

Article

A Novel Piecewise Cubic Hermite Interpolating Polynomial-Enhanced Convolutional Gated Recurrent Method under Multiple Sensor Feature Fusion for Tool Wear Prediction

Jigang He ¹, Luyao Yuan ², Haotian Lei ², Kaixuan Wang ², Yang Weng ^{2,*} and Hongli Gao ^{1,*} 

¹ School of Mechanical Engineering, Southwest Jiaotong University, Chengdu 610031, China; hjgssg@my.swjtu.edu.cn

² School of Mathematics, Sichuan University, Chengdu 610065, China; yuanluyao@stu.scu.edu.cn (L.Y.); hao@stu.scu.edu.cn (H.L.); eeentity@stu.scu.edu.cn (K.W.)

* Correspondence: wengyang@scu.edu.cn (Y.W.); hongli_gao@swjtu.edu.cn (H.G.)

Abstract: The monitoring of the lifetime of cutting tools often faces problems such as life data loss, drift, and distortion. The prediction of the lifetime in this situation is greatly compromised with respect to the accuracy. The recent rise of deep learning, such as Gated Recurrent Unit Units (GRUs), Hidden Markov Models (HMMs), Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Attention networks, and Transformers, has dramatically improved the data problems in tool lifetime prediction, substantially enhancing the accuracy of tool wear prediction. In this paper, we introduce a novel approach known as PCHIP-Enhanced ConvGRU (PECG), which leverages multiple—feature fusion for tool wear prediction. When compared to traditional models such as CNNs, the CNN Block, and GRUs, our method consistently outperformed them across all key performance metrics, with a primary focus on the accuracy. PECG addresses the challenge of missing tool wear measurement data in relation to sensor data. By employing PCHIP interpolation to fill in the gaps in the wear values, we have developed a model that combines the strengths of both CNNs and GRUs with data augmentation. The experimental results demonstrate that our proposed method achieved an exceptional relative accuracy of 0.8522, while also exhibiting a Pearson's Correlation Coefficient (PCC) exceeding 0.95. This innovative approach not only predicts tool wear with remarkable precision, but also offers enhanced stability.

Keywords: deep learning; cutter wear prediction; Piecewise Cubic Hermite Interpolating Polynomial (PCHIP); Convolutional Neural Network (CNN); Gated Recurrent Unit (GRU)



Citation: He, J.; Yuan, L.; Lei, H.; Wang, K.; Weng, Y.; Gao, H. A Novel Piecewise Cubic Hermite Interpolating Polynomial-Enhanced Convolutional Gated Recurrent Method under Multiple Sensor Feature Fusion for Tool Wear Prediction. *Sensors* **2024**, *24*, 1129. <https://doi.org/10.3390/s24041129>

Academic Editor: Jongmyon Kim

Received: 27 December 2023

Revised: 20 January 2024

Accepted: 25 January 2024

Published: 8 February 2024



Copyright: © 2024 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

In industrial scenarios, when equipment is used for processing, the lifetime and maintenance of the equipment are factors that must be considered. For example, in the context of Computerized Numerical Control (CNC) machine processing, the maintenance of the tool's lifetime is particularly important and has the highest priority. This is due to the necessity for more frequent replacement of severely worn tools, thereby increasing the downtime and maintenance expenses along the production line. Meanwhile, tool wear affects the quality of machined parts, leading to uneven machined surfaces, dimensional inaccuracies, and potential damage to the workpiece. Severely worn tools can even pose a safety hazard to the working environment and the operators. In essence, the prediction of tool wear not only contributes to heightened production efficiency, cost control, and product quality assurance, but also aligns with the trend towards intelligent manufacturing. This progression fosters the development of the manufacturing industry in a direction that is more advanced, sustainable, and intelligent.

The state of cutting tools has an important impact on production efficiency and surface processing quality. Therefore, online monitoring and real-time prediction of tool wear are of

great significance, and they also have become the most discussed and researched hot topic in the mechanical field. Over the years, researchers have explored various methodologies and techniques to predict tool wear, aiming to enhance productivity, optimize the tool lifetime, and minimize machine downtime [1–3]. The earliest monitoring of cutting tool conditions started with a single variable, known as direct measurement, and gradually evolved to fewer variables, known as indirect measurement. For instance, the optical image method was the earliest traditional method applied to tool wear monitoring [4,5]; it uses the reflectance of the worn surface to evaluate the wear of the tool. Contact resistance measurement is performed using electrical resistance and the radioactive elements [6]. However, a single signal has its own drawbacks. While some processes are too complicated, some are not suitable for large workpieces, some will be affected by noise, some signal acquisition will be delayed, and some are expensive (acoustic emission monitoring of the equipment). Therefore, multiple sensor signals are widely used to monitor tool wear. The incorporation of multi-signal conditions, which involves monitoring and analyzing a wide range of parameters including vibration, temperature, acoustic emission, and cutting force, among others, has provided a more-comprehensive understanding of the tool's behavior during machining processes. By considering a multitude of signals, engineers can gain a more nuanced insight into the complex interactions that affect tool wear and failure. This not only leads to more accurate predictions, but also enables proactive maintenance and optimization strategies. Multiple sensor signals mean multiple features, and their fusion starts to become the key [7,8].

In recent years, with the popularity of machine learning and deep learning, new directions have opened up for research on cutting tools, and numerous related studies have sprung up using methods such as Artificial Neural Networks (ANNs) [9,10], Support Vector Machines (SVMs) [11–13], the Hidden Markov Model (HMM) [14–17], Gaussian Process Regression (GPR), etc. [18,19]. With the rise of deep learning, these types of methods have advanced to a new level [20]. In the contemporary landscape of modern manufacturing, the incorporation of multi-signal conditions and the utilization of deep learning in tool lifetime prediction are essential for fostering efficiency, reliability, and competitiveness. Deep learning methods have shown significant promise in tool wear prediction for machining processes due to their ability to automatically learn complex patterns and relationships from large datasets [21]. They have the potential to outperform traditional analytical and empirical models by capturing intricate nonlinearities in the machining process. Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Attention networks, and so on, have skyrocketed in the mechanical field [22–24]. These models, primarily known for their remarkable achievements in areas such as computer vision and language translation, have found substantial relevance in the realm of mechanical production as well, underpinning the evolution of smart manufacturing.

It is important to note that deep learning methods require large, labeled datasets for effective training, which can be a challenge in some machining scenarios. Data augmentation is a technique widely used in machine learning to artificially increase the size of a training dataset by applying various transformations to the original data. This helps improve the model's generalization and robustness. When it comes to machine learning for tool wear prediction, data augmentation can be particularly beneficial in enhancing the model's ability to recognize patterns associated with different states of tool wear. Usually, in the collection of data on tool wear, only the values of the sensor signals (such as the cutting force, vibration, acoustic emission, and current) are collected, but the value of the tool wear is not measured. The main reason is that the signal acquisition sensors are attached to, for example, a CNC machine tool, so they can collect the data at a relatively high frequency, and the amount of wear of the tool is measured after the tool has been used for a constant interval, so the frequency of the obtained data is much lower. Therefore, we need to use data augmentation methods to improve the data availability. Data augmentation is a commonly used technique in machine learning, involving the transformation and expansion of training data to enhance their diversity and richness. It improves the

model's generalization capability, robustness, and accuracy. There are multiple methods available for data augmentation, including random erasing [25], data interpolation [26], and so on. By addressing issues such as overfitting, imbalanced data, missing data, and limited samples, data augmentation effectively enhances the performance and reliability of machine learning models [27].

In the context of tool wear prediction with deep learning, data augmentation refers to the technique of artificially increasing the size and diversity of the training dataset by applying various transformations to the original sensor data collected during machining processes, with the goal of enhancing the generalization and robustness of the deep learning model by exposing it to a wider range of variations and scenarios that may be encountered in real-world tool wear conditions. For tool wear prediction, the input data often consist of sensor readings, such as vibration signals, acoustic signals, current signals, or other sensor data collected during the cutting or machining process. In such operating conditions, tool wear stages, or machining scenarios simulated by introducing variations to sensor data using a data augmentation method, the Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) interpolation method can be used to obtain the uniformly spaced interpolation data series. In a modified grey model proposed by Wang F et al. to predict the RUL of rolling bearings based on vibration data, the PCHIP method was used to process the original data, and this method managed to maintain the trend characteristics of the original signal while improving the reliability of RUL prediction results [28].

The PCHIP-Enhanced ConvGRU (PECG) model we introduced adeptly merges CNN and GRU networks to effectively capture the time series characteristics of the tool wear data. In our study, thanks to the National Natural Science Foundation of China, we were able to utilize real industrial tool data versus synthetic datasets, which lends credence to the model's wear prediction results. In real scenarios, we employed the PCHIP method to interpolate and supplement the wear data, aimed at addressing incomplete tool wear data due to rapid sensor data acquisition. This approach alleviates the issue of high-dimensional but insufficient measurement data obtained from sensor-based tool wear measurements. Notably, PCHIP interpolation substantially elevated the relative prediction accuracy of the model from 0.8005 to 0.8522. Our methodology further involves the extraction of local features via the CNN layer, leveraging the resulting feature map as input for the GRU encoder to capture temporal dependencies. While fully exploiting the time series information processing capabilities of GRU, PECG effectively harnesses the spatial feature learning process of CNN, thereby organically combining and maximizing the strengths of both.

In summary, based on the research trend of multi-feature fusion in the industry, and the advantages of deep learning to mine data, a new PECG method under multiple feature fusion for tool wear prediction has been developed. Our proposed method has the following contributions:

- By employing the Piecewise Cubic Hermite Interpolating Polynomial method in tandem with an understanding of the patterns associated with missing tool wear data, we successfully interpolated and completed the wear data. This approach effectively resolves the challenge posed by high-dimensional tool wear measurement data collected by sensors, a scenario often characterized by relatively insufficient measurement data.
- We extract local features through the CNN layer, leveraging the feature map as input for the GRU encoder to capture temporal dependencies. The PECG model effectively harnesses the spatial feature learning capacity of CNN while fully optimizing the time series data processing abilities of GRU. This results in the seamless integration and maximization of the strengths of both models, making it particularly well-suited for processing data characterized by both time series and spatial features.
- These two aspects are combined to form a comprehensive PECG method.

The remainder of this paper is organized as follows. Section 2 introduces the data interpolation method, PCHIP. Section 3 describes the proposed wear prediction model

in detail. In Section 4, we conduct experimental studies to compare the proposed model with other methods and confirm its superiority. Section 5 provides conclusions. The abbreviations are listed at the end of this paper.

2. PCHIP Interpolation Method

In the data acquisition process, the varying methods of acquiring data have led to a significantly higher volume of sensor data compared to wear data, resulting in a lack of corresponding wear data for certain sensor readings. Consequently, there are missing values within the wear data. Previous approaches involved the deletion of sensor data lacking corresponding wear data, inadvertently discarding valuable information inherent in the sensor data. To address this issue, we have introduced the PCHIP interpolation method to substitute the missing wear data. Through this method, we establish a one-to-one correspondence between sensor data and tool wear data, ensuring the maximization of information encapsulated within the sensor data. This approach enables us to fully leverage the information gleaned from sensor data while circumventing the loss of valuable insights.

There are many interpolation methods. Among them, the simplest method is to define a piecewise linear function between each number of points. The linear method is fast and easy to implement, but linear interpolation does not produce a smooth curve. To solve this problem, a higher-order polynomial can be chosen between each pair of data points, and we can specify the gradient of this polynomial to ensure that the overall approximation function is continuous and has continuous derivatives. Cubic spline interpolation resolves sudden changes in gradients in the case of linear interpolation. But this also introduces a problem that the interpolation may be outside the range of our data point values, which can lead to overshooting issues.

We use Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) to avoid the above two problems. The cubic Hermite polynomial is defined as follows:

$$p(t) = h_{00}(t)p_0 + h_{10}(t)(x_{k+1} - x_k)m_0 + h_{01}(t)p_1 + h_{11}(t)(x_{k+1} - x_k)m_1 \quad (1)$$

where $h_{00}, h_{10}, h_{01}, h_{11}$ are Hermite basis functions. PCHIP interpolates using a piecewise cubic polynomial $P(x)$ with these properties:

- On each subinterval $x_k \leq x \leq x_{(k+1)}$, the polynomial $P(x)$ is a cubic Hermite interpolating polynomial for the given data points with specified derivatives at the interpolation points.
- $P(x)$ interpolates y , that is, $p(x_j) = y_j$, and the first derivative $\frac{dp}{dx}$ is continuous. The second derivative $\frac{d^2p}{dx^2}$ is probably not continuous, so jumps at x_j are possible.
- The cubic interpolant $P(x)$ is shape-preserving. The slopes at x_j are chosen in such a way that $P(x)$ preserves the shape of the data and respects monotonicity. Therefore, on intervals where the data are monotonic, so is $P(x)$, and at points where the data have a local extremum, so does $P(x)$.

These properties of the piecewise cubic polynomial maintain the monotonicity of the points on the interpolation curve [29]. They solve the problem of overshoot and the curve of the interpolation result is smooth at the same time.

3. Model Construction

Data-driven methods predict tool wear using predictive models trained by machine learning or pattern recognition algorithms [30]. When dealing with data-driven works, deep learning is able to learn from large amounts of data and identify subtle patterns and relationships between tool wear value and sensor data.

As shown in Figure 1, the proposed PECCG mainly includes two stages: data preprocessing and model construction. After the data preprocessing, we successfully resolved the problem of missing wear data by employing the PCHIP interpolation technique. The processed data were subsequently utilized to train the proposed model. The details of the model construction are illustrated below.

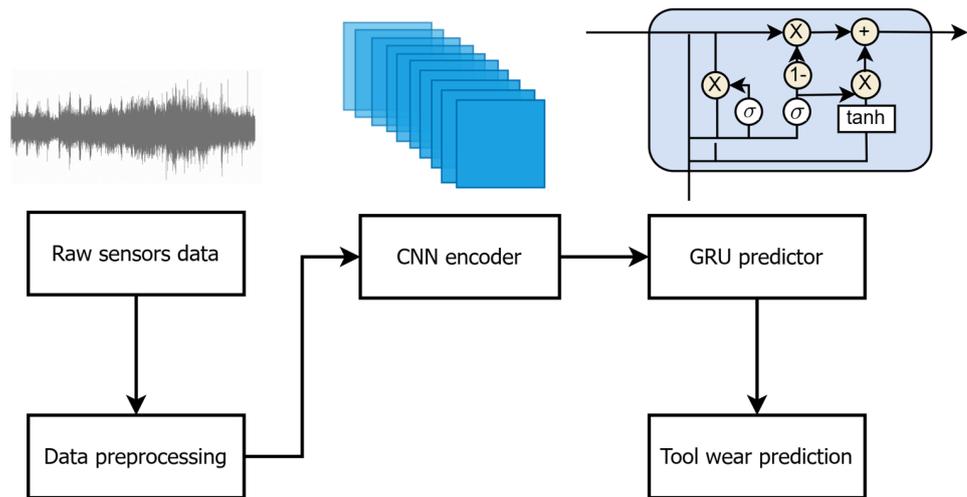


Figure 1. The framework of PECG method.

3.1. Convolutional Neural Network

CNNs are primarily used for image classification tasks and have become dominant in various computer vision tasks, but they can also be used for regression problems. A CNN has five basic layers: convolutional layer, pooling layer, activation layer, fully connected layer, and dropout layer. In this paper, we use a CNN as a feature extractor and pass the features to a GRU. In that case, the CNN in our method incorporates a convolutional layer followed by batch normalization and an activation layer. The equation for this process is as follows:

$$c_{ik} = \text{ReLU}(W_k * x_i + b_k) \quad (2)$$

where W_k indicates the convolutional filter, $*$ denotes the convolution operation, b_k is the bias, and the activation function is ReLU. Here, c_{ik} represents the encoding result, which is the extracted feature we use in the followed GRU.

3.2. Gated Recurrent Unit

The Gated Recurrent Unit (GRU) is a type of Recurrent Neural Network (RNN) architecture that has gained popularity in recent years due to its ability to model sequential data with greater efficiency and accuracy. In this paper, we use a GRU model after the CNN to obtain wear predictions. In a GRU model, there are two gates: an update gate and a reset gate. The update gate determines how much of the previous hidden state should be retained and how much of the current input should be added to the new hidden state, while the reset gate controls how much of the previous hidden state should be ignored. These gating mechanisms allow the GRU model to selectively remember or forget information from the past. Equations for this process are as follows:

$$r_t = \sigma(W_{ir}x_t + b_{ir} + W_{hr}h_{(t-1)} + b_{hr}) \quad (3)$$

$$z_t = \sigma(W_{iz}x_t + b_{iz} + W_{hz}h_{(t-1)} + b_{hz}) \quad (4)$$

$$n_t = \tanh(W_{in}x_t + b_{in} + r_t * (W_{hn}h_{(t-1)} + b_{hn})) \quad (5)$$

$$h_t = (1 - z_t) * n_t + z_t * h_{(t-1)} \quad (6)$$

where h_t is the hidden state at time t , x_t is the input at time t , $h_{(t-1)}$ is the hidden state of the layer at time $t - 1$ or the initial hidden state at time o , and r_t, z_t, n_t are the reset, update, and new gates, respectively. σ is the sigmoid function, and $*$ is the Hadamard product. Then, the result of the hidden state is imported to a fully connected layer and the output is the wear prediction result.

3.3. Model Framework

The framework of PECG is illustrated in Figure 2. And details of our model structure are shown in Table 1. In a CNN, the convolutional layers are used to extract features from the input data. CNN has the ability to capture complex patterns and relationships in the input data. In that case, after analyzing the data, which have high-dimensional sensor data as input, we first use a one-dimensional ten-layer CNN as an encoder to extract features and reduce the dimensionality of the data. The output of the CNN encoder is then imported to a GRU. Finally, the wear prediction is completed through a fully connected layer.

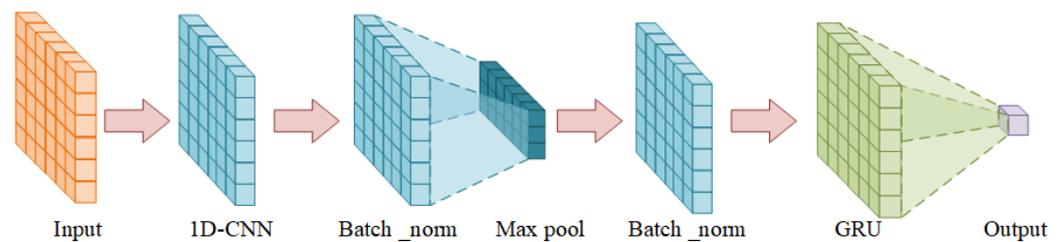


Figure 2. Structure of PECG.

Table 1. Details of model structure.

Layer	Feature Maps	Kernel Size	Parameter Number
CNN-Block_1	128	5	5504
CNN-Block_2-10	128	5	82,304
GRU	384	123	99,072

4. Experiment and Result

4.1. Experimental Conditions

We utilized tool data acquired with support from the National Natural Science Foundation of China, gathered from real industrial settings, as input for the model, rather than relying on virtual datasets available through networks. This approach significantly enhances the credibility of the wear prediction results. The milling cutter under consideration is the APMT1135 carbide cutter, a product of Duracarb. Its fundamental parameters include a tool tip angle of 85 degrees, a blade relief angle of 11 degrees, a blade length of 11 mm, a thickness of 3.5 mm, an inscribed circle diameter of 6.35 mm, and a maximum cutting depth of 9 mm. Figure 3 depicts the actual state of tool wear observed on the machinery.



Figure 3. Real situation of tool wear on machine tools.

There are four types of sensor signal collected during the cutting process: three-way force signals, three-way vibration signals, acoustic signals, and current signals. The sensor devices are shown in Figures 4 and 5. After one cutting path is completed (or after multiple cutting paths are completed), the experimental tool is removed and the wear amount is measured through a visual microscope. The measurement process of tool wear amount is illustrated in Figure 6.

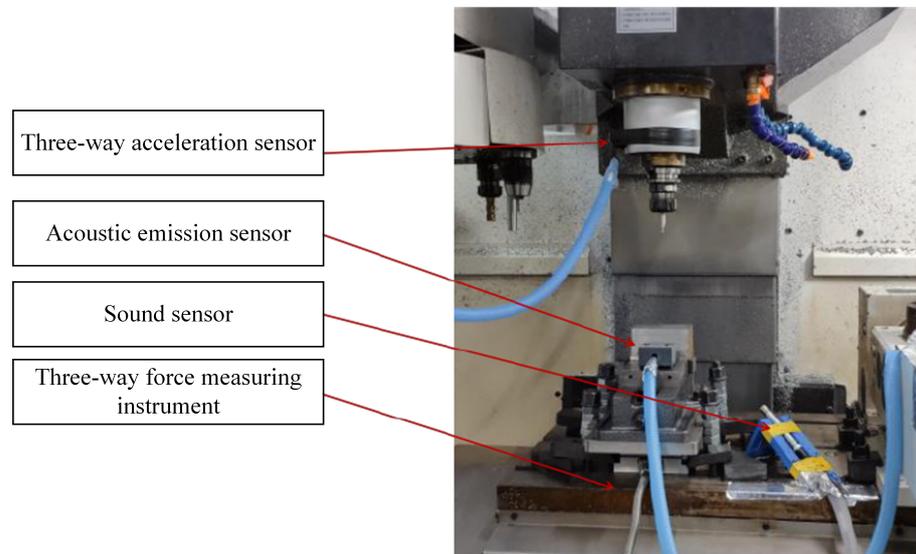


Figure 4. Sensor layout a. PHI Mechatronics Technology Laboratory, Chengdu, China.

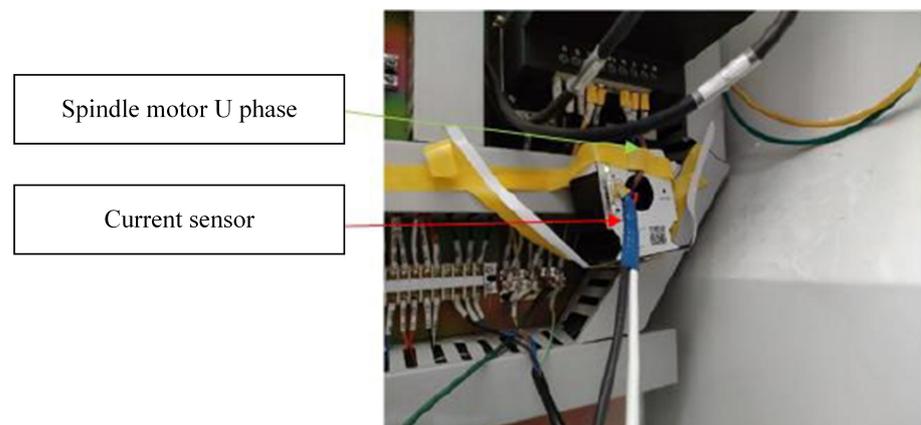


Figure 5. Sensor layout b. PHI Mechatronics Technology Laboratory, Chengdu, China.

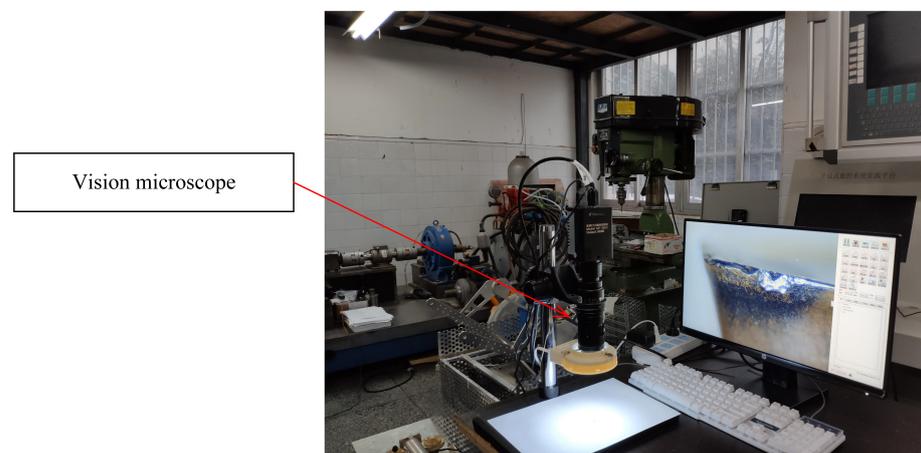


Figure 6. Wear amount collection. PHI Mechatronics Technology Laboratory, Chengdu, China.

The vibration signal is captured using the PCB365A15 three-way acceleration sensor, while the cutting force sensor employed is the KISTLER 9257B three-way load cell. Additionally, the setup includes the Bruel Kjaer's 4966-H-041 acoustic sensor, the PAC-WD acoustic emission sensor, and the POLARISMMI200B (current model: CSA201-P030T01) current sensor. These diverse datasets have been instrumental in supporting the publication of several articles on milling cutter life prediction and intelligent operation in esteemed journals [24,31–34]. Furthermore, these datasets represent the lifecycle patterns observed in carbide cutters.

4.2. Dataset

We carry out basic data cleaning for the collected wear data, which is divided into three parts: standardization, partial correction, and elimination. When we collect these eight types of data, we first standardize them:

$$x_s = \frac{x_i - \bar{x}}{\sigma} \quad (7)$$

where x_s is the standardized data. \bar{x} represents the mean of the data. σ represents the standard deviation of the data.

The reason for this is obvious: to scale the data so that they fall into a small, specific interval. Standardization solves the problem of small difference in working conditions by scaling according to variance. It is often used in some comparison and evaluation index processing to remove the unit restriction of the data and convert it into a dimensionless pure value, so that indicators of different units or magnitudes can be compared and weighted.

The signal drift of the data is shown in Figure 7. The obvious missing and drifting parts of life data monitoring are shown in the red and green boxes, respectively. We use Exponential Moving Average (EMA) to bring significantly drifting segments of the data back into the normal range:

$$v_t = \beta \cdot v_{t-1} + (1 - \beta) \cdot \theta_t \quad (8)$$

where v_t represents the average value of the first t bars ($v_0 = 0$), β is the weighting value (generally set to 0.9–0.999), and θ_t is the standardized data.

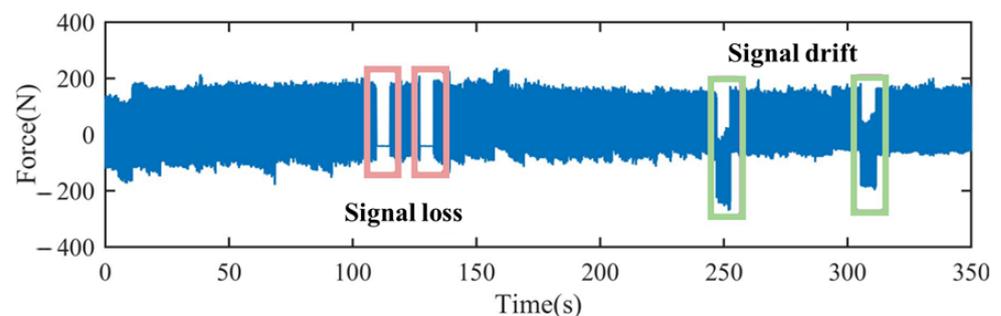


Figure 7. Signal loss and signal drift.

Furthermore, we eliminate obviously abnormal data [35]:

$$\delta = \sqrt{\frac{1}{m-1} \sum_{k=1}^m \left(S_K - \frac{\sum_{k=1}^m S_K}{m} \right)^2} \quad (9)$$

where δ is the abnormal data. The pseudocodes for describing the processes to display the data that cannot be used directly and need to be eliminated are shown in the following Algorithm 1.

Our dataset includes data collected from 28 milling cutters under eight different cutting conditions. The details of the cutting conditions of the cutters are shown in Table 2. In this table, Cm_n is the sign of a cutter, which means no. n cutter under condition m. The signals of cutters C4_3 and C7_9 are used as the test data, and the rest of the cutters are used as the training set. Deep learning makes it possible to involve all signals, making wear prediction accurate and efficient. Single sensor signals often have their own limitations, but deep learning possesses powerful feature learning capabilities. By using a multi-signal variable matrix for prediction, it is possible to extract rich information from multiple features, effectively avoiding the limitations associated with relying on a single feature. We use a total of eight variables: current, force (three directions, x,y,z), sound, and vibration (three directions, x,y,z) to predict the wear process.

Algorithm 1 Signal_Segment (Sig_{org}, L_w, d_w)

Inputs: Sig_{org} —original time-domain signal L_w —width of sliding window d_w —moving step length of sliding window**Outputs:** Mat_{window} —data window matrix1: **Calculate** cl 2: **Initialize** Mat_{window} 3: **for** $i = 1$ to cl **do**4: **if** $i = 1$ 5: Assign the data from 1 to L_w in Sig_{org} to the i th column of the Mat_{window} .6: **else if** $i! = cl$ 7: Assign the data who are located from the $(i * d_w + 1)$ th to the $(i * d_w + 1 + L_w)$ th in Sig_{org} to the i th column of Mat_{window} .8: **else**9: Assign the data who are located from the $(i * d_w + 1)$ th to the end of Sig_{org} to the i th column of Mat_{window} and replace the Null in the i th column with 0.11: **End if**12: **End for**

Table 2. Tool working conditions.

Condition	Cutter	Spindle Speed	Feed Rate	Depth of Cut
Condition 1	C1_1 C1_2	2750 rpm	220 mm/min	1.75 mm
Condition 2	C2_1	3000 rpm	200 mm/min	1.75 mm
Condition 3	C3_1	3000 rpm	240 mm/min	1.75 mm
Condition 4	C4_1 ... C4_6	3000 rpm	250 mm/min	1.75 mm
Condition 5	C5_1	3250 rpm	275 mm/min	1.75 mm
Condition 6	C6_1	3500 rpm	250 mm/min	1.75 mm
Condition 7	C7_1 ... C7_9	3500 rpm	300 mm/min	1.75 mm
Condition 8	C8_1 ... C8_7	4500 rpm	400 mm/min	1.5 mm

Figure 8 shows the used sensor signal of cutter C4_3, including the vibration signal, current signal, sound signal, and force signal. It can be found that there is no clear trend in the data. In that case, we employ our model to extract more information.

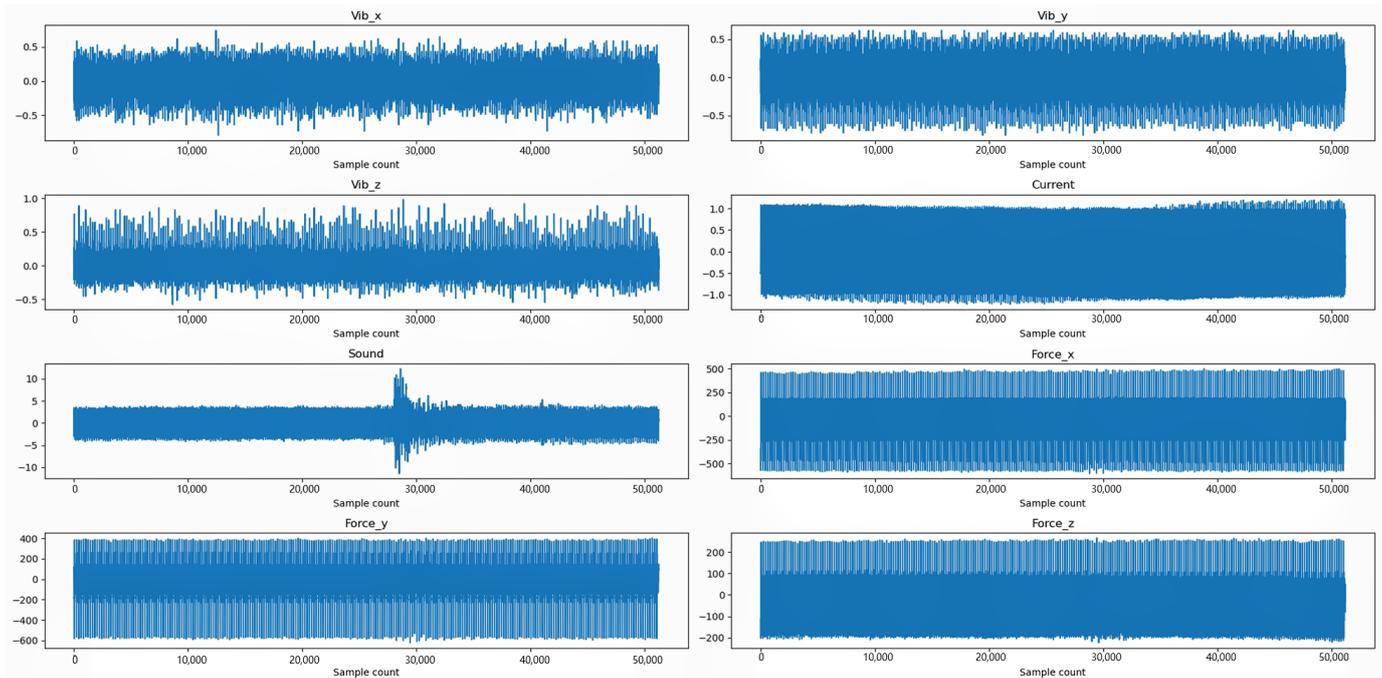


Figure 8. Monitoring signal data of C4_3.

The wear data of the tool indicate that the process of tool wear can be divided into three stages: the initial wear stage, the normal wear stage, and the rapid wear stage. First, there is the initial wear stage. Due to the regrinding of the tool, the cutting edge and tool surface are not smooth enough, resulting in a small actual contact area between the back surface of the tool and the cutting surface, but with high pressure. Therefore, the wear is rapid but for a short period of time. Next is the normal wear stage. After the initial wear, the contact area between the back surface of the tool and the workpiece increases, and the pressure per unit area decreases gradually. The micro-rough surface of the back surface of the tool is smoothed out, resulting in a slower wear rate. This stage represents the tool's effective working phase. Finally, there is the rapid wear stage. When the amount of tool wear reaches a certain limit, the cutting force and cutting temperature increase dramatically, leading to an accelerated tool wear rate until the tool loses its cutting ability. This stage is referred to as the rapid wear stage. The tool must be replaced before entering the rapid wear stage. As shown in Figure 9, three tools, C1_1, C2_1, and C4_3, demonstrate the three stages of wear. It can be observed that initially, the tool wear rapidly increases within a short period, then the growth rate slows down until the rapid wear stage, where the wear value starts to increase rapidly again.

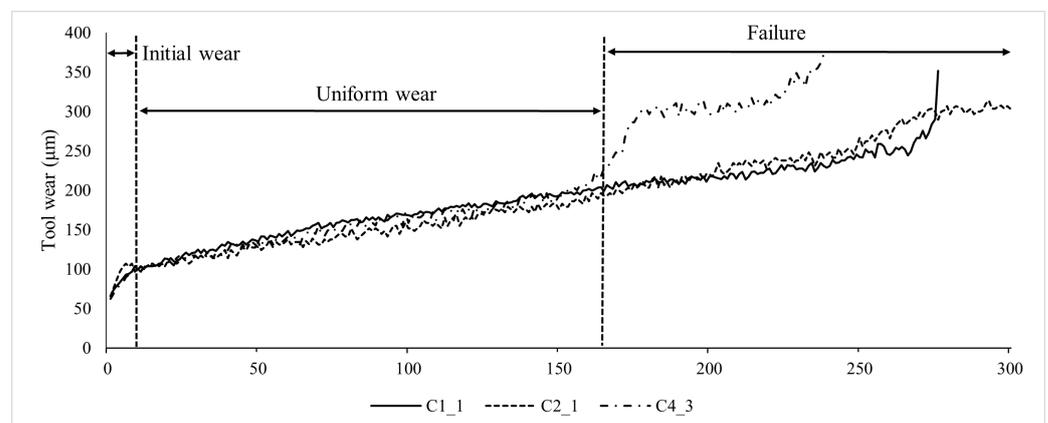


Figure 9. Wear stage of C1_1, C2_1, C4_3.

After data cleaning, we utilized the PCHIP method to conduct data augmentation on the missing portions of tool wear in the dataset, aiming to expand the application of information within the dataset. The interpolation results are depicted in Figures 10 and 11. Figure 10 showcases the interpolated outcome for C6_1, and Figure 11 displays the interpolated outcome for C8_1. The PCHIP interpolation method significantly resolves the problem of missing wear data, enabling the utilization of rich sensor feature information associated with the previously absent wear values.

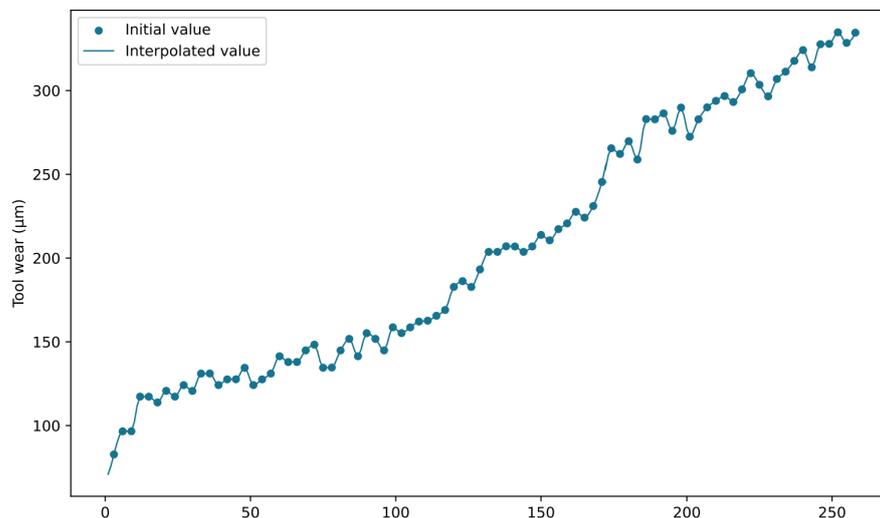


Figure 10. Interpolation results of C6_1.

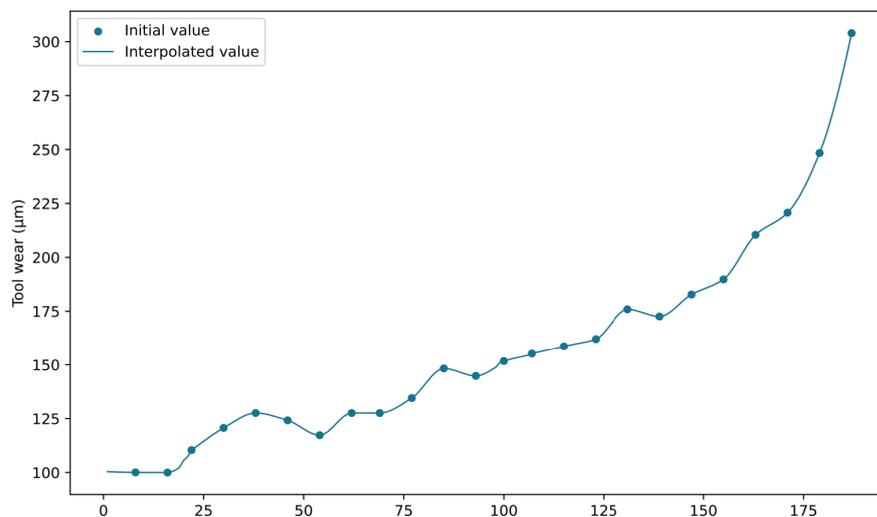


Figure 11. Interpolation results of C8_1.

We evaluate the performance of interpolation methods by choosing the data which do not have missing wear values and part of the true wear values, comparing the true wear values with the interpolation values. We compare PCHIP with three other common approaches—cubic spline, spline and linear by PCC, MAE, RMSE, MAPE, and standard deviation. The result is shown in Table 3. It shows that the result of PCHIP is the best among all methods below. According to the results below, it can be seen that PCHIP interpolation is better than other interpolation methods in PCC, MAE, RMSE, and standard deviation. PCHIP has the best interpolation performance. Therefore, we choose the PCHIP interpolation method to perform wear data interpolation work.

Table 3. Results of interpolation methods.

Methods	PCC	MAE	RMSE	MAPE	Standard Deviation
pchip [29]	0.9948	3.1701	4.7902	0.0191	4.7682
cubic spline	0.9942	3.2968	5.0278	0.0189	5.0265
spline	0.9932	3.3427	5.4407	0.0206	5.4278
linear	0.9934	3.2514	5.4306	0.0201	5.3822

Bold indicates optimal performance.

4.3. Prediction Results and Comparison

To demonstrate the effectiveness of the proposed methods, we compare it with the other three methods on the same test dataset. We represent the proposed method as PECCG. And the other three models are denoted as CNN, CNN Blocks, and GRU. The CNN method only uses a one-dimensional CNN. And the CNN Blocks method contains a configurable number of convolutional blocks. The GRU method only includes a GRU model. The tool wear prediction result of the cutters C4_3 and C7_9 using the four different models is shown in Figures 12 and 13. It can be seen that our combination of CNN and GRU is superior to the model which only uses CNN or RNN. It shows that PECCG can effectively extract features from high-dimensional data and as we can see, it can more accurately capture the underlying trend in the data. The CNN Blocks model captures the trend at first, but when the wear value suddenly changes, it fails to complete the prediction. This result also demonstrates that our model produces more robust and less volatile predictions compared to the other models.

To further quantify the effectiveness of our proposed model, we introduced five key evaluation metrics, including Pearson Correlation Coefficient (PCC), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Standard Deviation, and Relative Accuracy. These metrics were calculated using the same test set to assess the performance of our model. The equations for these metrics are as follows:

- **Pearson Correlation Coefficient (PCC)**
PCC measures the linear correlation between predicted and actual values, ranging from -1 to 1 .

$$\text{PCC} = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{\left[n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right] \left[n \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i \right)^2 \right]}} \quad (10)$$

- **Mean Absolute Error (MAE)**
MAE measures the average absolute difference between predicted and actual values.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (11)$$

- **Root Mean Squared Error (RMSE)**
RMSE measures the square root of the average squared difference between predicted and actual values.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (12)$$

- **Standard Deviation**
The standard deviation of errors is an indicator of the robustness of a model. A lower standard deviation signifies a higher degree of stability of the prediction performance.

$$\text{SD} = \sqrt{\frac{\sum_{i=1}^n (e_i - \bar{e})^2}{n}} \quad (13)$$

- Relative Accuracy

Relative accuracy is a measure of the error or difference between a measured or calculated value and the true value of a quantity, ranging from 0 to 1.

$$RA = 1 - \frac{1}{n} \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{\sum_{i=1}^n y_i} \quad (14)$$

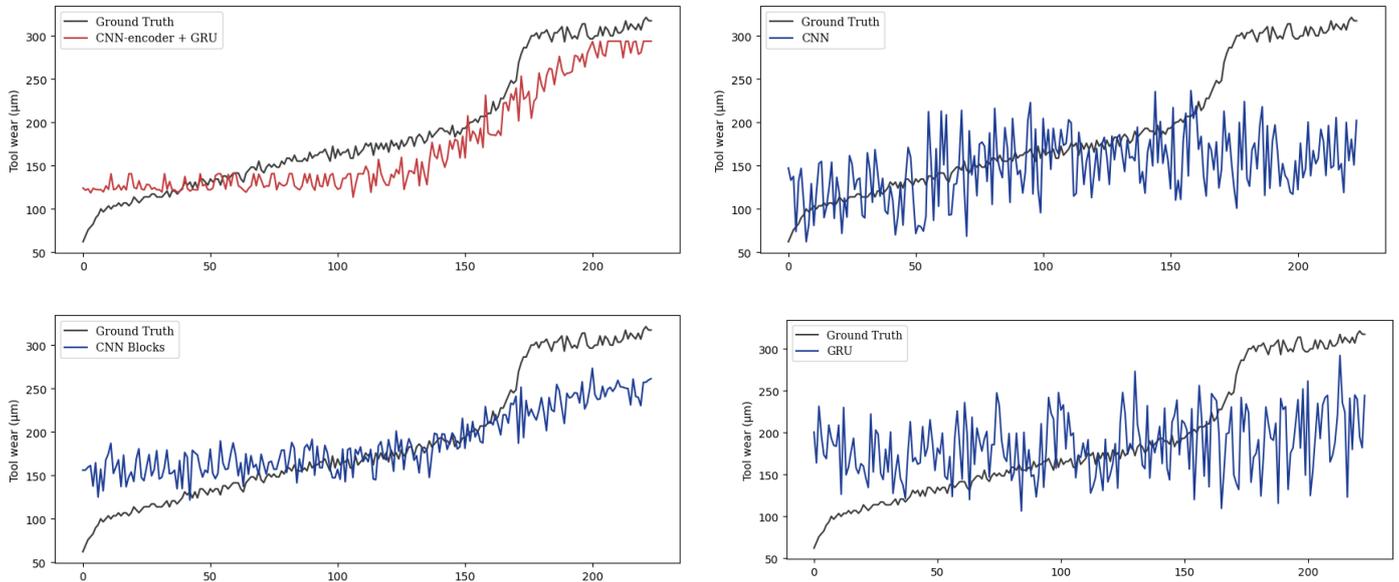


Figure 12. Tool wear prediction results of C4_3.

The metrics of the four models are illustrated in Table 4 and Figure 14. Among these models, PECC performs the best in all metrics. GRU performs the worst due to poor feature extraction. The PCC of GRU is quite low, at only 0.1947, and the relative accuracy is 0.6383, which is terrible, too. The results show that the single GRU is not suitable for performing regression. It can be seen that the CNN Blocks method performs better than CNN. Its PCC is 16.4% higher than CNN. Nevertheless, its performance can be improved. When we combine GRU with CNN Blocks, PECC outperforms those of all other models tested, providing strong evidence for its superior performance. The PCC of PECC is 0.9538, which highlights the strong correlation between predicted and actual wear values. Its standard deviation is about half of CNN. As a result of the integration of CNN Blocks and GRU, the relative accuracy of PECC is 0.8522, which is superior to the other three models. The design of PECC is less complex than the time–space attention model [24], while delivering superior performance outcomes. The relative accuracy of the time–space attention model is 0.7890. In comparison, PECC exhibits a relative accuracy that is 8% higher.

Table 4. Tool wear performance estimation results of four networks.

Methods	PCC	Relative Accuracy	MAE	RMSE	Standard Deviation
CNN [36]	0.7957	0.7898	40.2158	56.3456	44.3952
CNN Blocks	0.9258	0.8097	34.0152	41.8622	28.8696
GRU	0.1947	0.6383	56.6794	70.6391	70.6137
PECC	0.9538	0.8522	23.8362	28.5240	22.2840

Bold indicates optimal performance.

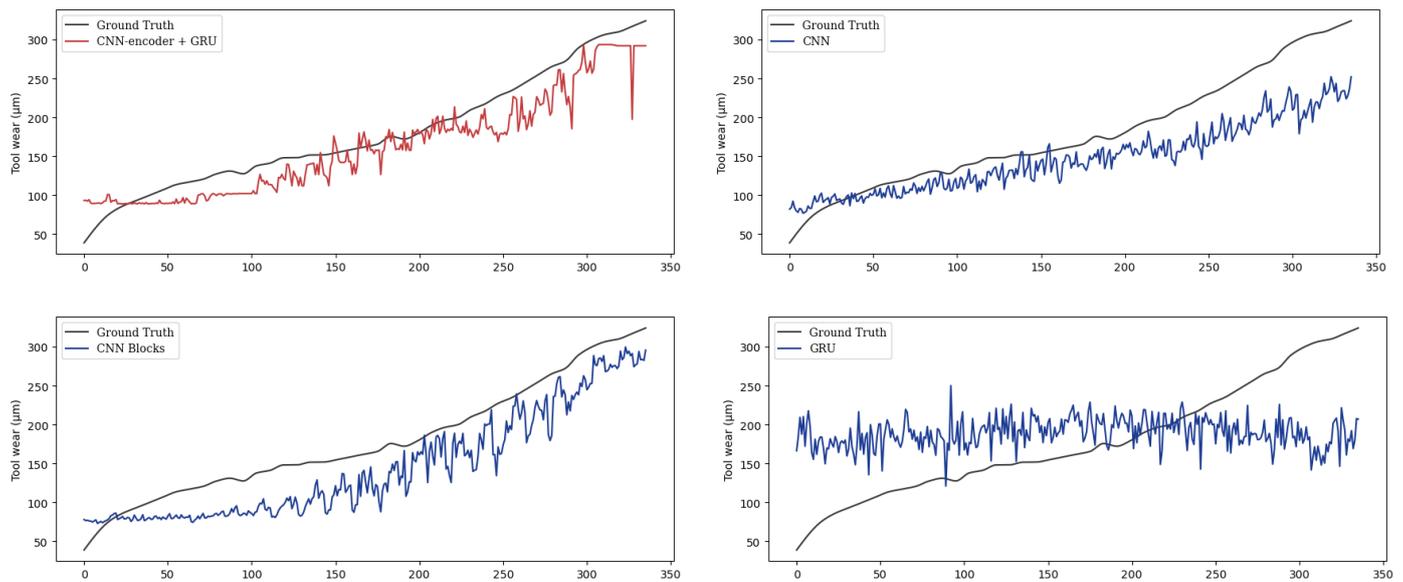


Figure 13. Tool wear prediction results of C7_9.



Figure 14. Tool wear performance estimation results of four networks.

In order to further illustrate the effectiveness of the PCHIP method, we take prediction results of four models trained without interpolation processing of missing data on the test set as the baseline. By comparing the predictive outcomes of interpolated and non-interpolated models, it can be inferred that the four models trained on interpolated data exhibit superior performance across all metrics when compared to the models trained on non-interpolated data. Results are shown in Figure 15. It can be seen from the dark blue bars that PEGC outperforms other models even when we do not use the PCHIP interpolation method. This demonstrates the superiority of our model architecture. When combined with the PCHIP method, all major metrics of the four models have been improved, further illustrating the effectiveness of the interpolation method we have adopted. The light blue bars in Figure 15 show that by incorporating the PCHIP interpolation method, noteworthy improvements are observed among the evaluated models. Specifically, the standard deviation of CNN Blocks decreases from 53.5264 to 28.8696, representing a significant reduction of approximately 46%. Similarly, the RMSE of PEGC decreased from 41.0460 to 28.5240, indicating a substantial decline of approximately 31%. These findings underscore the efficacy of the employed interpolation approach.

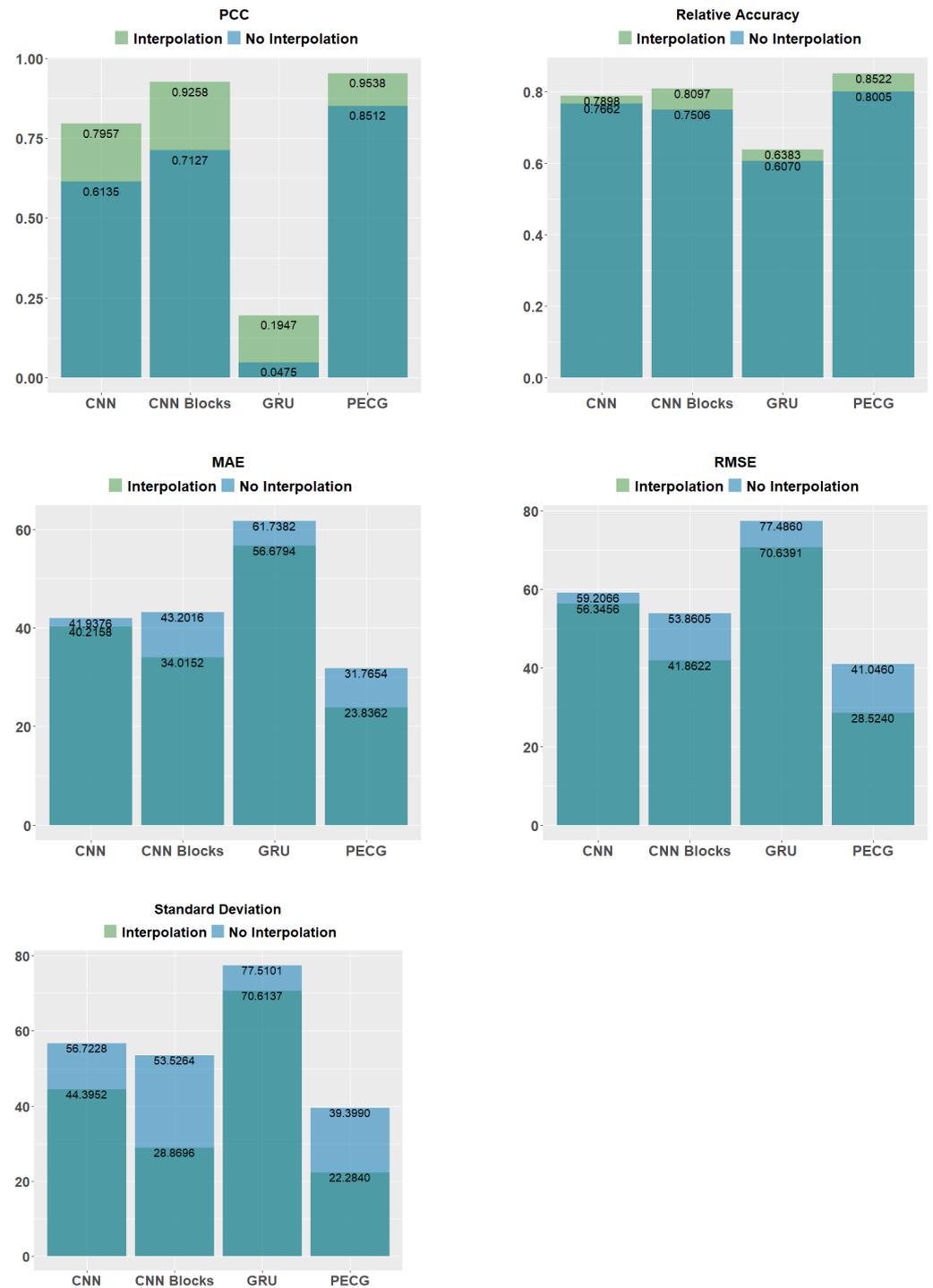


Figure 15. Comparison results with and without interpolation.

4.4. Phm 2010 Dataset Results

We validate the performance of the proposed method on the PHM 2010 dataset [37]. The platform of the PHM 2010 competition is shown in Figure 16. The cutting conditions of the dataset remain unchanged, utilizing a 6 mm ball nose tungsten carbide cutter to perform straight tool path cuts on the sidewall of an aluminum alloy blank. The experimental parameters are shown in Table 5. C1 and C4 are used as the training set, while C6 is used as the testing set. The tool scrap standard is 170 μm [18]. The results of different methods are shown below. As is shown in Figure 11, the proposed ConvGRU performs the best when the tool wear value is less than 170 μm . It seems that the other three models predict

more accurately than ConvGRU after the wear value exceeds 170 μm . However, in practice, these good performances have no practical significance, because when the tool wear value reaches 170 μm , it is considered to have reached the scrapping criteria and is no longer used. Therefore, the prediction results before reaching the tool scrap criteria are more important. As shown in Figure 17, the proposed ConvGRU model outperforms the other three models significantly in this aspect.

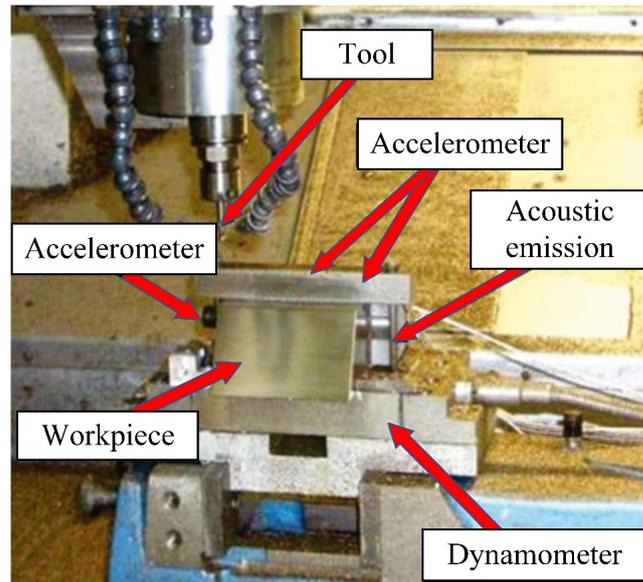


Figure 16. PHM 2010 competition platform [38].

Table 5. PHM 2010 competition experimental parameters.

Classification	Model/Value	Classification	Value
Machine model	Roders Tech RFM 760	Radial cutting depth	0.125 mm
Workpiece material	Nickel-based superalloy 718	Axial cutting depth	0.2 mm
Tool	3-tooth ball nose milling cutter	Number of sensors	3
Spindle speed	10,400 RPM	Number of sensing channels	7
Feed rate	1555 mm/min	Sampling frequency	50 kHz

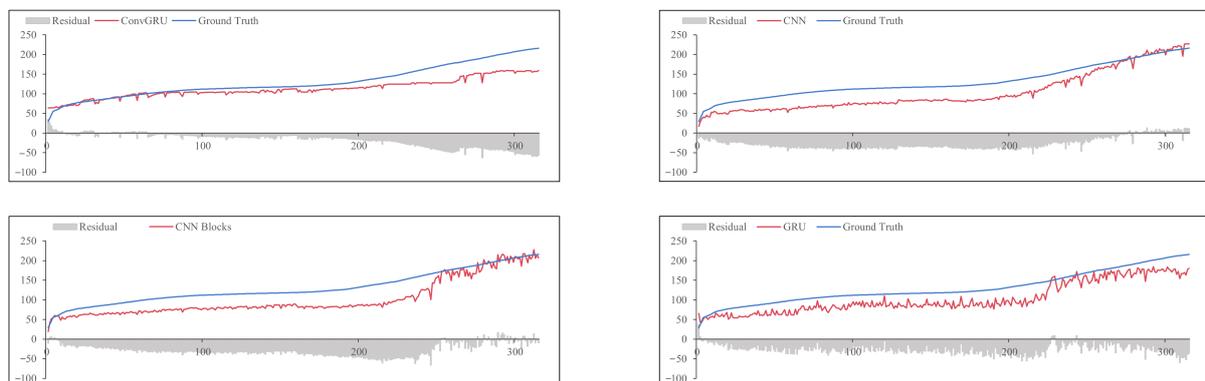


Figure 17. Comparison of results on PHM 2010 dataset.

We also prove the effectiveness of data augmentation on the PHM 2010 dataset. Firstly, we randomly selected 20% of the data from the dataset and removed these data points to simulate the scenario where tool wear values are missing in practice. Then, we applied PCHIP interpolation to fill in the missing data, and trained the ConvGRU model on the interpolated dataset. The results are shown in Table 6 and Figure 18. As shown in Table

6, after adopting the data augmentation method of PCHIP interpolation, the prediction performance of the PECC model outperforms the simple ConvGRU model, which is trained without interpolation, in terms of four evaluation metrics: PCC, Relative Accuracy, MAE, and RMSE. Furthermore, the prediction results of PECC are very close to the original dataset, indicating that the proposed PECC model with the data augmentation method PCHIP yields favorable results and exhibits a small gap compared to the results of the model trained by the actual dataset. Therefore, PECC effectively addresses the issue of missing wear data in practical applications while also reducing the cost of multiple wear value measurements.

Table 6. Tool wear prediction performance of PECC on PHM 2010 dataset.

Data	PCC	Relative Accuracy	MAE	RMSE
Original data	0.9793	0.8724	18.7152	24.8331
Missing data	0.9592	0.8344	22.4959	27.0046
Interpolated data	0.9690	0.8622	19.6308	25.4882

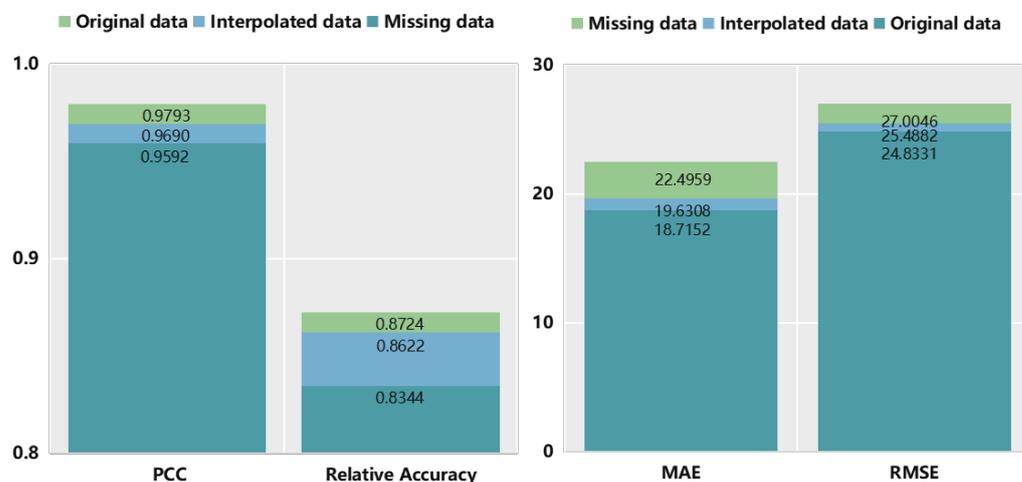


Figure 18. Tool wear prediction performance of PECC on PHM 2010 dataset.

5. Conclusions

In this paper, we introduce an efficient interpolation method known as PCHIP to address the challenge of missing data, specifically in the context of 397 wear prediction. Additionally, we present a novel model named PECC designed for wear prediction tasks.

CNNs possess a remarkable ability to learn hierarchical features from high-dimensional data, rendering them highly effective in capturing informative features for regression tasks. On the other hand, GRUs are known for their efficiency with fewer parameters and faster training speeds. GRUs, as a type of RNN, excel at capturing and modeling long-term dependencies within sequential data, making them particularly suited for time-series regression tasks. Moreover, the inclusion of gating mechanisms in GRUs means they need to learn a limited number of parameters, leading to accelerated training and improved generalization performance when compared to traditional RNNs.

In our approach, we unify CNN and GRU to craft the innovative PECC model. Initially, CNN plays a pivotal role in reducing the input data’s dimensionality and complexity, thus enabling precision in modeling temporal dependencies by the GRU. This fusion capitalizes on the exceptional feature extraction capabilities of CNN and the adeptness of GRU in handling time-series data. Consequently, PECC emerges as an effective predictive model for tool wear prediction, harnessing the strengths of both CNN and GRU.

Author Contributions: Conceptualization, J.H. and Y.W.; Methodology, J.H., L.Y. and H.L.; Validation, K.W.; Writing—original draft, J.H. and L.Y.; Writing—review & editing, Y.W. and H.G.; Supervision, Y.W. and H.G.; Funding acquisition, H.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China No. 51775452.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: PHM 2010 is available at https://phmsociety.org/phm_competition/2010-phm-society-conference-data-challenge.

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

Abbreviations

The following abbreviations are used in this manuscript:

SAE	Stacked Autoencoder
RUL	Remaining Useful Life
GRU	Gated Recurrent Unit
HMM	Hidden Markov Model
CNN	Convolutional Neural Network
CNN Block	Convolutional Neural Network Block
PCHIP	Piecewise Cubic Hermite Interpolating Polynomial
PECG	PCHIP-Enhanced ConvGRU
PCC	Pearson's Correlation Coefficient
CNC machine	Computerized Numerical Control Machine
MGRU	Multi-head gated recurrent unit
PCA	Principal component analysis
ANN	Artificial Neural Network
SVM	Support Vector Machine
GPR	Gaussian Process Regression
RNN	Recurrent Neural Network
DNN	Deep Neural Network
ReLU	Rectified Linear Unit
EMA	Exponential Moving Average
MAE	Mean Absolute Error
RMSE	Root Mean Square Error
MAPE	Mean Absolute Percentage Error

References

1. Bhattacharyya, P.; Sengupta, D.; Mukhopadhyay, S. Cutting force-based real-time estimation of tool wear in face milling using a combination of signal processing techniques. *Mech. Syst. Signal Process.* **2007**, *21*, 2665–2683. [[CrossRef](#)]
2. Song, W.; Liu, H.; Zio, E. Long-range dependence and heavy tail characteristics for remaining useful life prediction in rolling bearing degradation. *Appl. Math. Model.* **2022**, *102*, 268–284. [[CrossRef](#)]
3. Li, Q.; Ma, Z.; Li, H.; Liu, X.; Guan, X.; Tian, P. Remaining useful life prediction of mechanical system based on performance evaluation and geometric fractional Lévy stable motion with adaptive nonlinear drift. *Mech. Syst. Signal Process.* **2023**, *184*, 109679. [[CrossRef](#)]
4. Oguamanam, D.; Raafat, H.; Taboun, S. A machine vision system for wear monitoring and breakage detection of single-point cutting tools. *Comput. Ind. Eng.* **1994**, *26*, 575–598. [[CrossRef](#)]
5. Pedersen, K.B. Wear measurement of cutting tools by computer vision. *Int. J. Mach. Tools Manuf.* **1990**, *30*, 131–139. [[CrossRef](#)]
6. Cook, N.H. Tool wear sensors. *Wear* **1980**, *62*, 49–57. [[CrossRef](#)]
7. Beruvides, G.; Quiza, R.; Del Toro, R.; Haber, R.E. Sensing systems and signal analysis to monitor tool wear in microdrilling operations on a sintered tungsten-copper composite material. *Sens. Actuators A Phys.* **2013**, *199*, 165–175. [[CrossRef](#)]

8. Aliustaoglu, C.; Ertunc, H.M.; Ocak, H. Tool wear condition monitoring using a sensor fusion model based on fuzzy inference system. *Mech. Syst. Signal Process.* **2009**, *23*, 539–546. [[CrossRef](#)]
9. Arnaiz-González, Á.; Fernández-Valdivielso, A.; Bustillo, A.; López de Lacalle, L.N. Using artificial neural networks for the prediction of dimensional error on inclined surfaces manufactured by ball-end milling. *Int. J. Adv. Manuf. Technol.* **2016**, *83*, 847–859. [[CrossRef](#)]
10. Wiciak-Pikuła, M.; Felusiak-Czyryca, A.; Twardowski, P. Tool wear prediction based on artificial neural network during aluminum matrix composite milling. *Sensors* **2020**, *20*, 5798. [[CrossRef](#)] [[PubMed](#)]
11. Lu, M.C.; Wan, B.S. Study of high-frequency sound signals for tool wear monitoring in micromilling. *Int. J. Adv. Manuf. Technol.* **2013**, *66*, 1785–1792. [[CrossRef](#)]
12. Wang, J.; Xiang, Z.; Cheng, X.; Zhou, J.; Li, W. Tool Wear State Identification Based on SVM Optimized by the Improved Northern Goshawk Optimization. *Sensors* **2023**, *23*, 8591. [[CrossRef](#)] [[PubMed](#)]
13. Przybyś-Małaszczak, A.; Antoniuk, I.; Szymanowski, K.; Kruk, M.; Kurek, J. Application of Machine Learning Algorithms for Tool Condition Monitoring in Milling Chipboard Process. *Sensors* **2023**, *23*, 5850. [[CrossRef](#)] [[PubMed](#)]
14. Geramifard, O.; Xu, J.X.; Zhou, J.H.; Li, X. A physically segmented hidden markov model approach for continuous tool condition monitoring: Diagnostics and prognostics. *IEEE Trans. Ind. Inform.* **2012**, *8*, 964–973. [[CrossRef](#)]
15. Liao, Z.; Gao, D.; Lu, Y.; Lv, Z. Multi-scale hybrid HMM for tool wear condition monitoring. *Int. J. Adv. Manuf. Technol.* **2016**, *84*, 2437–2448. [[CrossRef](#)]
16. Li, W.; Liu, T. Time varying and condition adaptive hidden Markov model for tool wear state estimation and remaining useful life prediction in micro-milling. *Mech. Syst. Signal Process.* **2019**, *131*, 689–702. [[CrossRef](#)]
17. Yu, J.; Liang, S.; Tang, D.; Liu, H. A weighted hidden Markov model approach for continuous-state tool wear monitoring and tool life prediction. *Int. J. Adv. Manuf. Technol.* **2017**, *91*, 201–211. [[CrossRef](#)]
18. Li, D.; Li, Y.; Liu, C. Gaussian process regression model incorporated with tool wear mechanism. *Chin. J. Aeronaut.* **2022**, *35*, 393–400. [[CrossRef](#)]
19. Kong, D.; Chen, Y.; Li, N. Gaussian process regression for tool wear prediction. *Mech. Syst. Signal Process.* **2018**, *104*, 556–574. [[CrossRef](#)]
20. Serin, G.; Sener, B.; Ozbayoglu, A.M.; Unver, H.O. Review of tool condition monitoring in machining and opportunities for deep learning. *Int. J. Adv. Manuf. Technol.* **2020**, *109*, 953–974. [[CrossRef](#)]
21. Liu, M.; Yao, X.; Zhang, J.; Chen, W.; Jing, X.; Wang, K. Multi-sensor data fusion for remaining useful life prediction of machining tools by IABC-BPNN in dry milling operations. *Sensors* **2020**, *20*, 4657. [[CrossRef](#)]
22. Wu, X.; Liu, Y.; Zhou, X.; Mou, A. Automatic identification of tool wear based on convolutional neural network in face milling process. *Sensors* **2019**, *19*, 3817. [[CrossRef](#)] [[PubMed](#)]
23. Cai, W.; Zhang, W.; Hu, X.; Liu, Y. A hybrid information model based on long short-term memory network for tool condition monitoring. *J. Intell. Manuf.* **2020**, *31*, 1497–1510. [[CrossRef](#)]
24. Feng, T.; Guo, L.; Gao, H.; Chen, T.; Yu, Y.; Li, C. A new time–space attention mechanism driven multi-feature fusion method for tool wear monitoring. *Int. J. Adv. Manuf. Technol.* **2022**, *120*, 5633–5648. [[CrossRef](#)]
25. Zhong, Z.; Zheng, L.; Kang, G.; Li, S.; Yang, Y. Random erasing data augmentation. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34, pp. 13001–13008.
26. Shukla, S.N.; Marlin, B.M. Interpolation-prediction networks for irregularly sampled time series. *arXiv* **2019**, arXiv:1909.07782.
27. Long, Y.; Zhou, W.; Luo, Y. A fault diagnosis method based on one-dimensional data enhancement and convolutional neural network. *Measurement* **2021**, *180*, 109532. [[CrossRef](#)]
28. Wang, F.; Liu, X.; Liu, C.; Li, H.; Han, Q. Remaining useful life prediction method of rolling bearings based on Pchip-EEMD-GM (1, 1) model. *Shock Vib.* **2018**, *2018*, 3013684. [[CrossRef](#)]
29. Fritsch, F.N.; Carlson, R.E. Monotone piecewise cubic interpolation. *SIAM J. Numer. Anal.* **1980**, *17*, 238–246. [[CrossRef](#)]
30. Li, Z.; Liu, R.; Wu, D. Data-driven smart manufacturing: Tool wear monitoring with audio signals and machine learning. *J. Manuf. Process.* **2019**, *48*, 66–76. [[CrossRef](#)]
31. He, J.; Gao, H.; Li, S.; Guo, L.; Lei, Y.; Cao, A. An intelligent maintenance decision-making based on cutters economic life. *Int. J. Prod. Econ.* **2024**, *267*, 109075. [[CrossRef](#)]
32. You, Z.; Gao, H.; Guo, L.; Liu, Y.; Li, J.; Li, C. Machine vision based adaptive online condition monitoring for milling cutter under spindle rotation. *Mech. Syst. Signal Process.* **2022**, *171*, 108904. [[CrossRef](#)]
33. Sun, Y.; Li, S.; Gao, H.; Zhang, X.; Lv, J.; Liu, W.; Wu, Y. Transfer learning: A new aerodynamic force identification network based on adaptive emd and soft thresholding in hypersonic wind tunnel. *Chin. J. Aeronaut.* **2023**, *36*, 351–365. [[CrossRef](#)]
34. Sun, Y.; He, J.; Gao, H.; Song, H.; Guo, L. A New Semi-supervised Tool-wear Monitoring Method using Unreliable Pseudo-Labels. *Measurement* **2023**, *226*, 113991. [[CrossRef](#)]
35. Li, C.; Guo, L.; Gao, H.; Li, Y. Similarity-measured isolation forest: Anomaly detection method for machine monitoring data. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–12. [[CrossRef](#)]
36. Dai, W.; Liang, K.; Wang, B. State monitoring method for tool wear in aerospace manufacturing processes based on a convolutional neural network (CNN). *Aerospace* **2021**, *8*, 335. [[CrossRef](#)]

-
37. 2010 PHM Society Conference Data Challenge. Available online: https://phmsociety.org/phm_competition/2010-phm-society-conference-data-challenge (accessed on 1 April 2023).
 38. Benkedjouh, T.; Zerhouni, N.; Rechak, S. Tool wear condition monitoring based on continuous wavelet transform and blind source separation. *Int. J. Adv. Manuf. Technol.* **2018**, *97*, 3311–3323. [[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.