Real-time monitoring method for surface roughness of γ-TiAl alloy based on deep learning of time-frequency diagram

Yongxian Wu
Linyan Liu
Lei Huang
zhenhua wang (✉ niatwzh17@163.com)
Nanjing University of Science and Technology

Research Article

Keywords: γ-TiAl alloy, Surface roughness, Deep learning, Time-frequency diagram, Transfer learning

Posted Date: May 18th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2929517/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 October 20th, 2023. See the published version at https://doi.org/10.1007/s00170-023-12453-3.
Real-time monitoring method for surface roughness of $\gamma$-TiAl alloy based on deep learning of time-frequency diagram

Yongxian Wu\textsuperscript{a}, Linyan Liu\textsuperscript{a,*}, Lei Huang\textsuperscript{a}, Zhenhua Wang\textsuperscript{a,*}

\textsuperscript{a}School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China

Abstract: $\gamma$-TiAl alloy is a typically difficult material to machine, and machining defects such as grain pull-out and material spalling are common during machining, resulting in component scrap. As a result, it is critical to investigate the real-time monitoring method of surface roughness during milling. A new model for predicting surface roughness is proposed in this article. Based on the idea of deep learning, the prediction problem of surface roughness is turned into a classification problem. The features of the force signal are extracted using a continuous wavelet transform, and the one-dimensional signal is converted into a two-dimensional time-frequency diagram to obtain more information. The transfer learning mechanism is introduced to make the model run faster and improve prediction accuracy. The overfitting problem is solved and the model's generalization ability is improved through batch normalization and dropout layer, and the prediction accuracy is increased by 3%. The results show that the method has high recognition accuracy, with a maximum classification accuracy of 98% and an average accuracy of 96.12%, a 6% improvement over the traditional model, allowing it to accurately realize real-time monitoring of surface quality in milling processing.

Keywords: $\gamma$-TiAl alloy; Surface roughness; Deep learning; Time-frequency diagram; Transfer learning

1 Introduction

The high specific strength, elastic modulus, low density, and good oxidation resistance of $\gamma$-TiAl alloy make it suitable for the fabrication of some complex structural and functional parts in the
field of aircraft engines, such as low-pressure turbines and engine blades\textsuperscript{[1,2]}. Despite the good mechanical and thermal properties of $\gamma$-TiAl alloy, the material's unique lamellar structure and mechanical characteristics, such as high room temperature brittleness and low thermal conductivity, cause it to generate large cutting forces and cutting heat during machining\textsuperscript{[3,4]}. This easily results in phenomena like delamination fracture, material pull-out, and spalling, which leads to a sharp decrease in the surface quality and fatigue resistance of the workpiece. Surface roughness is the primary index used to evaluate the surface quality of the workpiece\textsuperscript{[5]}, which is closely related to the fatigue strength and wear resistance of mechanical parts and has a significant impact on the product's stability and service life. As a result, real-time monitoring of the machined surface roughness of $\gamma$-TiAl alloy is critical.

Traditional surface roughness measurement employs a contact stylus device\textsuperscript{[6]}, which has drawbacks in terms of lengthy measurement times and challenging setup, and it is simple for materials with a low hardness to develop damage on the surface of the workpiece, which impairs the workpiece's quality\textsuperscript{[7]}. Many researchers developed mathematical models to predict the surface roughness of the workpiece to avoid the impact of contact measurement on the workpiece's surface quality. Gao et al.\textsuperscript{[8]} proposed a workpiece surface topography prediction model based on Z-MAP method that considers the parameter variations of the cutting system and tool wear. He et al.\textsuperscript{[9]} from Southeast University proposed a hybrid model that uses a Hidden Markov model based on Bayesian inference and a support vector machine to evaluate surface roughness. Han et al.\textsuperscript{[10]} constructed a two-dimensional finite element model of rough surface roughness using the W-M fractal function to numerically simulate the roughness evolution of the contact surface of the formed part. A surface roughness prediction model was proposed. The model can be applied to metal-formed parts where it is difficult to improve the surface quality by machining. However, these methods make it difficult to take into account external environmental factors, and it is challenging to estimate the surface roughness of the workpiece. Machine learning and deep learning are becoming more popular in machining as artificial intelligence technology advances\textsuperscript{[11]}. Wu et al.\textsuperscript{[12]} performed analysis, calculation, and frequency normalization on milling process vibration signals, and used the resulting high correlation features as input to an artificial neural network to predict surface roughness. Pimenov et al.\textsuperscript{[13]} used a neural network model to predict the surface roughness of a part by real-time monitoring of a CNC machine tool's spindle power. Cheng et al.\textsuperscript{[14]} created an extreme learning
machine model for surface roughness prediction and compared it to the response surface method and the support vector machine network; the results showed that the extreme learning machine network can achieve relatively higher accuracy in surface roughness prediction. Guo et al.\cite{15} proposed a hybrid feature selection method that selects features based on their correlation with surface roughness as well as hardware and time costs. A long short term memory network is used to predict the surface roughness. Yin et al.\cite{16} predicted the surface roughness of Ti-6Al-4V alloy during grinding based on a hybrid algorithm combining fuzzy neural network, and Taguchi’s empirical analysis with a compressed air measuring head.

The above review of the literature indicates that advanced artificial intelligence methods are more accurate for surface roughness prediction, but most current studies predict surface roughness values based on machining parameters, vibration parameters, and so on, and their work is primarily focused on finding decision coefficients. Second, the characteristics used for surface roughness prediction are typically chosen empirically, with no systematic review. Furthermore, the majority of existing forecasts are regression problems that are substantially hampered by data noise and erroneous predictions. Classification models, on the other hand, are less susceptible to noise interference and require less data\cite{17}. There has been little research on predicting surface roughness by monitoring cutting force signals in machining, and even fewer on predicting surface roughness for novel materials such γ-TiAl alloys.

To address these issues, a new surface roughness prediction method is suggested that collects cutting force signals during machining and uses a one-dimensional sliding window for overlapping sampling to increase the data scale and improve the accuracy of the neural network. The one-dimensional data is turned into a two-dimensional image by extracting a large number of features using the continuous wavelet transform, allowing the model to learn additional information. The improved residual neural network and transfer learning are used to perform online prediction of milling surface roughness, and the surface quality diagnosis problem is changed into a classification problem. Experiment validation was also carried out.

2 Theoretical basis

2.1 Residual network

As a typical deep learning model, the convolutional neural network (CNN) is powerful in
image processing and plays a key role in many relevant domains\textsuperscript{[18]}. It consists primarily of a convolutional layer, activation, pooling layers, a fully connected layer, and a classifier\textsuperscript{[19]}, and differs from a general neural network in that it has a local sensory field of view, shares weights throughout network training, and employs downsampling of the input signal. Fig. 1 depicts the basic structure:

**Fig. 1** Convolutional neural network structure

The structure of the convolutional neural network gradually expands in a deeper and wider direction as a result of advances in computer processing speed and data volume. Even though more information can be obtained as the network level deepens, the model will eventually hit saturation or degrade, reducing model performance\textsuperscript{[20,21]}. By directly transmitting the information from the image input, the residual structure of the residual network (ResNet) can get over the limits of standard convolutional neural networks and allow the network to learn fewer features, deepen to a deeper level, and extract deeper information features.

2.2 Continuous wavelet transform

To ensure that signal characteristics can be included in the frequency domain information, it is frequently required to capture both the high-frequency and low-frequency portions of the signal when performing time-frequency transformations on signals\textsuperscript{[22]}. The Fourier transform, one of the popular time-frequency transform techniques, has a fixed window size and cannot extract both high- and low-frequency data. The continuous wavelet transform (CWT) approach, on the other hand, modifies the window size in accordance with the signal's frequency to extract the high and low frequency signals. This technique can simultaneously achieve continuity for all panning and scaling
operations, ensuring the precision of time-frequency analysis. To classify surface roughness, this thesis uses the continuous wavelet transform as a feature extraction method\textsuperscript{[23]}. The equation for the continuous wavelet transform is defined in (1):

\[
WT_f(\alpha, \tau) = f(t), \varphi_{\alpha, \tau}(t) \leq \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t) \varphi^* \left(\frac{t-\tau}{a}\right) dt
\]  

\text{(1)}

Where \( WT_f(\alpha, \tau) \) denotes the wavelet transform coefficients, \( \alpha, \tau \) are two parameters of the scale and translation of wavelet transform. One-dimensional force signals can be converted into two-dimensional feature information through CWT, which is convenient for the later model to extract the key information points.

2.3 Transfer learning

To improve prediction performance, a lot of data must be tagged for training during the model-training process. Yet, this calls for a sizable workload and several evenly distributed datasets. As a result, for networks with unbalanced dataset categories and a limited number of datasets, training with transfer learning can significantly save time and increase accuracy.

Transfer learning is a deep learning extension that places more emphasis on the distributional disparities between various features. By training the network model and changing the weights of the network model to minimize disparities in the distribution of the feature domains between the origin and the target domain data, transfer learning can enhance the predictive power of neural networks\textsuperscript{[24,25]}.

3 Experimental design and data pre-processing

3.1 Introduction to the experimental platform

\( \gamma \)-TiAl alloy (Ti-47.5Al-2.5V-1.0Cr) developed by the General Iron and Steel Institute serves as the experimental material for this study. Table 1 shows the main mechanical properties of \( \gamma \)-TiAl alloy at high temperatures (800°C)\textsuperscript{[26]}. For the sake of the experiment, \( \gamma \)-TiAl alloy specimen was a wire cut into a rectangle with a square cross-section of 20mm × 20mm × 5mm. For the test, a carbide milling cutter with a tool diameter of \( d = 6\text{mm} \) was used. For the cutting force test, a force-measuring device supplied by Kistler was used, which included a Kistler9119AA2 force gauge, a Kistler5080 charge amplifier, and a Kistler5697A data collector. The sampling frequency
of cutting force is 10 kHz, which means 10,000 points per second. The force-measuring device is shown in Fig. 2(a), and the milling device is shown in Fig. 2(b).

<table>
<thead>
<tr>
<th>Properties</th>
<th>Ti-47.5Al-2.5V-1.0Cr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yield strength (MPa)</td>
<td>380</td>
</tr>
<tr>
<td>Tensile strength (MPa)</td>
<td>500</td>
</tr>
<tr>
<td>Elastic modulus (GPa)</td>
<td>151</td>
</tr>
<tr>
<td>Thermal conductivity (W/(m · °C))</td>
<td>23.1</td>
</tr>
<tr>
<td>Elongation (%)</td>
<td>6</td>
</tr>
<tr>
<td>Density (g/cm³)</td>
<td>3.87</td>
</tr>
</tbody>
</table>

*Table 1* Mechanical properties of γ-TiAl alloy

![Fig. 2 Experimental setup (a) The force-measuring device, (b) The milling device](image)

3.2 Experimental Settings

3.2.1 Basic Framework of the Experiment

Fig. 3 shows the construction of the experimental, which is divided into five modules: (1) Milling of γ-TiAl alloy; (2) Force single recording and $R_a$ measurement; (3) Wavelet time-frequency diagram and $R_a$ values; (4) surface roughness labels; (5) Deep learning model.
3.2.2 Data acquisition

This experiment intends to obtain workpieces with different surface roughness grades by changing machining parameters and under different tool wear states. To investigate the correlation between surface roughness and cutting force signals, a three-factor, three-level orthogonal test with variable spindle speed, feed rate, and depth of cut was developed. The experimental parameters are shown in Table 2. To get as numerous and thorough surface roughness values for γ-TiAl alloy as possible for each parameter, a milling tool was experimented with from a wear value of zero until the maximum wear value was obtained. Surface roughness values should, in theory, gradually increase as tool wear increases. The surface roughness of the workpiece was measured using a laser microscope (KEYENCE, VK-XX 100 series) after each experiment.
<table>
<thead>
<tr>
<th>No.</th>
<th>Spindle Speed (r/min)</th>
<th>Feed Rate (mm/r)</th>
<th>Depth of cut (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3000</td>
<td>0.010</td>
<td>0.10</td>
</tr>
<tr>
<td>2</td>
<td>3000</td>
<td>0.015</td>
<td>0.20</td>
</tr>
<tr>
<td>3</td>
<td>3000</td>
<td>0.020</td>
<td>0.15</td>
</tr>
<tr>
<td>4</td>
<td>3500</td>
<td>0.010</td>
<td>0.20</td>
</tr>
<tr>
<td>5</td>
<td>3500</td>
<td>0.015</td>
<td>0.15</td>
</tr>
<tr>
<td>6</td>
<td>3500</td>
<td>0.020</td>
<td>0.10</td>
</tr>
<tr>
<td>7</td>
<td>4000</td>
<td>0.010</td>
<td>0.15</td>
</tr>
<tr>
<td>8</td>
<td>4000</td>
<td>0.015</td>
<td>0.10</td>
</tr>
<tr>
<td>9</td>
<td>4000</td>
<td>0.020</td>
<td>0.20</td>
</tr>
</tbody>
</table>

The force-measuring system collected force signals in the X, Y, and Z directions, and the force signals in three directions, as well as the wavelet time-frequency diagrams, are displayed in Table 3. The signals in the three directions are stable and have a certain periodicity, but the wavelet time-frequency diagrams of the X and Y directions lack significant spectral characteristics, making it difficult to distinguish the different roughness levels. The Z direction in this experiment is the direction perpendicular to the workpiece milling surface, which is closely related to the depth of cut and has the largest influence on surface roughness. In summary, the Z direction force signal $F_z$ is chosen as the cutting force signal characterization for this experiment.
Table 3 Three directions of force signals and their corresponding time-frequency diagrams

<table>
<thead>
<tr>
<th>Image Type</th>
<th>X direction</th>
<th>Y direction</th>
<th>Z direction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Force signals</td>
<td><img src="image" alt="Diagram" /></td>
<td><img src="image" alt="Diagram" /></td>
<td><img src="image" alt="Diagram" /></td>
</tr>
<tr>
<td>Time-frequency diagrams</td>
<td><img src="image" alt="Diagram" /></td>
<td><img src="image" alt="Diagram" /></td>
<td><img src="image" alt="Diagram" /></td>
</tr>
</tbody>
</table>

3.3 Surface roughness grade classification

The arithmetic mean deviation $R_a$ of the profile is the most often used criterion for evaluating the surface roughness of a part[27], hence $R_a$ is utilized as an evaluation criterion for the surface roughness of the measured part in this experiment. The surface roughness formula is defined in (2):

$$R_a = \frac{1}{L} \int_{0}^{L} |z(x)| \, dx$$  \hspace{1cm} (2)

Where $L$ is the evaluation length and $z(x)$ is the profile height function.

The surface roughness range for this measurement is 0.2μm to 1.0μm, and it is classified into four classes, which are shown in Table 4 along with class division intervals and related surface quality.

Table 4 Surface roughness classification

<table>
<thead>
<tr>
<th>Classification</th>
<th>$R_a$ range(μm)</th>
<th>Surface quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class1</td>
<td>0.2–0.4</td>
<td>Fine</td>
</tr>
<tr>
<td>Class2</td>
<td>0.4–0.6</td>
<td>Smooth</td>
</tr>
<tr>
<td>Class3</td>
<td>0.6–0.8</td>
<td>Rough</td>
</tr>
<tr>
<td>Class4</td>
<td>0.8–1.0</td>
<td>Coarse</td>
</tr>
</tbody>
</table>

Fig. 4 depicts how the surface roughness $R_a$ of the milled workpiece changes as the total material removal increases during the experiment. As the experiment progresses, the tool wears more and more, and $R_a$ increases from 0.234μm to 0.885μm, which is consistent with the three stages of tool wear: initial, intermediate, and severe wear. This process yielded four types of workpiece surface roughness classes, and their related surface morphologies are illustrated in Table 5.
The surface microscopic microstructure of γ-TiAl alloy was imaged using a laser microscope at a magnification of 20 times. As shown in Table 5, the surface morphology and three-dimensional profiles of γ-TiAl alloy with different roughness under the laser microscope. The surface quality of different roughness can be seen to differ significantly; the lower the surface roughness, the smoother and higher the quality of the workpiece surface, and the larger the surface roughness, the more likely to generate processing defects such as grain pull-out and material spalling.
Table 5 Microscopic surface morphology and Three-dimensional profiles of γ-TiAl alloy at four surface roughness levels

<table>
<thead>
<tr>
<th>Surface roughness grade</th>
<th>Microscopic surface morphology</th>
<th>Three-dimensional profiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1:</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>$R_a=0.2\sim0.4$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 2:</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>$R_a=0.4\sim0.6$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 3:</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>$R_a=0.6\sim0.8$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 4:</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>$R_a=0.8\sim1.0$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.4 Data Enhancement

Although the proposed deep learning model is capable of learning, it requires a huge quantity of data for training, which is costly in terms of data acquisition. Hundreds of milling experiments were conducted separately under each parameter in the milling orthogonal experiments, and there were about hundreds of data for each roughness class, which did not reach the amount of data for training. To address this issue, data augmentation is required to process the existing data and increase the amount of data required for experimental samples to ensure the accuracy of data classification.

In computer vision, data enhancement is often performed by flipping, cropping, and distorting images. However, because these methods are inadequate for processing one-dimensional time-domain signals, the data in this part is augmented by the addition of a fixed-length sliding window data slice. The data length of the force signal obtained from each knife walk in the experiment is approximately 40,000 data points, and the length of the training samples collected each time is set to 1024, allowing for a large amount of data to be obtained under various working conditions, as shown in Fig. 5.

![Window sliding for data enhancement](image)

**Fig. 5** Window sliding for data enhancement

4 The proposed Method

Based on time-frequency diagrams and an upgraded ResNet paired with a transfer learning mechanism, a method for real-time monitoring of the surface roughness of γ-TiAl alloy is proposed. Fig. 6 depicts the basic procedure, which is divided into two parts: modeling and prediction. The experiments are carried out through a designed orthogonal test scheme, and the cutting force signals gathered are translated into wavelet time-frequency diagrams before being fed into a deep learning
model for training. The trained model can be called online during the machining process to take the
signal as input and automatically transform it into a time-frequency diagram for surface roughness
identification. This technology makes it possible to monitor surface roughness in real time and
adjust parameters in time to obtain better surface roughness of the workpiece during machining.

Fig. 6 Flow chart of the proposed method

4.1 Input dataset

Table 6 depicts the distribution of time-frequency diagrams of cutting force signals at four
different roughness levels. CWT maps the force signal features that can reflect surface roughness
information into a two-dimensional time-frequency diagram. The time-frequency diagrams of four
different levels of roughness signals reveal more obvious differences in their distribution
characteristics. The obvious frequency domain features are mainly reflected in the low-frequency
domain, when \( R_a \) is small, the spectral amplitude is low, and as \( R_a \) increases, the amplitude increases
accordingly. The time-frequency diagrams corresponding to the force signals of the four surface
roughness reveal significant variances in their distribution patterns in the time-frequency domain.
The low-energy band, for example, eventually expands to the high-frequency zone as the surface
roughness rises (the light blue area in the white box in the figure). Furthermore, the bigger \( R_a \), the
darker the red color (the red area in the dashed circle in the figure), suggesting that the energy is
concentrated in the frequency range of 0–1000 Hz. They can be used as input images to the model.
<table>
<thead>
<tr>
<th>Surface roughness grade</th>
<th>Cutting force signals</th>
<th>Time-frequency diagrams</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1: $R_a=0.2~0.4$</td>
<td><img src="image1" alt="Signal for Class 1" /></td>
<td><img src="image2" alt="TFD for Class 1" /></td>
</tr>
<tr>
<td>Class 2: $R_a=0.4~0.6$</td>
<td><img src="image3" alt="Signal for Class 2" /></td>
<td><img src="image4" alt="TFD for Class 2" /></td>
</tr>
<tr>
<td>Class 3: $R_a=0.6~0.8$</td>
<td><img src="image5" alt="Signal for Class 3" /></td>
<td><img src="image6" alt="TFD for Class 3" /></td>
</tr>
<tr>
<td>Class 4: $R_a=0.8~1.0$</td>
<td><img src="image7" alt="Signal for Class 4" /></td>
<td><img src="image8" alt="TFD for Class 4" /></td>
</tr>
</tbody>
</table>
4.2 Structure of the Proposed Model

The ResNet's structural foundation is the residual basic module, which employs shortcut connections to skip convolutional layers. This effectively solves the issue of gradient disappearance or gradient explosion brought on by extending the depth of the neural network and enables more flexible construction of CNN structures. Fig. 7 depicts the fundamental structure.

![Basic residual module](image)

**Fig. 7 Basic residual module**

The ResNet50 network model's building blocks are essentially residual block structures\(^{30}\), with a convolutional layer and several basic building blocks forming the entire residual network. ReLU activation functions are used at each level, and Adam optimizer is used to improve network recognition accuracy. The set of weight parameters obtained from ImageNet training is brought into its model network in conjunction with transfer learning. Table 4 shows the size parameters of each network block, as well as the two-dimensional size of each block's output, and a common structure of the residual block, is defined in (3):

\[
y = F(x, \{W_i\}) + x
\]  

(3)

Where and \( W_i \) is the weight in the weight matrix, \( F \) is the residual mapping to be learned, and \( x \) and \( y \) represent the input and output of the residual function \( F \).

The weight parameters of the pre-trained model were trained on the ImageNet dataset\(^{31}\), and in this study, the images to be analyzed are two-dimensional time-frequency diagrams, which are very different from those in ImageNet, so the last 3 layers of pre-trained ResNet50, namely the fully connected layer, the softmax layer, and the classification layer, were modified so that the modified ResNet50 can be applied to classify 4 classes of surface roughness. The modified ResNet50 was then retrained to generate new parameters, and Table 7 shows the corresponding structure of the
modified ResNet50. The structure of the proposed model is shown in Fig. 8. The two-dimensional features are tiled into a one-dimensional vector starting with a $7 \times 7$ pooling layer, and a dropout layer is added after the fully connected layer with a 50% dropout rate to reduce model overfitting. The final layer is a softmax function, which is used to calculate the likelihood of four final result classifications, which are fine, smooth, rough, and coarse.

<table>
<thead>
<tr>
<th>Layer name</th>
<th>Net</th>
<th>Output Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conv1</td>
<td>$7 \times 7, 64, \text{stride}2$</td>
<td>112 $\times$ 112</td>
</tr>
<tr>
<td></td>
<td>$3 \times 3 \text{max pool, stride}2$</td>
<td></td>
</tr>
<tr>
<td>Conv2_x</td>
<td>$[1 \times 1 \ 64]$</td>
<td>$56 \times 56$</td>
</tr>
<tr>
<td></td>
<td>$[3 \times 3 \ 64] \times 3$</td>
<td></td>
</tr>
<tr>
<td>Conv3_x</td>
<td>$[1 \times 1 \ 128]$</td>
<td>$28 \times 28$</td>
</tr>
<tr>
<td></td>
<td>$[3 \times 3 \ 128] \times 4$</td>
<td></td>
</tr>
<tr>
<td>Conv4_x</td>
<td>$[1 \times 1 \ 256]$</td>
<td>$14 \times 14$</td>
</tr>
<tr>
<td></td>
<td>$[3 \times 3 \ 256] \times 6$</td>
<td></td>
</tr>
<tr>
<td>Conv5_x</td>
<td>$[1 \times 1 \ 512]$</td>
<td>$7 \times 7$</td>
</tr>
<tr>
<td></td>
<td>$[3 \times 3 \ 512] \times 3$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$[1 \times 1 \ 2048]$</td>
<td></td>
</tr>
</tbody>
</table>

Average pool, 4-d fc, softmax, classification layer

5 Results and Analysis

5.1 Classifier performance evaluation

The performance evaluation metrics of the classifier are mainly Accuracy, Precision, Recall,
and F-Score\textsuperscript{[32]}, which are defined as follows:

\[
\text{Accuracy(\%)} = \frac{TP + TN}{TP + TN + FN + FP} \times 100
\]  
(4)

\[
\text{Precision(\%)} = \frac{TP}{TP + FP} \times 100
\]  
(5)

\[
\text{Recall(\%)} = \frac{TP}{TP + FN} \times 100
\]  
(6)

\[
F - \text{Score(\%)} = \left(2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}}\right) \times 100
\]  
(7)

Where \(TP\), \(TN\), \(FP\), and \(FN\) represent true positive, true negative, false positive, and false negative, respectively.

The Precision-Recall (P-R) curve describes the interconstraint between true positive rate and positive predictive value. \(P\) stands for Precision, and \(R\) stands for Recall, and neither can be maximized at the same time because one comes at the expense of the other. Class 4 performs poorly, as shown in Fig. 9(a), with the precision rate decreasing from the point where it exceeds recall 0.7, most likely due to a lack of training data. The other three classes produce excellent results even with recall close to one and accuracy greater than 0.8, indicating that the models are properly trained.

The Receiver Operator Characteristic (ROC) curves depict the model's performance at the classification level\textsuperscript{[33]}. The dashed line in Fig. 9(b) represents the ROC curve of a random classifier. A good classifier will be located near the upper left corner, away from the center dashed line. The ROC curve's overall trend corresponds to the precision and recall curves. The area under the curve (AUC) measures the performance of the classifier. The AUC of a perfect classifier is one. All four roughness classes have AUC values close to one, indicating that the model generalizes well. Due to the reasons stated above, the visual inspection results show that Class 4 has the smallest AUC.

Fig. 9(c) depicts the model's confusion matrix, with the most notable feature being that Class 4 has the worst prediction, with 7\% of Class 4 being predicted as Class 3. This is also consistent with the trend of the P-R curve and ROC curve performance, indicating that Class 4 data may have a low amount of data or an unclear classification level.
5.2 Network optimization results based on the proposed model

The accuracy of the proposed model on the test set is shown in Fig. 10. When compared to the basic ResNet50, the traditional EfficientNet, and the new Vision Transformer proposed in recent years, the improved model has significantly improved the classification accuracy of surface roughness, with a classification accuracy of 96.12%, demonstrating that the improved method effectively improves the model's performance.
5.3 Key factors affecting model performance

5.3.1 Effect of different residual structures

ResNet18, ResNet34, ResNet50, and ResNet101 are the most common residual structures, and the structures become more complex as the network levels are deepened in turn. As shown in Fig. 11 and Fig. 12, although the ResNet101 network is the deepest, it does not achieve better results in terms of accuracy for the not-so-large dataset in this paper, but it will have the problems of long training time and excessive memory usage. ResNet50 performs best in both loss rate and accuracy and has the highest prediction accuracy on the test set with a maximum of 98%. As a result, the ResNet50 residual structure is chosen as the best in this paper.
5.3.2 Effect of different Batch size

The effect of batch size on model performance is shown in Fig. 13. When the batch size is increased from 8 to 32, the classification accuracy rises; when it is increased further to 64, the classification accuracy falls slightly, but at this point, the training loss rate also drops, and the training pace quickens. As can be seen, selecting 64 for the batch size is preferable.
5.3.3 Effect of different learning rates

The learning rates (LR) range is set to [0.0001, 0.001], and the constructed model's performance is evaluated. Fig. 14 shows that for the same epoch, the training loss is greater and the LR is lower, and the training loss rate is greatest when the LR is 0.0001. As the LR increases, so does the training accuracy. When the epoch is greater than 40, the accuracy and loss remain constant. It can be seen from this that selecting 0.001 as the initial LR is more reliable.

Fig. 15 depicts the classification precision, F-score, and recall of the test samples with various LRs. It can be seen that with LR=0.01, the classification precision is the highest, up to 95.5%, and the F-score and recall are also the highest, 95.1% and 95%, respectively, confirming the correctness of selecting LR=0.01.
5.4 Further experimental validation of the model

In the previous sections, we validated the model using a three-factor, three-level orthogonal test and received good results. To further validate the model's generalization performance, the machining parameters are changed and the cutting force signals under random parameters are collected, and the collected force signals are converted into time-frequency diagrams as the model's input to predict its surface roughness, and the predicted results are compared with the real surface roughness values to judge the prediction's accuracy. The results of the real surface roughness values with the model prediction under numerous different cutting parameters are reported in Table 8, and the force signals with different parameters are shown in Fig. 16. Except for Class 3 and Class 4, which include one set of prediction mistakes each, all of the nine groups of randomly selected experiments are correct. The confusion matrix of the test results is shown in Fig. 17, and it can be observed that Class 1 has the highest prediction accuracy, while Class 3 and Class 4 are inferior, but all of them exceed 80%, indicating that the model has good generalization performance.

<table>
<thead>
<tr>
<th>No.</th>
<th>Spindle Speed (r/min)</th>
<th>Feed Rate (mm/r)</th>
<th>Depth of Cut (mm)</th>
<th>True Value (μm)</th>
<th>Prediction</th>
<th>True Or False</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2500</td>
<td>0.015</td>
<td>0.2</td>
<td>0.307</td>
<td>Class 1</td>
<td>TRUE</td>
</tr>
<tr>
<td>2</td>
<td>3000</td>
<td>0.0075</td>
<td>0.2</td>
<td>0.809</td>
<td>Class 3</td>
<td>FALSE</td>
</tr>
<tr>
<td>3</td>
<td>3000</td>
<td>0.010</td>
<td>0.15</td>
<td>0.316</td>
<td>Class 1</td>
<td>TRUE</td>
</tr>
<tr>
<td>4</td>
<td>3000</td>
<td>0.010</td>
<td>0.15</td>
<td>0.356</td>
<td>Class 1</td>
<td>TRUE</td>
</tr>
<tr>
<td>5</td>
<td>3000</td>
<td>0.020</td>
<td>0.2</td>
<td>0.599</td>
<td>Class 2</td>
<td>TRUE</td>
</tr>
<tr>
<td>6</td>
<td>3000</td>
<td>0.025</td>
<td>0.2</td>
<td>0.722</td>
<td>Class 4</td>
<td>FALSE</td>
</tr>
<tr>
<td>7</td>
<td>3500</td>
<td>0.02</td>
<td>0.15</td>
<td>0.326</td>
<td>Class 1</td>
<td>TRUE</td>
</tr>
<tr>
<td>8</td>
<td>4000</td>
<td>0.025</td>
<td>0.2</td>
<td>0.837</td>
<td>Class 4</td>
<td>TRUE</td>
</tr>
<tr>
<td>9</td>
<td>4000</td>
<td>0.025</td>
<td>0.2</td>
<td>0.548</td>
<td>Class 2</td>
<td>TRUE</td>
</tr>
</tbody>
</table>
Fig. 16 Cutting force signals at nine sets of verification parameters

Fig. 17 Confusion matrix for surface roughness classification under random parameter
6 Conclusion

Surface roughness is a significant indicator to monitor the surface quality of workpieces in machining, and research on the real-time monitoring method of surface roughness boosts machining efficiency and reduces machining costs. In this research, we offer a real-time surface roughness monitoring method based on time-frequency diagrams and a transfer learning model that can successfully determine whether the roughness of the machined surface matches the machining requirements, and we reach the following conclusions:

(1) Converting the one-dimensional force signal into a two-dimensional time-frequency diagram can provide additional information about surface roughness and is also a data enhancement method. The differences in the spectrum features of the time-frequency diagrams can be utilized to more accurately discern the characteristics of varied surface roughness in machining.

(2) A new method of intelligent monitoring of surface roughness based on ResNet with transfer learning is proposed by combining continuous wavelet transform with deep learning theory. This method does not require manual feature extraction, overcomes the information loss problem of traditional methods, and can obtain more comprehensive surface roughness information.

(3) The proposed method has fast convergence and high accuracy for all four classes of surface roughness. For this model, the highest accuracy is 98% for the Class 1 roughness and the lowest is 90% for Class 4, and the average test accuracy is improved by about 6% from 89.93% to 96.12% compared with the traditional neural network. During the model training process, the generalization ability of the model is improved by introducing batch normalization and Dropout, which can improve the model's accuracy by about 2%.

(4) According to the findings of this investigation, the force signal can be saved in real-time in the database, and the surface roughness value of the workpiece during machining can be monitored online by invoking the deep learning model. According to Section IV, generating a time-frequency diagram only takes about 0.1s of processing time, i.e., it is possible to determine whether the machining accuracy meets the requirements in a very short time, so that the tool can be replaced or the cutting parameters can be adjusted to reduce the surface roughness in time, and the surface quality of the workpiece during machining can be ensured to the greatest extent possible.
References


Funding This work was supported by the National Natural Science Foundation of China (52175416, 51775280).

Competing Interests The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author contribution All authors contributed to this work.