Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Prediction of surface roughness in aluminum alloy milling based on dynamic and static data fusion

Boyang Meng

bymeng@hrbust.edu.cn

Harbin University of Science and Technology https://orcid.org/0000-0003-1418-7374

Song Li
Zhongjie Li
Fuxing Xiang
Rui Ding
Caixu Yue
Xianli Liu

Research Article

Keywords: Dynamic and static data, Surface roughness, Linear modulation, Bidirectional cross-attention fusion

Posted Date: April 2nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9018628/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Version of Record: A version of this preprint was published at The International Journal of Advanced Manufacturing Technology on July 25th, 2026. See the published version at https://doi.org/10.1007/s00170-026-18800-4.

Prediction of surface roughness in aluminum alloy milling based on dynamic and static data fusion

Boyang Meng[1]*,Song Li[1],Zhongjie Li[1],Fuxing Xiang[1],Rui Ding[1], Caixu Yue[1],Xianli Liu[1]

[1]Key Laboratory of Advanced Manufacturing and Intelligent Technology, Ministry of Education, Harbin University of Science and Technology, Harbin 150080, China

*Corresponding author: Boyang Meng. E-mail: bymeng@hrbust.edu.cn.

Abstract

In advanced manufacturing, surface roughness is a critical metric for evaluating the quality of machined workpieces, which directly impacts product quality and manufacturing efficiency. However, traditional roughness measurement methods, such as contact profilometers and optical inspection technologies, require precise calibration, are susceptible to environmental influences, and have limited data acquisition capabilities. These methods result in high costs, low efficiency, and an inability to meet the demands of real-time monitoring in actual production. To address these challenges, this study proposes a multi-source signal fusion model architecture that integrates parameter conditions, bidirectional cross-attention, and global encoding mechanisms. First, the collected sensor signals undergo preprocessing, feature extraction, and principal component analysis dimensionality reduction. The reduced-dimensional current and cutting force features serve as dynamic inputs, and process parameters serve as static conditions. This approach provides the model with high-quality data, thereby enhancing prediction accuracy. Second, the reduced-dimensional current and cutting force features, along with the process parameters, are uniformly encoded into token sequences. The FiLM linear modulation mechanism is then applied to adaptively scale and translate these features, improving prediction accuracy under varying parameter conditions. Finally, bidirectional cross-attention learns the relationship between the current and the cutting force. This relationship is then fused with the process parameters for encoding. A shallow neural network is used for regression prediction. The results validate the model's effectiveness in predicting surface roughness, offering a novel solution for this application.

Keywords Dynamic and static data; Surface roughness; Linear modulation; Bidirectional cross-attention fusion

1 Introduction

As a key indicator of machined surface quality, surface roughness significantly impacts the wear and corrosion resistance, as well as the mechanical properties, of processed products[1-2]. Excessive roughness accelerates wear in mechanical products, leads to poor assembly quality, and causes fatigue failure. These issues adversely affect the performance and service life of the entire machine[3-6]. In demanding and complex application environments, such as those in the aerospace

and precision manufacturing industries, surface quality issues in mechanical components can lead to significant losses. Therefore, predicting product surface roughness is crucial for improving product quality, minimizing losses, and extending service life.

In recent years, numerous scholars have conducted extensive research on predicting surface roughness. These methods currently fall into three categories: traditional mechanical modeling, conventional data mining, and deep learning techniques[7]. Traditional mechanical modeling methods establish mathematical models of surface roughness based on the physical mechanisms involved in the machining process, thereby enabling prediction[8]. These methods have the advantages of high interpretability and relatively low data requirements. However, these methods have high computational demands and are inefficient. They also introduce various idealized assumptions during model development, which can result in inaccurate final predictions[9-10]. Additionally, simplifications in theoretical surface roughness models often result in discrepancies between predicted and expected values.

Traditional surface roughness prediction methods in data mining use existing machining experiments or data to determine the relationship between process parameters and surface roughness through feature extraction and classical machine learning algorithms. Hakan et al.[11]developed a multilayer perceptron (MLP), learning vector quantization (LVQ), and fuzzy Mamdani approach to evaluate the surface roughness of SLM-Ti6Al4V components. Zhou et al.[12]proposed a feed-forward, multilayer artificial neural network (ANN) roughness prediction model for Al/SiC particle composites that incorporates cutting parameters and the volume fraction of the composites as inputs. Based on this model, they developed a method to accurately determine specific target roughness values for precise cutting parameter settings. However, as communication and sensor technologies become more integrated, an increasing number of researchers are incorporating sensor signals into model inputs. By analyzing time-, frequency-, and time-frequency domain data during machining, they extract latent dynamic machining information to further enhance prediction accuracy[13]. Chen et al.[14]proposed a prediction model for the surface roughness of complex-structured workpieces during turning based on vibration signals using gaussian process regression (GPR) technology. Zhao et al.[15]proposed an ensemble learning support vector machine (ELSVM)-based surface roughness prediction model for LUVAG alumina ceramic surfaces. Kristin et al.[16]combined a kinematic model with an ensemble enhanced regression tree machine learning algorithm. This combination uses machining parameters and sensor data as inputs to accurately predict surface roughness. These studies have advanced surface roughness prediction technology and provided new insights and methodologies for optimizing manufacturing processes. However, these approaches often require extensive sample data, and parameters are usually set through trial and error or empirically, resulting in limited flexibility and applicability. Additionally, incorrect parameter settings can lead to biased prediction results, compromising accuracy.

The rapid advancement of deep learning has provided a novel approach to tackling complex, high-dimensional, nonlinear problems in high-dimensional data mining and feature learning[17]. In industrial applications, deep learning has been widely used in fault diagnosis and remaining life prediction research, achieving significant progress[18-19]. Consequently, researchers have introduced deep learning into surface roughness prediction studies as well. Fang et al.[20]proposed a machine vision inspection method based on an improved model agnostic meta-learning (MAML+) network to address the issue of traditional deep learning methods requiring large amounts of training data. Guo et al.[21]introduced a novel surface roughness prediction model based on fused

signal features and a deep learning architecture. They employed variational modal decomposition to extract time-frequency features from force and vibration signals. genetic algorithm (GA)-based model selection (MI) and optimal roughness feature subsets were selected as model inputs. deep belief networks (DBNs) were used for roughness prediction, and the model parameters were optimized using the incremental stability seeking algorithm (ISSA). Wang et al.[22]proposed a surface roughness prediction method that accounts for variation in tool wear with cutting parameters. They used tool wear status and sensor signals as inputs and designed a stacked autoencoder and long short-term memory network (SAE-LSTM) as the fundamental surface roughness prediction model. Transfer learning strategies were employed to enable online surface roughness prediction under variable cutting parameters. In complex, multi-parameter, coupled machining processes, however, different input features contribute differently to surface roughness. Introducing an attention mechanism allows for the adaptive highlighting of key factors, thereby enhancing the model's interpretability and predictive performance. Xiao et al.[23]proposed the CAT-Net multi-hidden-layer deep learning prediction model. To improve the model's predictability, they incorporated an attention mechanism to assign weights to grinding parameters and analyzed its impact on roughness prediction. Wang et al.[24]proposed the mutual convergence fusion network (MCFN) for predicting surface roughness in robotic machining using internal and external signals. Based on a dual cross-modal attention mechanism (Dual-MCM), the MCFN enables the collaborative interaction of cross-modal information, segmented aggregation, and subsequent roughness prediction output. Zhang et al.[25]proposed the attention mechanism spatial-multiscale parallel network (AMS-Net), a deep learning model that incorporates channel and spatial attention mechanisms within a multiscale parallel framework for predicting surface roughness under dual-light conditions. This network uses two branches to extract surface roughness features from images captured under different light sources. Each branch constructs a multi-scale parallel convolution module (MSP) to extract feature information from roughness images at different scales. Dual-source spatial attention modules (DSA) and dual-source channel attention modules (DCA) perform spatial and channel projections, respectively, and the roughness feature maps output by the MSP interact with each other. Despite advances in deep learning methods for predicting surface roughness, practical, complex operating conditions often reveal that process parameters and signals from different sensors are not effectively integrated or sufficiently interacted within models. This limits predictive performance. This study addresses these limitations by simultaneously incorporating process parameters and multi-sensor signals (current and force) into the model. Leveraging attention mechanisms for interaction and holistic integration harnesses complementary information from diverse sources to enhance the stability and accuracy of surface roughness prediction under complex operating conditions.

The main innovations of this study are as follows:

（1）We propose a framework for process parameter and multi-sensor signal fusion aimed at predicting surface roughness. Integrating current signals, cutting force signals, and process parameters within the model effectively mitigates prediction instability and inaccuracy caused by the utilization of insufficient information.

（2）Using the FiLM mechanism to adjust signal characteristics through process parameters allows the model to more accurately reflect the impact of machining parameters on signals. This enables the model to better adapt to scenarios involving multiple parameter variations.

（3）A cross-signal deep interaction fusion module has been constructed. Bidirectional

cross-attention allows for the fusion of information in both directions between current and cutting force features. When combined with a Transformer encoder for global feature learning, the module's information fusion capabilities and prediction performance are enhanced.

The remainder of this paper is structured as follows: Section 2 introduces the fundamental theory and overall framework underlying the methods employed in this model. Section 3 details the model's specific modules. Section 4 covers the collection, preprocessing, and feature extraction of milling experimental data, as well as the tuning of the model's hyperparameters. Section 5 analyzes the results of training the model and validates the effectiveness of the proposed model through comparative experiments and ablation studies. Section 6 concludes the research.

2 Methodology

2.1 Data preparation

2.1.1 Data preprocessing

Preprocessing is required before feature extraction to enhance data quality. The experimentally acquired data includes entry and exit processes, resulting in raw signals that contain idle-state components irrelevant to this study and must be filtered out. For example, the raw cutting force signal from the seventh milling operation is shown in Fig.1. The signals in the X, Y, and Z directions each have segments at the beginning and end where the signal amplitude is significantly lower than the stable amplitude. This occurs because signal acquisition began before milling started and ended after milling finished, resulting in the collection of additional no-load signal segments that must be removed.

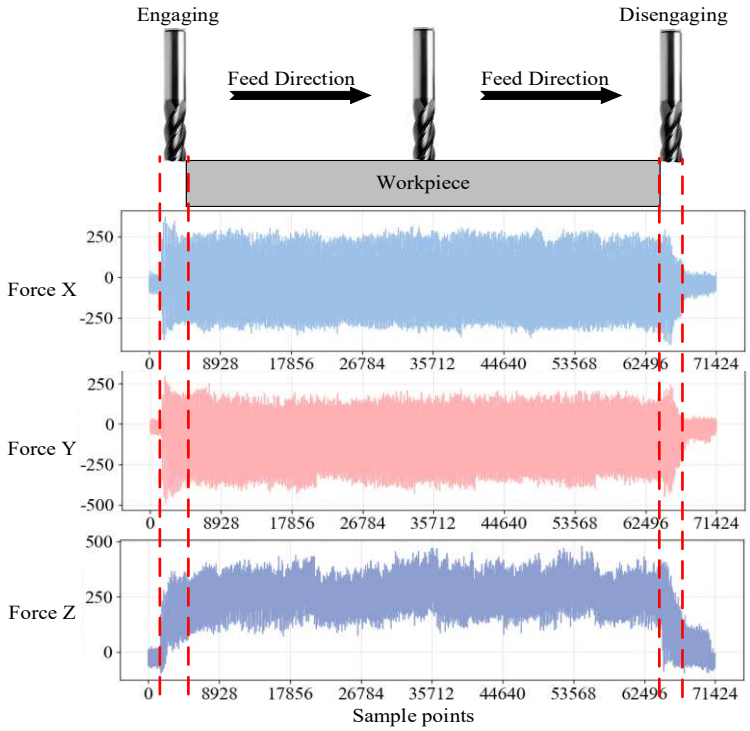


Fig.1 Cutting Force Signal with No-Load Component Removed

Fig.2 displays the spectral characteristics of the cutting force and current signals from the 10th, 30th, 50th, and 70th milling operations. The spectral energy of each signal is concentrated primarily in the low-to-mid frequency range, with low overall amplitude in the high-frequency region. This indicates that there is no significant noise interference. Thus, no additional denoising

algorithms are applied during preprocessing; only no-load signals are removed to prevent the loss of valuable information.

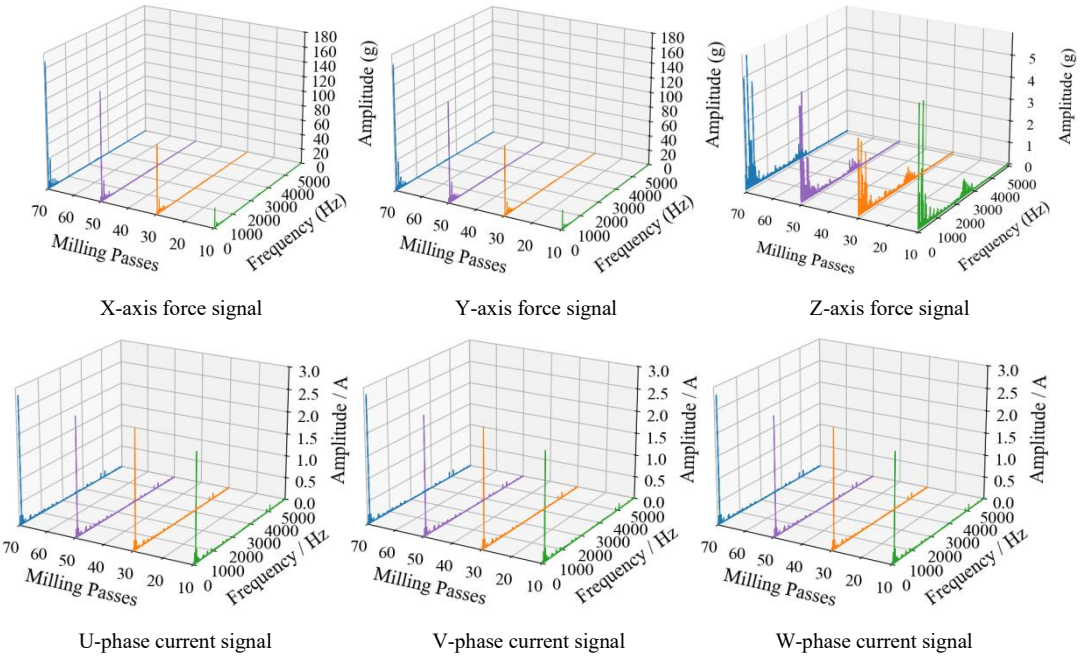


Fig.2 Spectra of Cutting Force Signal and Current Signal

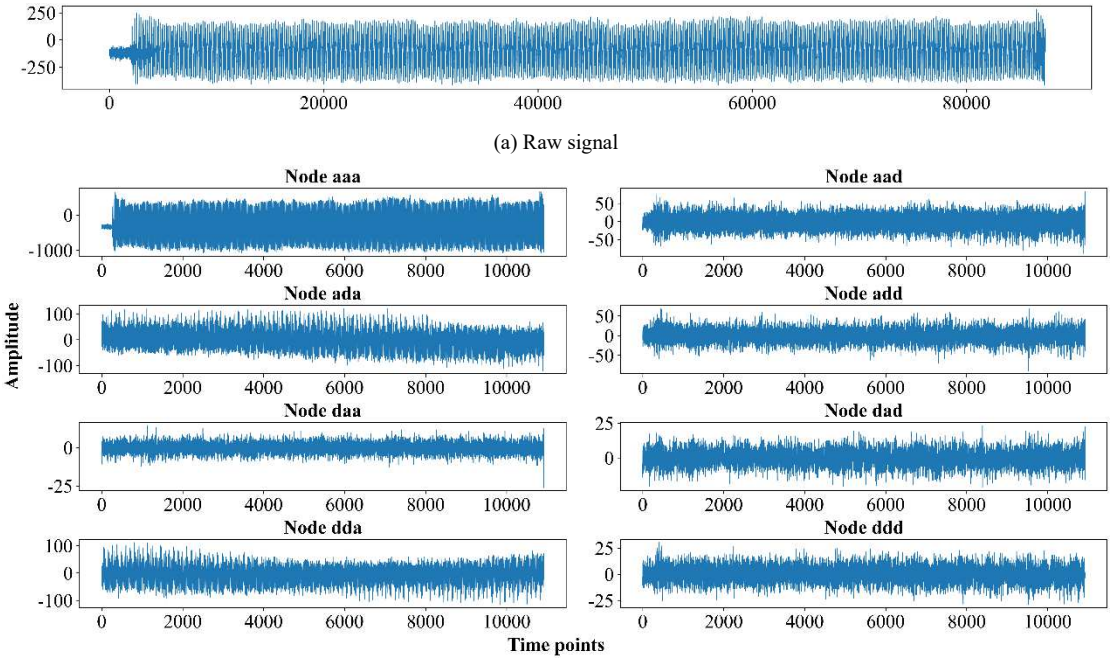
2.1.2 Feature extraction

The correlation between milling surface roughness and signal characteristics can be further analyzed based on the preprocessed signals. In the field of machining, surface roughness and signals during the cutting process exhibit correlations in the time domain, the frequency domain, and the time-frequency domain. Time-domain features directly characterize signal amplitude fluctuations and impact characteristics, which facilitates describing dynamic stability and trends during the machining process as operating conditions change. This study therefore extracts 13 time-domain features, including seven dimensioned indicators: mean, absolute mean amplitude, peak-to-peak amplitude, variance, standard deviation, root mean square (RMS), and root amplitude, as well as six dimensionless indicators: kurtosis, skewness, pulse factor, peak factor, waveform factor, and margin factor.

Time-domain signals directly reflect fluctuations and transient changes over time. However, they are limited in their ability to characterize periodic components and energy distribution. Therefore, this study uses frequency-domain features to provide a more comprehensive description of the signals. The frequency-domain features selected for this study are center frequency, root mean square frequency, and frequency variance.

Time-frequency domain feature extraction captures signal information in both the time and frequency dimensions simultaneously, providing a more comprehensive understanding of dynamic behavior during the milling process. This study uses Wavelet Packet Decomposition (WPD) to analyze time and frequency. The optimal wavelet basis function is selected based on the maximum ratio of energy to Shannon entropy. Computational evaluation revealed that coif5 is the optimal basis function; therefore, full-band multiscale partitioning was performed under the coif5 condition. Fig.3 shows the results of three-layer WPD decomposition and demonstrates complete coverage of the signal frequency bands. Fig.4 shows the statistical analysis of the energy at each node in the third layer. Node AAA exhibits dominant energy, indicating that the signal energy is con-

centrated primarily in the low-frequency subbands and that low frequencies contribute more significantly to the signal. Thus, reconstruction was performed on node AAA, and the wavelet packet energy, absolute wavelet packet mean, and wavelet packet mean were extracted as time-frequency domain features.



(a) Raw signal

(b) Three-level wavelet packet decomposition results

Fig.3 Original Signal and Three-Level Wavelet Packet Decomposition Results

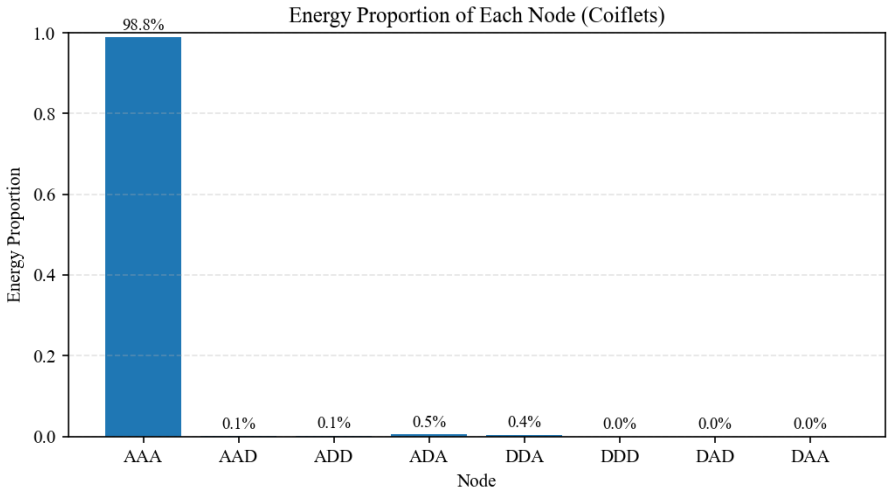
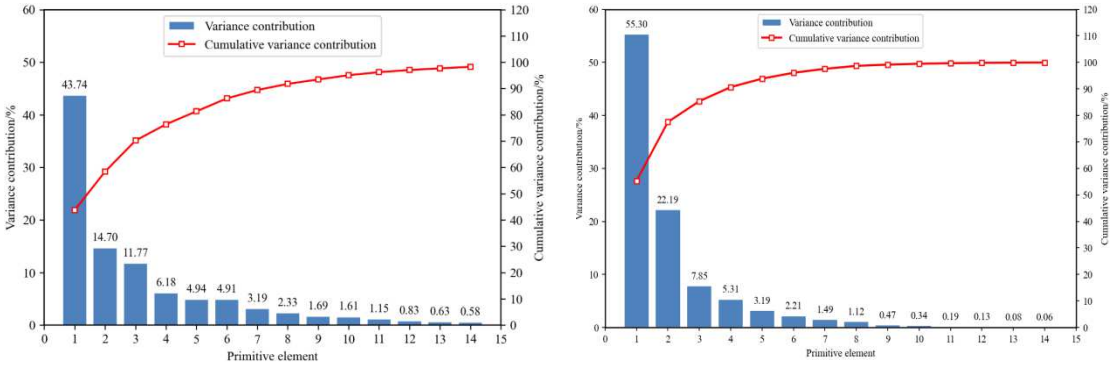


Fig.4 Energy Distribution Proportion of Third-Level Wavelet Packet Decomposition

2.1.3 PCA dimension reduction

A total of 57 features were retained for each sample by extracting features from force and current signals in the time, frequency, and time-frequency domains. However, these features may not necessarily exhibit a strong correlation with surface roughness, and there is redundant overlap among them. This increases the number of parameters during model training and impacts model accuracy and training efficiency. To address these issues, this study employs principal component analysis (PCA) to reduce the dimensionality of the extracted features. First, the features are normalized within the range of 0 to 1 through dimensionless processing. Next, the covariance matrix among the features is computed to derive their eigenvalues and eigenvectors. Finally, significant

eigenvalues with substantial influence on signal features are identified based on the variance contribution rate of eigenvalues. Fig.5(a) and 5(b) show the contribution rates of principal components for force and current signals, respectively.



(a) Cutting force signal (b) Current signal

Fig.5 Variance Contribution and Cumulative Variance Contribution of Each Component in Cutting Force and Current Signals

Figure 8 shows that the cumulative variance contribution of the top 10 principal components of the cutting force signal is 95.06%, while the cumulative variance contribution of the top six principal components of the current signal is 96.05%. These percentages represent 95.06% and 96.05%, respectively, of the information from the original factors. Thus, this study selected the top 10 principal components of the cutting force signal and the top six principal components of the current signal as inputs for the FiBiCAT model to predict surface roughness.

2.2 FiBiCAT prediction model framework

This study proposes a method for predicting surface roughness that integrates dynamic current signals, cutting force signals, and static process parameter information, as shown in Fig.6. A milling experimental platform was established using 7075 aluminum alloy as the research subject. The collected cutting force and current signals were analyzed and preprocessed to eliminate environmental interference. Feature extraction was then performed in the time and frequency domains using wavelet packets. Then, Principal Component Analysis (PCA) was applied to calculate the weight distribution of each feature and assign corresponding weight coefficients. This approach reduced feature complexity while retaining key information, yielding a high-quality dataset. To address the instability of the evaluation results caused by the limited number of signal samples, this study employed 5-fold cross-validation for model training and averaged the 5-fold results to determine the final performance metric. Building upon this foundation, this paper proposes a FiLM-modulated, bidirectional, cross-attention fusion Transformer network for predicting surface roughness. The model converts current and cutting force features into token sequences and achieves cross-signal fusion through bidirectional cross-attention, which fully exploits the complementary information between the two signal types. The model simultaneously incorporates process parameters to generate the FiLM modulation term. The Transformer encoder captures sequence correlations. Finally, a regression head outputs the surface roughness prediction value.



Fig.6 Overall Framework Diagram

2.3 FiLM-modulated bidirectional cross-attention fusion transformer network

This study introduces FiBiCAT, a bidirectional cross-attention fusion Transformer network that incorporates feature linear modulation. FiBiCAT consists of a serial connection of the FiLM linear modulation module, the bidirectional cross-attention fusion module, and the Transformer global fusion regression module. Its overall structure is shown in Fig.7. After PCA dimensionality reduction, the model takes three process parameters—current features, cutting force features, and cutting depth, feed rate, and machining sequence—as inputs, where the two signal types complete token sequence encoding. The process parameters generate modulation parameters β and γ through a multilayer perceptron network. These parameters then apply FiLM modulation to both signal types. Bidirectional cross-attention learns latent correlations between the current and cutting force signals while enhancing feature expression. The enhanced signal features and token-encoded process parameters then jointly feed into the Transformer encoder, which captures sequence correlations. The regression head outputs prediction results.

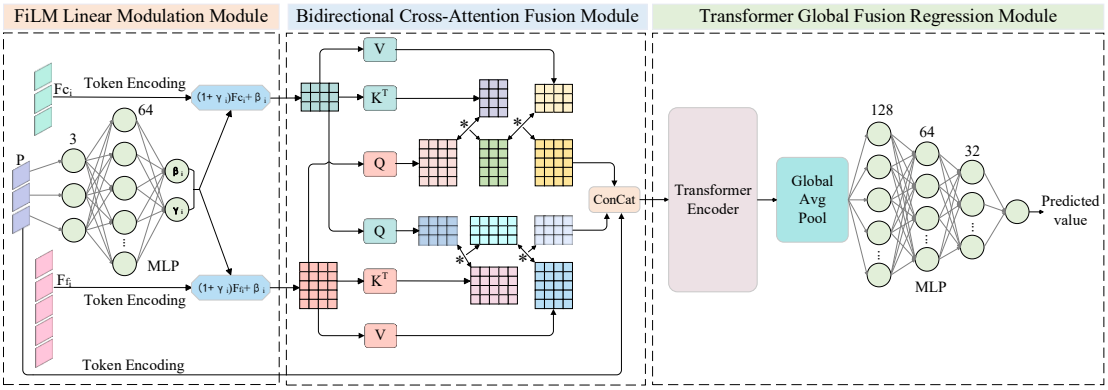


Fig.7 Overall Structure Diagram of the Model

2.3.1 FiLM linear modulation module

The current features, which are reduced via principal component analysis (PCA), cutting force features, and three process parameters—cutting depth, feed rate, and machining sequence—are uniformly encoded into token sequences. This process maps the original signal's numerical features into a vector representation that can be processed within the model. The cutting depth, feed rate, and machining sequence are computed using a multilayer perceptron network with an input layer of three, a hidden layer of 64, and an output layer of two. This yields two modulation parameters: β and γ. β controls the offset of the input signal, altering its baseline to adapt the overall position of the input signal to varying machining conditions. γ controls the scaling of the input signal by amplifying or reducing its amplitude to better align with the process parameters. The generated β and γ parameters uniformly adjust the encoded token sequences representing current and cutting force features. These parameters enable the model to adapt the representation of the input signal flexibly based on varying process parameter information, thereby facilitating accurate predictions across different machining conditions. The adjusted current and cutting force features are then transmitted as outputs to the subsequent module.

For example, the current feature after PCA is:

$$\mathbf{x}^{(c)} = \left[x_1^{(c)}, x_2^{(c)}, \ldots, x_{F_c}^{(c)} \right] \tag{1}$$

For the jth scalar feature $x_j^{(c)}$, an independent linear mapping is used to obtain a token vector:

$$\mathbf{t}_j^{(c)} = \mathbf{W}_j^{(c)} x_j^{(c)} + \mathbf{b}_j^{(c)} \in R^D \tag{2}$$

Second, arrange all the tokens in a sequence to create a token sequence matrix:

$$\mathbf{T}^{(c)} = \left[\mathbf{t}_1^{(c)}; \mathbf{t}_2^{(c)}; \ldots; \mathbf{t}_{F_c}^{(c)} \right] \in R^{F_c \times D} \tag{3}$$

Similarly, the cutting force token sequence：

$$\mathbf{t}_j^{(f)} = \mathbf{W}_j^{(f)} x_j^{(f)} + \mathbf{b}_j^{(f)} \in R^D, \quad \mathbf{T}^{(f)} \in R^{F_f \times D} \tag{4}$$

Process Parameter Token Sequence：

$$\mathbf{t}_k^{(p)} = \mathbf{W}_k^{(p)} p_k + \mathbf{b}_k^{(p)} \in R^D, \quad \mathbf{T}^{(p)} \in R^{3 \times D}, k \in \{1,2,3\} \tag{5}$$

β and γ adjust the sequence of features for unified encoding：

$$Y = (1+\gamma) \square X + \beta \tag{6}$$

2.3.2 Bidirectional cross-attention fusion module

The current features and the cutting force features that are output after the linear modulation of the previous module, FiLM, are passed as inputs to the bidirectional cross-attention module. These features undergo a linear transformation to map into the query vector (Q_c), the key vector (K_c), and the value vector (V_c). Similarly, the cutting force features undergo a linear transformation to map into the query vector (Q_f), the key vector (K_f), and the value vector (V_f). The core of the attention mechanism involves calculating the similarity between each query and all keys, and then weighting the corresponding value based on this similarity. This study employs a bidirectional cross-attention mechanism because the current and cutting force features collectively reflect the tool-workpiece interaction during machining. Incorporating cutting force features when analyzing current features supplements cutting state information that is difficult to characterize directly through current features alone. Likewise, integrating current features when analyzing cutting force features helps capture variations in system load and energy consumption. Through bidirectional feature interaction, key characteristics related to surface roughness can be characterized more comprehensively. The specific process is as follows: First, we calculate the similarity between the current feature, Q_c, and all cutting force features, K_f. This yields the similarity values between each current feature, Q_c, and all corresponding cutting force features, K_f. Then, the attention weights are derived using the Softmax function. A higher similarity value indicates a stronger correlation between a cutting force feature (K_f) and a specific current feature (Q_c). Consequently, the V_f corresponding to that cutting force feature receives greater weight during the weighted fusion process, signifying its enhanced contribution to current feature representation updates and feature fusion. The opposite is also true, enabling bidirectional information exchange and feature enhancement. Finally, the weighted current and cutting force features are concatenated to form a new sequence feature. This sequence feature is then concatenated with the process parameters encoded by the token sequence, yielding a more comprehensive feature representation. This representation is then input into the next module. These steps enable the module to achieve mutual learning and information fusion between current and cutting force features. This improves the model's ability to understand the relationships between features, thereby enhancing the accuracy of surface roughness prediction.

The similarity between the Q_c characteristic of the current and the K_f characteristic of the cutting force is as follows:

$$Similarity_{cf} = Q_c \cdot K_f \tag{7}$$

The similarity between the cutting force characteristic (Q_f) and the current characteristic (K_c) is as follows:

$$Similarity_{fc} = Q_f \cdot K_c \tag{8}$$

The similarity is normalized into attention weights using the softmax function to ensure that each feature's contribution to the current target is represented as a probability. This guarantees that each feature's influence falls between 0 and 1.The attention weights for the current feature, Q_c, and the cutting force feature, K_f, are:

$$Attention\ Weight_{cf} = Softmax(Q_c \cdot K_f) \tag{9}$$

The attention weights for the cutting force feature (Q_f) and the current feature (K_c) are as fol-

lows:

$$Attention\ Weight_{fc} = Softmax(Q_f \cdot K_c) \tag{10}$$

The model focuses more on information relevant to surface roughness prediction by weighting more important features to enhance them, thereby improving its accuracy. The current feature V_c and the cutting force feature V_f are weighted as follows:

$$Weighted\ Value_c = Attention\ Weight_{cf} \times V_f \tag{11}$$

$$Weighted\ Value_f = Attention\ Weight_{fc} \times V_c \tag{12}$$

Concatenate the Weighted $Value_c$ of the current feature with the Weighted $Value_f$ of the cutting force feature to create a new feature sequence.

$$Fusion = Concat(Weighted\ Value_c, Weighted\ Value_f) \tag{13}$$

The newly merged feature sequence is concatenated with the token-encoded process parameters, forming a more comprehensive feature representation.

$$Final\ Fusion = Concatenate(Fusion, P) \tag{14}$$

The concatenated features will ultimately be passed to the next module as its input.

2.3.3 Transformer global fusion regression module

The concatenated feature vectors output from the previous module are fed into the Transformer encoder. As illustrated in Figure 8, the Transformer encoder comprises three components: multi-head self-attention, a feedforward neural network, and layer normalization. First, the input feature vector passes through the multi-head self-attention layer. This layer analyzes the relationship between the current, cutting force, and process parameters from multiple distinct perspectives, with each perspective focusing on different aspects of the information. The results are then fused and fed into the feedforward neural network. Multi-head attention primarily facilitates interaction between different features, enabling each feature to incorporate information from the others while maintaining a linear combination as the core mechanism. The feedforward neural network performs a further nonlinear transformation on the features fused by the multi-head attention layer, thereby enhancing the model's expressive capability. Finally, layer normalization standardizes the output results, improving the model training process. The feature vectors output by the Transformer are converted into a fixed-length global feature vector via global average pooling, which is calculated using the following formula: This global feature vector is fed into a multilayer perceptron regression network with an input layer of 128, two hidden layers of 64 and 32 neurons, respectively, and an output layer of 1 to predict the surface roughness value.

$$Global\ Feature\ Vector = \frac{1}{N}\sum_{i=1}^{N} Feature_i \tag{15}$$

Here, N represents the number of features, and Feature_i represents the value of each feature.

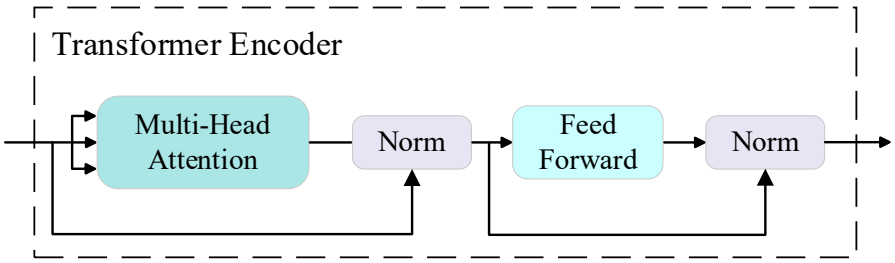


Fig.8 Transformer Encoder Structure Diagram

3 Experimental research

3.1 Experimental preparation

To validate the proposed method of predicting surface roughness in aluminum alloy milling, an independent milling test platform was built to collect the necessary experimental data for the prediction model. The experiments were conducted on a VDL-1000E CNC milling machine (Fig.9). The milled workpiece was made of 7075 aluminum alloy and had the following dimensions: 100 mm × 100 mm × 100 mm. A 10-mm-diameter, four-flute, tungsten carbide end mill was used to perform 80 pass milling operations, each 100 mm long, during the milling process.



Fig.9 Experimental Setup

During flat milling operations, the tool angle remains constant and perpendicular to the workpiece surface throughout the process. Therefore, the three primary determinants of surface quality are spindle speed, feed rate, and cutting depth. To account for practical milling conditions and improve the universality of the proposed surface roughness prediction model, the process parameters must closely resemble the actual machining conditions. The parameter ranges established for this study are as follows: Spindle speed: n = 3000 rpm, Cutting depth: ap = 0.1–0.8 mm, Feed per tooth: f = 0.02–0.11 mm. These parameters yielded 80 sets of sample data and 80 sets of surface roughness data. Selected process parameter schemes are shown in Table 1.

Table 1 Selected Process Parameters

Experiment Number	n（r/min）	a_p（mm）	f（mm）
1	3000	0.1	0.02
2	3000	0.1	0.03
⋮	⋮	⋮	⋮
10	3000	0.1	0.11

11	3000	0.2	0.02
12	3000	0.2	0.03
⋮	⋮	⋮	⋮
20	3000	0.2	0.11
⋮	⋮	⋮	⋮
71	3000	0.8	0.02
72	3000	0.8	0.03
⋮	⋮	⋮	⋮
80	3000	0.8	0.11

The most critical dynamic signals affecting surface roughness are force, vibration, and current signals. This study uses a Kistler 9257B rotary force transducer to acquire force signals and uses current transformers to acquire current signals. The sampling frequency is 10 kHz to ensure accurate and comprehensive information is obtained from the actual machining process. The collected data can be categorized as either dynamic signal data or static process data. Dynamic signal data includes force and current signals, while static process data includes spindle speed, feed rate, cutting depth, and sequence number of the cutting operation. To measure surface roughness, a portable roughness tester was used to determine the Ra value of the machined surface after each milling operation. Accurate surface roughness measurements were achieved by taking measurements at various points on the workpiece surface and calculating the arithmetic mean of these results.

3.2 Experimental setup

3.2.1 Cross-validation

After undergoing feature extraction in the time, frequency, and time-frequency domains, followed by dimensionality reduction, the current signal and cutting force signal, along with process parameter information, are used as inputs. Surface roughness serves as the output. Since the experimental dataset comprises only 80 sets, the limited data volume necessitates caution. If the training and test sets are divided into a single 7:3 split, model performance can only be evaluated based on this partition. This approach is susceptible to significant fluctuations caused by the uneven distribution of certain operating conditions, which makes it difficult to accurately reflect the model's generalization capability. To reduce randomness and utilize the limited samples more effectively, this study employs 5-fold cross-validation[26], the principle of which is illustrated in Fig.10. The 80 datasets in this experiment are divided into five mutually exclusive subsets, each containing 16 datasets. For each iteration, four subsets are selected for model training, and the remaining subset is used for performance evaluation. This process is repeated five times, ensuring that each sample participates in exactly one evaluation. The final model performance is measured by averaging the results from the five tests, yielding a more reliable and robust evaluation of the model.

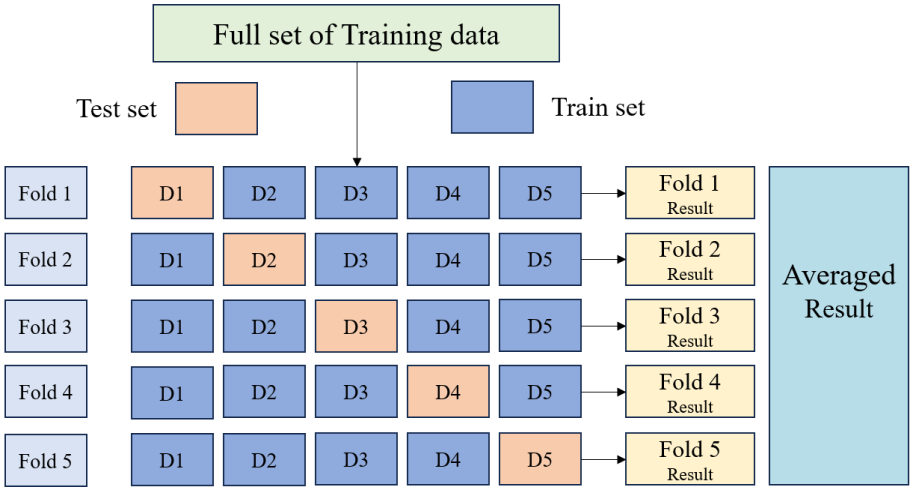


Fig.10 Schematic diagram of five-fold cross-validation.

3.2.2 Parameter and hyperparameter configuration

Numerous hyperparameters are involved during model training and prediction, and their configuration significantly impacts the final results. This paper's experimental environment was built using the Python 3.9 programming language and the PyTorch 2.0.1 deep learning framework. Computational acceleration was achieved via CUDA 11.8. The hardware platform features an NVIDIA GeForce GTX 1650 Ti graphics card with 4 GB of VRAM. To achieve optimal model convergence, the following parameters were set after multiple rounds of pre-training and optimization adjustments: Initial learning rate of 4e-4, Adam optimizer, 300 epochs, and batch size of 8. A StepLR learning rate scheduling strategy was adopted with a decay step of 10 epochs and gamma of 0.5. This strategy was maintained consistently across all comparative experiments. Table 2 shows detailed training parameter configurations.

Table 2 Hyperparameter Configuration Table

Hyper-parameters	Value
Batch size	8
Epoch	300
Optimizer	Adam
Loss function	SmoothL1
Learning rate	4e-4
Learning rate scheduler	StepLR

3.2.3 Performance metrics

This study uses mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R^2) as the primary evaluation metrics to analyze and compare prediction outcomes. Specific formulas are presented in equations (16) through (18). MAE is characterized by its rapid computation and ease of operation; a smaller MAE indicates greater computational accuracy of the model. RMSE is more sensitive to outliers in predictions; lower RMSE values indicate more stable and reliable prediction results from the neural network model. R^2 represents the model's resilience to variations; values closer to 1 indicate stronger resistance to change. Integrating these metrics provides a comprehensive evaluation of the model's predictive performance, ensuring its accuracy, reliability, and robustness.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \bar{y}| \tag{16}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \tag{17}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(\hat{y}_i - y_i\right)^2}{\sum_{i=1}^{n}\left(\bar{y} - y_i\right)^2} \tag{18}$$

In this formula, y_i represents the actual value, $\hat{y}_i$ represents the predicted value, y represents the sample mean, and n represents the size of the dataset.

4 Results and discussion

4.1 FiBiCAT prediction model

Fig.11 shows that the surface roughness prediction curves generated by FiBiCAT through 5-fold cross-validation on 80 sample sets are highly consistent with the actual curves across all folds. This demonstrates the model's ability to accurately predict the overall surface roughness trend, indicating stable performance with different data partitions. Additionally, the overall error is minimal, with overestimation and underestimation occurring alternately. This suggests that there is no significant systematic bias in the model. Most errors are concentrated at local extrema or rapidly changing sections of the true curve. This indicates larger prediction deviations under extreme conditions or sudden operational changes. Nevertheless, the model shows high consistency and reliability for most samples.

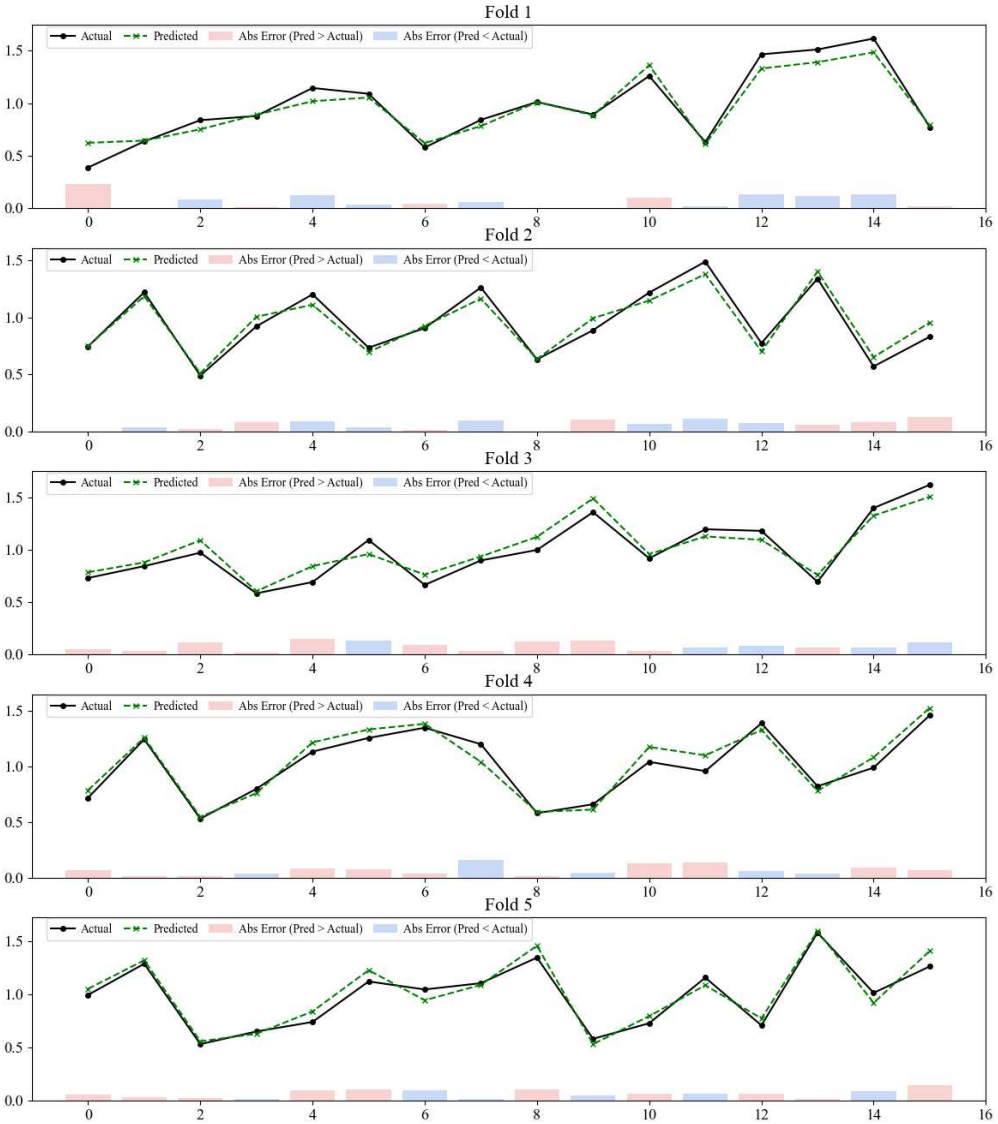


Fig.11 FiBiCAT Surface Roughness Prediction Results

According to Table 3, the model achieved the following average metrics in the 5-fold cross-validation test: MAE = 0.0709, RMSE = 0.0842, and R^2 = 0.9210. R^2 = 0.9210. Specifically, the MAE ranged from 0.0634 to 0.0846; the RMSE, from 0.0739 to 0.0957; and the R^2, from 0.8953 to 0.9348. These results suggest that the model has high predictive accuracy and minimal fluctuation across different cross-validation partitions, indicating strong generalization capability and stability.

Table 3 FiBiCAT Surface Roughness Prediction Performance Metrics

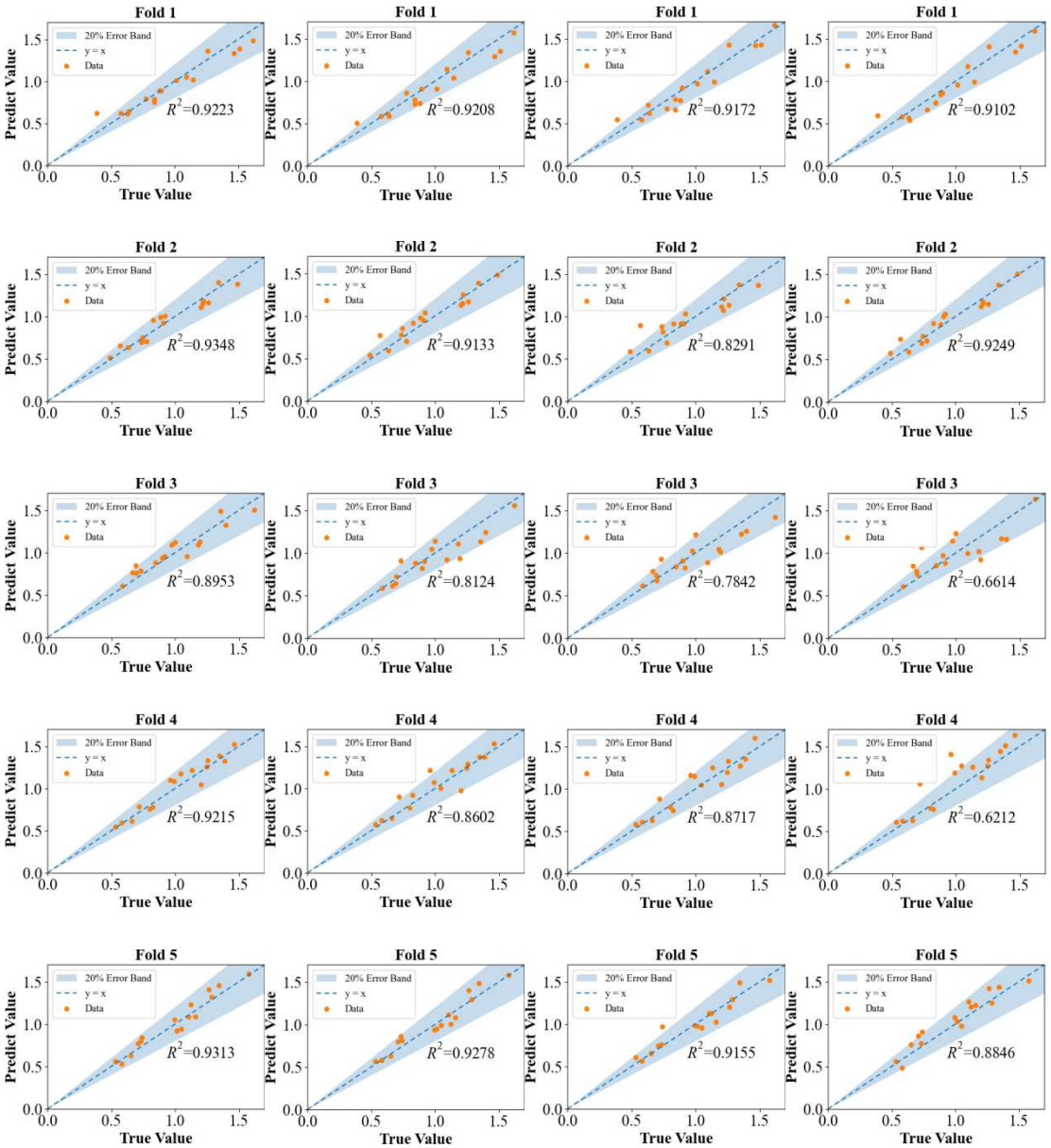
Fold	MAE	RMSE	R^2
1	0.0711	0.0957	0.9223
2	0.0634	0.0739	0.9348
3	0.0846	0.0934	0.8953
4	0.0680	0.0809	0.9215
5	0.0675	0.0773	0.9313
Mean	0.0709	0.0842	0.9210

4.2 Model comparison

To further validate this model's superiority, this study compares the FiBiCAT model with

three mainstream neural network models: MLP, GRU, and Transformer. The sample data, inputs, and outputs for each prediction model are identical to those of the FiBiCAT model. Fig.12 (a–d) shows the R^2 prediction results for surface roughness across models. The blue dashed line shows where the predicted and actual surface roughness values coincide. The blue shaded area represents a ±20% error band. Orange scatter points illustrate the relationship between the predicted and true values for the test samples. Points that are closer to the blue dashed line indicate higher prediction accuracy. Fig.13(a–b) compares the mean absolute error (MAE) and root mean square error (RMSE) of surface roughness predictions across models. Performance metrics are summarized in Table 4.

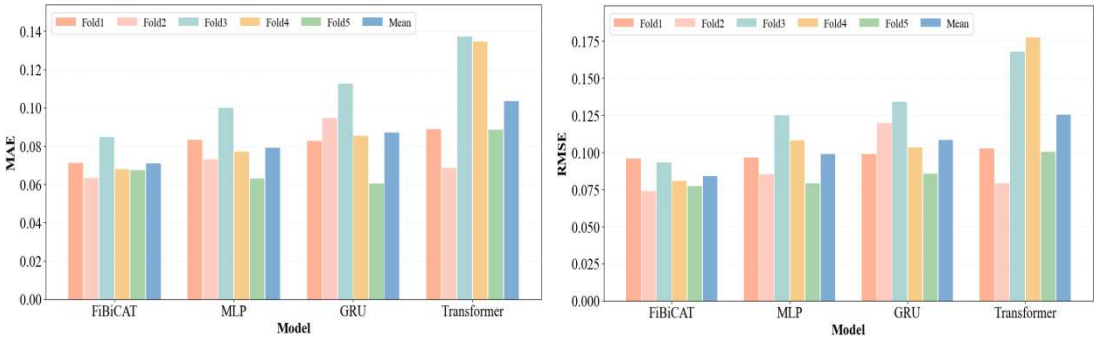
Fig.12 shows that the FiBiCAT model exhibits tighter scatter around the blue dashed line in each fold. More sample points fall within the ±20% error band. This indicates that the model's predictions are closer to the actual values. In contrast, the scatter points for the MLP and GRU models are more dispersed, and the Transformer model is even more dispersed. These results suggest that these models have relatively weaker fitting and generalization capabilities for this task, which further validates the accuracy of the FiBiCAT model established in this study.



(a) FiBiCAT **(b) MLP** **(c) GRU** **(d) Transformer**

Fig.12 Comparison of R^2 Values for Surface Roughness Predictions Across Different Models

As shown in Fig.13 and Table 4, the prediction results indicate that FiBiCAT maintains a low error level across all folds, with relatively small fluctuations among the five folds, compared with other models. The average performance metrics of the FiBiCAT model, MAE and RMSE, are 0.0709 and 0.0842, respectively. Both are lower than those of other models. R^2 is 0.9210, which is closer to 1 than that of other models. These results suggest that the FiBiCAT model effectively reduces average prediction errors and significantly enhances interpretability regarding roughness variations. Consequently, the proposed FiBiCAT model provides more stable and precise surface roughness predictions under small sample sizes and varying cross-validation splits, offering significant advantages in surface roughness prediction.



(a) MAE plot (b) RMSE plot

Fig.13 Comparison of MAE and RMSE for Surface Roughness Predictions Across Different Models

Table 4 Average Metrics for Surface Roughness Prediction Performance Across Different Models

Models	MAE	RMSE	R^2
Transformer	0.1035	0.1256	0.8004
GRU	0.0871	0.1084	0.8635
MLP	0.0778	0.0988	0.8869
FiBiCAT	0.0709	0.0842	0.9210

4.3 Ablation experiment

This study conducted ablation experiments on the FiBiCAT model to validate the effectiveness of the multi-signal fusion and linear modulation strategy. To control variables and ensure experimental rigor, only one module was modified or removed at a time while keeping all other parameters consistent. The results of the ablation experiments are presented in Table 5.

Table 5 Ablation Experiment Results

Models	MAE	RMSE	R^2
NoFiLM	0.0863	0.1068	0.8709
NoGating	0.0786	0.0999	0.8773
NoCrossAttn	0.0751	0.0953	0.8957
NoSelf-Attention	0.0772	0.0970	0.8886
NoModalityDropout	0.0762	0.0978	0.8839
ParamsOnly	0.0826	0.1063	0.8656
Force+Params	0.0711	0.0926	0.9019
Current+Params	0.0784	0.0970	0.8912
FiBiCAT	**0.0709**	**0.0842**	**0.9210**

1）NoFiLM：Remove the process parameter-based FiLM linear modulation.
2）NoGating：Remove learnable feature gating.
3）NoCrossAttn：Eliminate the cross-signal attention between the current signals and the force signals.
4）NoSelf-Attention：Remove the self-attention mechanism from the fusion transformer.
5）NoSignalDropout：Remove the signal-level dropout regularization.
6）ParamsOnly：Only enter process parameter information. Set the current and force signals to zero.
7）Force+Params：Only input force signals and process parameter information. Set the current signal to 0.
8）Current+Params：Only input current signals and process parameter information. Set the force signal to zero.

As shown in Table 8, removing the process-parameter-based FiLM linear modulation and learnable feature gating significantly degrades model performance. These results suggest that adaptive modulation of signal features based on process parameters substantially improves prediction accuracy, and that feature recalibration helps suppress redundant and noisy features. Eliminating the cross-signal cross-attention mechanism also degrades model performance, suggesting that cross-signal fusion learning between current and force signals is crucial for roughness prediction. Canceling the signal random dropout strategy also leads to model performance degradation, demonstrating that regularization enhances model robustness. Single-signal comparisons reveal that roughness prediction accuracy is higher with force signals than with current signals. Using only process parameters yields the poorest results, indicating that force signals have stronger representational capabilities. Using multi-sensor signal inputs significantly improves prediction accuracy compared to single-sensor inputs, showing that the two signal types provide complementary representations of the machining state. Ablation results reveal the FiLM linear modulation module, the bidirectional cross-attention module, and the Transformer global fusion module work synergistically, further validating the effectiveness of the proposed FiBiCAT model.

5 Conclusion

This paper presents FiBiCAT, a multi-source surface roughness prediction model that integrates dynamic current signals, cutting force signals, and static process parameters. Through preprocessing the acquired current and cutting force signals, feature extraction, and applying principal component analysis (PCA) for dimensionality reduction, key feature information is retained while reducing feature complexity. This provides high-quality input data for the model. The following conclusions are drawn:

（1）We have established a model architecture that integrates current signals, cutting force signals, and process parameters. This architecture uses features extracted from multiple signal sources and considers the effect of process parameters, such as cutting depth and feed rate, on cutting conditions. This allows the model to learn more comprehensive and effective information.

（2）Linearly modulating the current and cutting force signal features using process parameters assigns different weights to the two signal types. Bidirectional cross-attention combines the two signals to complement each other, thereby enhancing the accuracy of roughness prediction.

（3）Using the 5-fold cross-validation dataset partitioning method, the FiBiCAT model achieved an average R^2 performance metric of 0.9210, with respective MAE and RMSE values of 0.0709 and 0.0842. Compared to other mainstream neural network models, the FiBiCAT model demonstrates higher prediction accuracy.

Declarations

Funding This work was supported by the National Natural Science Foundation of China: Outstanding Youth Science Fund of Heilongjiang Province (JQ2024E008).2023 Domestic Numerical Control System Application Demonstration Project: 2023ZY01076-11

Competing interests The authors declare no competing interests.

Data availability The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Author Contributions All authors contributed to the study conception and design.Song Li designed the methodology framework, wrote the original draft, and revised and corrected the manuscript.Boyang Meng provided content guidance and critically reviewed the manuscript.Zhongjie Li supervised the coding and figure drawing.Fuxing Xiang and Rui Ding performed the investigation and data management.Caixu Yue revised the content, reviewed the format, and acquired funding.Xianli Liu was responsible for project management.

References

[1] Yang C A O, Yejun Z H U, Wenfeng D, et al (2022) Vibration coupling effects and machining behavior of ultrasonic vibration plate device for creep-feed grinding of Inconel 718 nickel-based superalloy. CJA 35(2): 332-345.https://doi.org/10.1016/j.cja.2020.12.039

[2] Kuang W, Qing M, Wenfeng D, et al (2022) Fretting wear behaviour of machined layer of nickel-based superalloy produced by creep-feed profile grinding. CJA 35(10): 401-411.https://doi.org/10.1016/j.cja.2021.10.007

[3] Roy B, Mukhopadhyay T, Dey S (2022) Polynomial neural network based probabilistic hydrodynamic analysis of two-lobe bearings with stochasticity in surface roughness. Tribol.Int 174: 107733.https://doi.org/10.1016/j.triboint.2022.107733

[4] Kim D H, Kim T J Y, Wang X, et al (2018) Smart machining process using machine learning: A review and perspective on machining industry. INT J PR ENG MAN-GT 5(4): 555-568.https://doi.org/10.1007/s40684-018-0057-y

[5] Pimenov D Y, Kiran M, Khanna N, et al (2023) Review of improvement of machinability and surface integrity in machining on aluminum alloys. Int.J.Adv.Manuf.Technol 129(11): 4743-4779.https://doi.org/10.1007/s00170-023-12630-4

[6] Liu C, Huang Z, Huang S, et al (2023) Surface roughness prediction in ball screw whirlwind milling considering elastic-plastic deformation caused by cutting force: Modelling and verification. MEAS 220: 113365.https://doi.org/10.1016/j.measurement.2023.113365

[7] Yeganefar A, Niknam S A, Asadi R (2019) The use of support vector machine, neural network, and regression analysis to predict and optimize surface roughness and cutting forces in milling. Int.J.Adv.Manuf.Technol 105(1): 951-965.https://doi.org/10.1007/s00170-019-04227-7

[8] Nagai S, Yoshida I, Oshiro K, et al (2022) Acceleration of surface roughness evaluation using RANSAC and least squares method for Running-in wear process analysis of plateau surface. MEAS 203:

111912.https://doi.org/10.1016/j.measurement.2022.111912

[9] Kong D, Zhu J, Duan C, et al (2020) Bayesian linear regression for surface roughness prediction. Mech.Syst.Sig.Process 142: 106770.https://doi.org/10.1016/j.ymssp.2020.106770

[10] Tian W, Zhao F, Sun Z, et al (2022) Prediction of surface roughness using fuzzy broad learning system based on feature selection. Manuf.Syst 64: 508-517.https://doi.org/10.1016/j.jmsy.2022.07.012

[11] Dedeakayoğulları H, Kaçal A, Keser K (2022) Modeling and prediction of surface roughness at the drilling of SLM-Ti6Al4V parts manufactured with pre-hole with optimized ANN and ANFIS. MEAS 203: 112029.https://doi.org/10.1016/j.measurement.2022.112029

[12] Zhou G, Xu C, Ma Y, et al (2020) Prediction and control of surface roughness for the milling of Al/SiC metal matrix composites based on neural networks. ADV MANUF 8(4): 486-507.https://doi.org/10.1007/s40436-020-00326-x

[13] Yao Z, Shen J, Wu M, et al (2023) Position-dependent milling process monitoring and surface roughness prediction for complex thin-walled blade component. Mech.Syst.Sig.Process 198: 110439.https://doi.org/10.1016/j.ymssp.2023.110439

[14] Chen J, Lin J, Zhang M, et al (2024) Predicting surface roughness in turning complex-structured workpieces using vibration-signal-based gaussian process regression. SENSORS-BASEL 24(7): 2117.https://doi.org/10.20944/preprints202402.0486.v1

[15] Zhao M, Xue B, Li B, et al (2024) Ensemble learning with support vector machines algorithm for surface roughness prediction in longitudinal vibratory ultrasound-assisted grinding. PRECIS ENG 88: 382-400.https://doi.org/10.1016/j.precisioneng.2024.02.018

[16] Bennett K S, DePaiva J M, Veldhuis S C (2024) An integrated framework for a multi-material surface roughness prediction model in CNC turning using theoretical and machine learning methods. Int.J.Adv.Manuf.Technol 131(7): 3579-3598.https://doi.org/10.1007/s00170-024-13201-x

[17] Xiao Y, Zheng S, Shi J, et al (2023) Knowledge graph-based manufacturing process planning: A state-of-the-art review. J.Manuf.Syst 70: 417-435.https://doi.org/10.1016/j.jmsy.2023.08.006

[18] Yu Y, Karimi H R, Shi P, et al (2024) A new multi-source information domain adaption network based on domain attributes and features transfer for cross-domain fault diagnosis. Mech.Syst.Sig.Process 211: 111194.https://doi.org/10.1016/j.ymssp.2024.111194

[19] Bernini L, Malguzzi U, Albertelli P, et al (2024) Hybrid prognostics to estimate cutting inserts remaining useful life based on direct wear observation. Mech.Syst.Sig.Process 210: 111163.https://doi.org/10.1016/j.ymssp.2024.111163

[20] Fang R, Yi H, Shu A, et al (2022) Illumination-robust milling surface roughness machine vision inspection based on MAML++ network. Opt.Eng 61(12): 124105-124105.https://doi.org/10.1117/1.oe.61.12.124105

[21] Guo M, Zhou J, Li X, et al (2023) Prediction of surface roughness based on fused features and ISSA-DBN in milling of die steel P20. Sci. Rep 13, 15951 [EB/OL].https://doi.org/10.1038/s41598-023-42968-4

[22] Wang Y, Wang Y, Zheng L, et al (2022) Online surface roughness prediction for assembly interfaces of vertical tail integrating tool wear under variable cutting parameters. SENSORS-BASEL 22(5): 1991.https://doi.org/10.3390/s22051991

[23] Xiao G, Zhu B, Zhang Y, et al (2023) CAN-net: a multi-hidden layer attention deep learning method for surface roughness prediction during abrasive belt grinding of superalloy with local weights. Int.J.Artif.Intell.Tools 32(06): 2350024.https://doi.org/10.1142/S0218213023500240

[24] Wang Z, Gao D, Lu Y, et al (2025) A mutual cross-attention fusion network for surface roughness prediction in robotic machining process using internal and external signals. J.Manuf.Syst 82: 284-300.https://doi.org/10.1016/j.jmsy.2025.06.018

[25] Zhang T, Guo X, Fan S, et al (2022) AMS-Net: Attention mechanism based multi-size dual light source network for surface roughness prediction. J.Manuf.Processes 81: 371-385.https://doi.org/10.1016/j.jmapro.2022.07.009

[26] Val S, Lambán M P, Lucia J, et al (2024) Analysis and Prediction of Wear in Interchangeable Milling Insert Tools Using Artificial Intelligence Techniques. Appl.Sci.-Basel 14(24): 11840.https://doi.org/10.3390/app142411840