bearing; the specific model was NSK 6804DD (Saint-Cyr-sur-Loire, France). The testbed was developed by the FEMTO-ST Institute (Besançon, France) specifically to test and validate bearing fault detection, diagnosis, and prediction methods. The working condition information related to the experimental bearing is shown in Table 2. The Bearing 1_1 and Bearing 2_1 life cycle horizontal vibration signals are shown in Figure 6. The trend of the Bearing 1_1 acceleration amplitude over its whole life cycle showed that the bearing had experienced a smooth running-in period from 1 s to 15,000 s, during which the bearing could operate normally, usually including the normal operation period and the slight degradation period. Between 15,000 s and 23,000 s, the bearing entered the initial

degradation stage, and although the performance was reduced, it could still meet the needs of normal operation. After more than 23,000 s, the bearing entered a rapid descent stage, until it could not work at all.

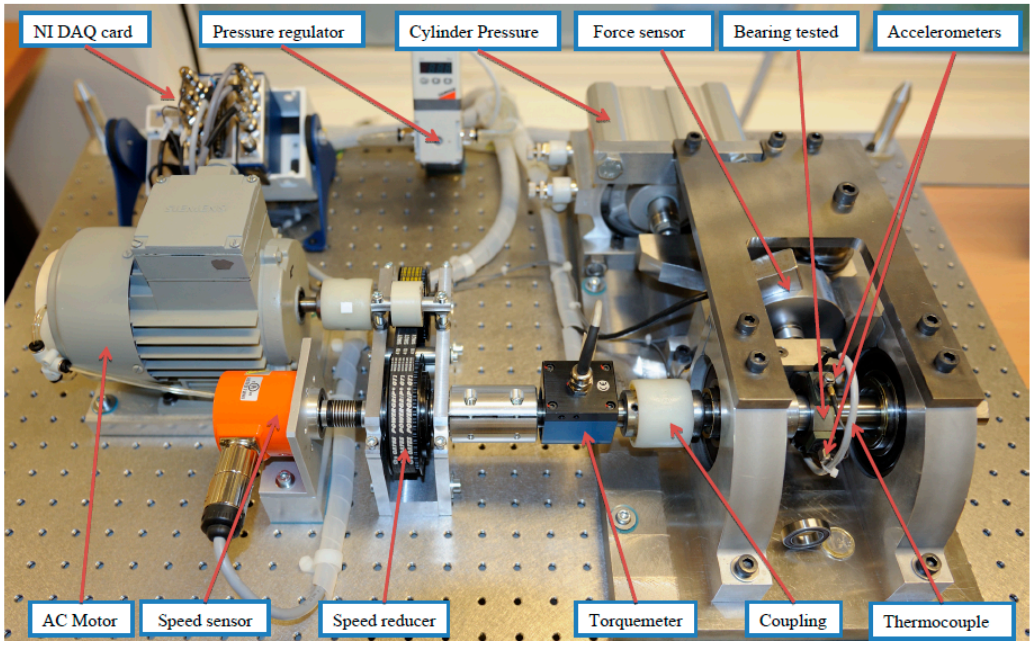


Figure 5. PRONOSTIA bearing life test platform.

Table 2. Test bearing condition information.

Condition	Load (N)	Rotational Speed (r/min)	Training Set	RUL (s)	Testing Set	RUL (s)
A	4000	1800	Bearing 1_1 Bearing 1_2	28,030 8710	Bearing 1_3 Bearing 1_4 Bearing 1_5 Bearing 1_6 Bearing 1_7	23,750 14,280 24,630 24,480 22,590
B	4200	1650	Bearing 2_1 Bearing 2_2	9110 7970	Bearing 2_3 Bearing 2_4 Bearing 2_5 Bearing 2_6 Bearing 2_7	19,550 7510 23,110 7010 2300
C	5000	1500	Bearing 3_1 Bearing 3_2	5150 16,370	Bearing 3_3	4340

3.2. Data Preparation

In order to control the input data, the collected vibration data were preprocessed. A total of 32 key features were extracted from the original data, as shown in Table 3. The horizontal feature visualization of bearing 1_1 is partly shown in Figure 7. F1 through F32 represent visualizations of the 32 different feature indicators in Table 3. In order to optimize the training speed, the features were normalized. Their value range was uniformly adjusted to [0,1] to effectively eliminate the amplitude difference between different features. Different time domain features had different sensitivities to various faults or their degradation degree. For example, kurtosis usually easily revealed early faults, while the root mean square value could reflect the long-term degradation degree. The use of time–frequency domain features allowed us to comprehensively consider information over the whole spectrum range, rather

than relying only on a few specific frequency bands to avoid information loss. Trigonometric features provided more accurate trend analyses by converting input signals to different scales. Because the frequency resolution of vibration signal was low, this method did not involve the extraction of frequency domain features.

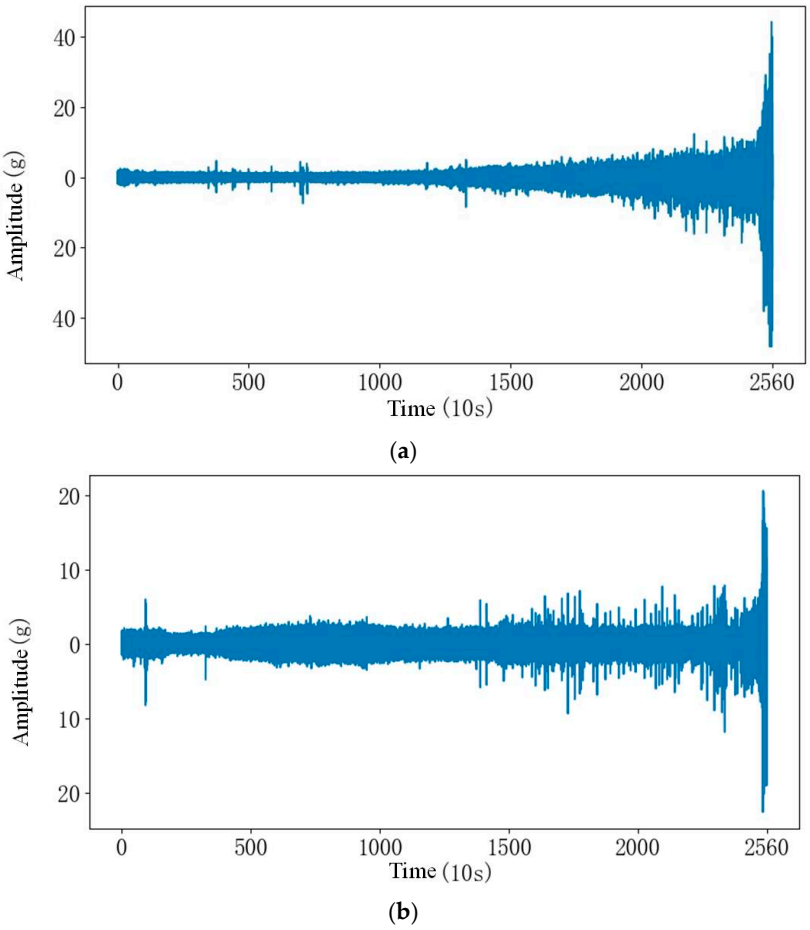


Figure 6. PRONOSTIA bearing life test platform: (**a**) Bearing 1_1 and (**b**) Bearing 1_1 horizontal life cycle vibration signals.

Table 3. Feature indicator.

Type	Characteristic		
Time domain feature	F1: Average F4: Maximum value F7: Skewness F10: Crest factor F13: Coefficient of variation	F2: Root mean square value F5: Minimum value F8: Kurtosis F11: Pulse factor	F3: Standard deviation F6: Peak–peak value F9: Waveform factor F12: Margin factor
Time–frequency domain characteristics	F14–F29: Sixteen-band energy		
Trigonometric function characteristics	F30: Inverse hyperbolic sine standard deviation	F31: Inverse hyperbolic cosine standard deviation	F32: Standard deviation of inverse hyperbolic tangent

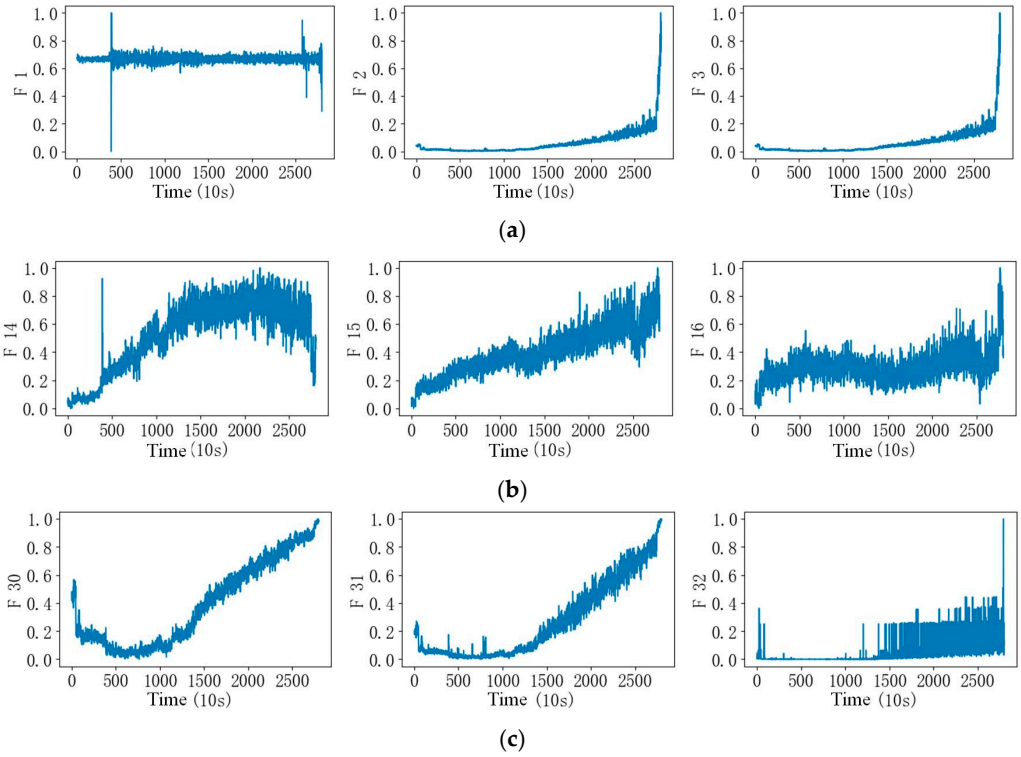


Figure 7. Bearing 1_1 character visualization chart in the horizontal direction: (**a**) time domain features; (**b**) time–frequency domain characteristics; and (**c**) trigonometric function characteristics.

Due to the small number of experimental data samples for testing bearings under Working Condition C in Table 2, this experiment selected bearing vibration data under two different operating conditions, Working Condition A and Working Condition B, dividing them into a training set, a verification set, and a test set. The specific division is shown in Table 4.

Table 4. Data set partitioning.

Experimental Group	Training Set	Verification Set	Test Set
S1	Bearing 1_1, Bearing 1_2, Bearing 1_3, Bearing 1_4, Bearing 1_6	Bearing 1_7	Bearing 1_5
S2	Bearing 2_1, Bearing 2_2, Bearing 2_3, Bearing 2_4, Bearing2_6	Bearing 2_7	Bearing 2_5

3.3. Data Preprocessing

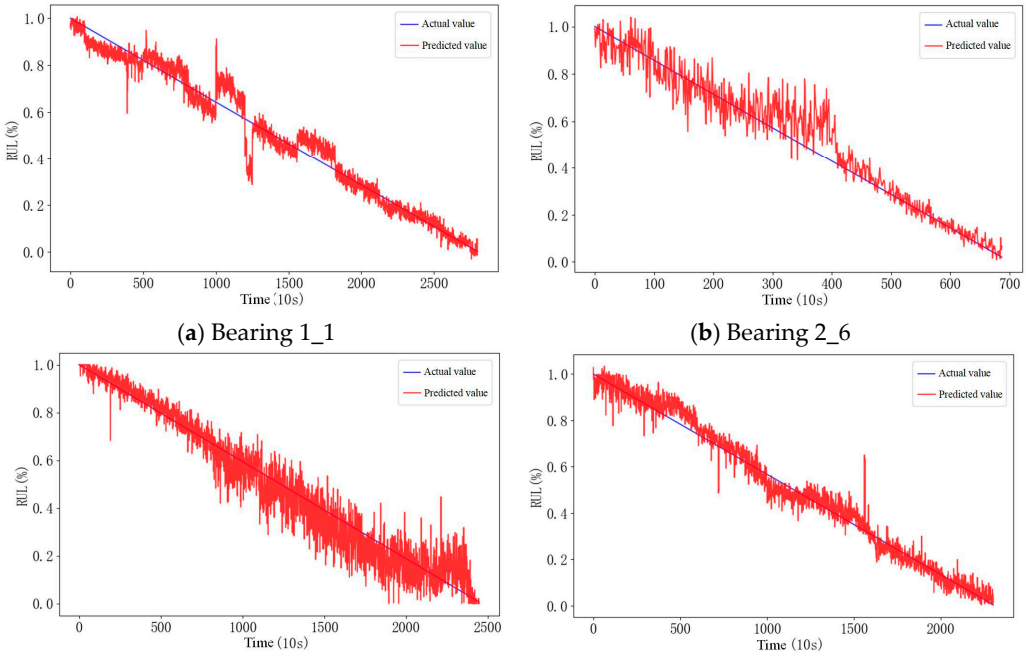
The experiment was implemented using a Windows system with i7 11800H CPU and GTX 3060 GPU. The optimizer for model training was Adam, the learning rate was 0.001, and the loss function was MSE. The epoch was set to 30.

4. Experimental Results and Discussion

4.1. Prediction Effect of the Proposed Method

Figure 8 shows the remaining useful life prediction results for training sets Bearing 1_1 and Bearing 2_6 and test sets Bearing 1_5 and Bearing 2_5 in two experimental groups. The red curve in the figure represents the RUL prediction results of these bearings, and the blue line represents the actual RUL values. Because the accelerated degradation vibration data of these bearings were normalized to the range of 0 to 1 during network training,

when the bearings started to run, their RUL was marked as 1; then, when the bearings ran to the end of failure, their RUL was marked as 0, and the RUL during operation was divided equidistantly. Therefore, the actual RUL value on the prediction result graph below appears as a diagonal line from 1 to 0. It can be observed from the figure that the predicted RUL value fluctuates around the actual RUL value, which has good monotonicity, fits the degradation trend of bearings well, and meets the basic requirements of prediction as a whole.



(**a**) Bearing 1_1 (**b**) Bearing 2_6

(**c**) Bearing 1_5 (**d**) Bearing 2_5

Figure 8. MITCN model's RUL prediction results.

In this study, the following indexes were used to evaluate the model quantitatively: root mean square error (RMSE), mean absolute error (MAE) and coefficient of determination (R). These indexes were helpful to quantitatively analyze the model's prediction accuracy of remaining service life. The results of the output test set's RMSE error, mean absolute error, and determination coefficient are shown in Table 5.

Table 5. Prediction results.

Experimental Group	RMSE	MAE	R^2
S1	0.0834	0.0542	0.9613
S2	0.1253	0.0917	0.9165

The experimental results showed that MITCN model performed well in the test set. In the S1 test set, the values of RMSE and MAE were small, and R was close to 1, which indicated that the model had a high degree of curve fitting. In the S1 experimental group, because the bearing ran for a long time under the actual load of 4000 N, more training data could be collected, which significantly enhanced the training effect of the model. On the other hand, the actual load of bearings in the S2 experimental group increased to 4200 N, which led to the decrease in the speed and running time, thus reducing the amount of data collected. Although the performance of the model in the S1 experimental group was reduced, the results in Table 4 show that the MITCN model trained using the S2 experimental group still had better test results.

4.2. Comparison Experiments with Other Methods

In order to further demonstrate the superiority of this model, it was compared with the TCN, MTCN (multi-scale TCN without the ITCN module), long short-term memory (LSTM) CNN, and transformer models. RMSE, MAE, and R2 were used as evaluation indexes, and Bearing1_5 of the S1 experimental group test set was used for comparative analysis among models. The other parameters were consistent, as shown in Table 6, outlining the model comparison results.

Table 6. Comparison results of different models.

Model	RMSE	MAE	R^2
TCN	0.1678	0.1058	0.9421
MTCN	0.1165	0.0989	0.9498
LSTM	0.2465	0.1963	0.9254
CNN-LSTM	0.2134	0.1574	0.9375
Transformer	0.1007	0.0859	0.8797
MITCN	**0.0834**	**0.0542**	**0.9613**

The MITCN model was the best under every performance index: the MAE and RMSE values of MITCN were the lowest, and the R^2 of MITCN was 0.9% higher than that of the transformer network, which fully shows that this model has certain advantages in predicting the remaining life of bearings.

5. Conclusions

In this study, the authors proposed a residual life prediction method called MITCN. MITCN could not only capture information at different time scales using a multi-scale extended causal convolution residual structure but also adaptively enhance useful feature information and suppress useless feature information through a soft threshold function and an attention mechanism. By paying attention to the correlation of time series, MITCN could deeply mine feature information in the whole life cycle signal and had excellent residual life prediction performance. The key features, including 13 time domain features, 16 time–frequency domain features, and 3 trigonometric function features, were extracted from the original data as inputs. The RUL prediction curve of the trained model on the test set fit well with the actual life curve in the degradation stage. The experimental results showed that the overall effect of MITCN on RMSE, MAE, and R2 was better than that of other comparative models, which proves the feasibility and superiority of MITCN in predicting the remaining service life of rolling bearings. Although MITCN achieved good results, further analysis is still needed in the case of large data sets and lifetime prediction of different research objects. In future research, the authors will adopt the method of data enhancement to solve the problem of residual life prediction with few data samples.

Author Contributions: Formal analysis, Q.L.; Investigation, Z.R.; Data curation, L.H.; Writing—original draft, H.L.; Writing—review & editing, K.X. All authors have read and agreed to the published version of the manuscript.

Funding: The project is supported by the Natural Science Foundation of China (No. 52275091), Natural Science Foundation of Liaoning Province (No. 2022-MS-125), Shenyang Natural Science Foundation (No. 23-503-6-02) and Fundamental Research Funds for the Central Universities (No. N2403005, No. N2303011).

Data Availability Statement: The data presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tang, Y.F.; Xu, F.; Xu, D.L. Life Prediction of Rolling Bearing Based on Stacked Denoising Self-encoder. *Autom. Instrum.* **2024**, *39*, 124–130.
2. Eleonora, D.; Aleksandar, A.; Ljiljana, R.; Palinkaš, I.; Kiss, I. Selection, Dimensioning and Maintenance of Roller Bearings. In Proceedings of the 31st International Conference on Organization and Technology of Maintenance (OTO 2022), Osijek, Croatia, 12 December 2022; pp. 133–142.
3. Alexander, P.; Evgeny, T.; Damjan, S. Temperature Conditions and Diagnostics of Bearings. *Appl. Eng. Lett.* **2023**, *8*, 45–51.
4. Zhou, Y.G.; Zhang, J.C.; Sun, Y.P.; Yu, C.; Zhou, L. Prediction of Remaining Service Life of Rolling Bearings Based on TET and DSRNet-AttBiLSTM. *J. Vib. Shock* **2024**, *43*, 163–173.
5. Yuan, X.; Zhu, Y.S.; Zhang, Y.Y. Fault Diagnosis of Rolling Bearing Based on Model Identification. *Vib. Test Diagn.* **2013**, *33*, 12–17.
6. Liang, Y.; Jia, L.M.; Cai, G.Q. Research on Nonlinear Dynamic Fault Model of Rolling Bearing. *Chin. Railw. Sci.* **2014**, *35*, 98–104.
7. Tong, Z.Y.; Li, J.X.; Qiao, Z.R. Fault Diagnosis of Motor Bearing Based on Improved Bayesian Network. *Electron. Meas. Technol.* **2022**, *45*, 48–55.
8. Hoang, D.T.; Kang, H.J. A survey on deep learning based bearing fault diagnosis. *Neurocomputing* **2019**, *335*, 327–335. [CrossRef]
9. Cheng, C.; Ma, G.; Zhang, Y.; Sun, M.; Teng, F.; Ding, H.; Yuan, Y. A Deep Learning-Based Remaining Useful Life Prediction Approach for Bearings. *IEEE/ASME Trans. Mechatron.* **2020**, *25*, 1243–1254. [CrossRef]
10. Yu, W.; Kim, I.Y.; Mechefske, C. An Improved Similarity-Based Prognostic Algorithm for RUL Estimation Using an RNN Autoen Coder Scheme. *Reliab. Eng. Syst. Saf.* **2020**, *199*, 106926. [CrossRef]
11. Liu, J.; Lei, F.; Pan, C.; Hu, D.; Zuo, H. Prediction of Remaining Useful Life of Multi-Stage Aero-Engine Based on Clustering and LSTM Fusion. *Reliab. Eng. Syst. Saf.* **2021**, *214*, 107807. [CrossRef]
12. Li, X.; Jiang, H.; Liu, Y.; Wang, T.; Li, Z. An Integrated Deep Multiscale Feature Fusion Network for Aeroengine Remaining Useful Life Prediction with Multi-Sensor Data. *Knowl.-Based Syst.* **2022**, *235*, 107652. [CrossRef]
13. Yang, X.; Zheng, Y.; Zhang, Y.; Wong, D.S.-H.; Yang, W. Bearing Remaining Useful Life Prediction Based on Regression Shapalet and Graph Neural Network. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 3505712. [CrossRef]
14. Yakout, M.; Nassef, M.G.A.; Backar, S. Effect of Clearances in Rolling Element Bearings on Their Dynamic Performance, Quality and Operating Life. *J. Mech. Sci. Technol.* **2019**, *33*, 2037–2042. [CrossRef]
15. Cui, L.L.; Wang, X.; Wang, H.Q.; Jiang, H. Remaining Useful Life Prediction of Rolling Element Bearings Based on Simulated Performance Degradation Dictionary. *Mech. Mach. Theory* **2020**, *153*, 103967. [CrossRef]
16. Wang, F.; Tang, S.; Sun, X.; Li, L.; Yu, C.; Si, X. Remaining Useful Life Prediction Based on Nonlinear Random Coefficient Regression Model with Fusing Failure Time Data. *J. Syst. Eng. Electron.* **2023**, *34*, 247–258. [CrossRef]
17. Lei, Y.; Li, N.; Gontarz, S.; Lin, J.; Radkowski, S.; Dybala, J. A Model-Based Method for Remaining Useful life Prediction of Machinery. *IEEE Trans. Reliab.* **2016**, *65*, 1314–1326. [CrossRef]
18. Guo, Y.R.; Mao, J.; Zhao, M. Prediction Method of Bearing Remaining Useful Life Based on CNN and Attention BiLSTM. *J. Shanghai Univ. Eng. Sci.* **2023**, *37*, 96–104.
19. Azyus, A.F.; Wijaya, S.K.; Naved, M. Prediction of Remaining Useful Life Using the CNN-GRU Network: A study on Maintenance Management. *Softw. Impacts* **2023**, *17*, 100535. [CrossRef]
20. Qiu, P.Z.; Wu, D.R.; Xia, Y. Review on The Application of Recurrent Neural Networks in The Prediction of Remaining Service Life. *Mod. Inf. Technol.* **2023**, *7*, 61–66.
21. Sloukia, F.; El Aroussi, M.; Medromi, H.; Wahbi, M. Bearings Prognostic Using Mixture of Gaussians Hidden Markov Model and Support Vector Machine. In Proceedings of the 2013 ACS International Conference on Computer Systems and Applications (AICCSA), Ifrane, Morocco, 27–30 May 2013; pp. 1–4.
22. Yan, M.; Wang, X.; Wang, B.; Chang, M.; Muhammad, I. Bearing Remaining Useful Life Prediction Using Support Vector Machine and Hybrid Degradation Tracking Model. *ISA Trans.* **2020**, *98*, 471–482. [CrossRef] [PubMed]
23. Rohani, B.A.; Aasi, A.; Arghand, H.A. Estimation of Remaining Useful Life of Rolling Element Bearings Using Wavelet Packet Decomposition and Artificial Neural Network. *Iran. J. Sci. Technol. Trans. Electr. Eng.* **2019**, *43*, 233–245. [CrossRef]
24. Zhang, J.; Tian, J.; Li, M.; Leon, J.I.; Franquelo, L.G.; Luo, H.; Yin, S. A Parallel Hybrid Neural Network with Integration of Spatial and Temporal Features for Remaining Useful Life Prediction in Prognostics. *IEEE Trans. Instrum. Meas.* **2022**, *72*, 3501112. [CrossRef]
25. Guo, J.; Wang, J.; Wang, Z.; Gong, Y.; Qi, J.; Wang, G.; Tang, C. A CNN-BiLSTM-Bootstrap Integrated Method for Remaining Useful Life Prediction of Rolling Bearings. *Qual. Reliab. Eng. Int.* **2023**, *39*, 1796–1813. [CrossRef]
26. Wang, R.; Shi, R.; Hu, X.; Shen, C. Remaining Useful Life Prediction of Rolling Bearings Based on Multiscale Convolutional Neural Network with Integrated Dilated Convolution Blocks. *Shock Vib.* **2021**, *2021*, 6616861. [CrossRef]
27. Guo, D.; Cao, Z.; Fu, H.; Li, Z. Remaining Useful Life Estimation for Rolling Bearings Using MSGCNN-TR. *IEEE Sens. J.* **2022**, *22*, 24333–24343. [CrossRef]

28. Jiang, F.; Ding, K.; He, G.; Lin, H.; Chen, Z.; Li, W. Dual-attention-based Multiscale Convolutional Neural Network with Stage Division for Remaining Useful Life Prediction of Rolling Bearings. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 3525410. [CrossRef]
29. Bai, S.; Kolter, J.Z.; Koltun, V. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. *arXiv* **2018**, arXiv:1803.01271.
30. He, K.; Su, Z.Q.; Tian, X.Q.; Yu, H.; Luo, M. RUL Prediction of Wind Turbine Gearbox Bearings Based on Self-Calibration Temporal Convolutional Network. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 3501912. [CrossRef]
31. Deng, F.; Bi, Y.; Liu, Y.; Yang, S. Remaining Useful Life Prediction of Machinery: A New Multiscale Temporal Convolutional Network Framework. *IEEE Trans. Instrum. Meas.* **2022**, *72*, 2516913. [CrossRef]
32. Li, C.B.; Sun, X.; Hou, X.B. On-line monitoring method of tool wear in CNC milling driven by digital twins. *China Mech. Eng.* **2022**, *33*, 78–87. [CrossRef]
33. Yu, Y.; Ma, J.; Wang, X.D.; Zhu, J.; Liu, G. Improved CNN Fault Diagnosis Model Based on Depth Separable Small Convolution Kernel and CBAM. *Electron. Meas. Technol.* **2022**, *45*, 171–178.
34. Woo, S.H.; Park, J.; Lee, J.Y.; Kweon, I.S. CBAM: Convolutional Block Attention Module. In Proceedings of the European Conference on Computer Vision (ECCV) 2018, Munich, Germany, 8–14 September 2018; Volume 11211, pp. 3–19.
35. Benkedjouh, T.; Medjaher, K.; Zerhouni, N.; Rechak, S. Remaining Useful Life Estimation Based on Nonlinear Feature Reduction and Support Vector Regression. *Eng. Appl. Artif. Intell.* **2013**, *26*, 1751–1760. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.