Proceedings of the ASME 2022
International Design Engineering Technical Conferences and
Computers and Information in Engineering Conference
IDETC/CIE2022
August 14-17, 2022, St. Louis, Missouri

DETC2022-XXXX

A Material Removal Prediction Method Based On Multi-Scale Attention Mechanism

Zhihang Li[1] **Qian Tang[1]** **Ying Liu[2]** **Xiao Li[1]**

1. State Key Laboratory of Mechanical Transmissions
Chongqing University
Chongqing, 400044, China

2. Institute of Mechanical and Manufacturing Engineering
School of Engineering
Cardiff University
Cardiff, CF24 3AA, UK

ABSTRACT

The exact removal of material in abrasive belt grinding determines the final machining quality of the workpiece. However, it is difficult to determine the removal state of materials in actual processing, which is affected by factors such as abrasive belt wear and processing errors. Therefore, a multi-scale attention convolutional neural network for material removal state prediction method is proposed based on the analysis of displacement data. First, the first-order difference and sliding-window expansion methods for displacement data are adopted, making it possible to use displacement data for deep learning, which is the premise of material removal state prediction. Then, the multi-scale convolutional neural network is Employed to extract important features of the displacement data. Due to the different importance of different features, Squeeze-and-Excitation Networks is used to independently assign the importance of features based on the loss function, so that the model pays more attention to those main features and ignores the secondary features, which can improve the convergence speed and prediction accuracy of the model. The K6 cross-validation of experiment results shows that this method can accurately predict the material removal state with an average prediction accuracy of 87.9%, which can be practically applied to the online prediction of the material removal state in industrial processing to further control the processing quality.

Keywords: material removal, displacement data, multi-scale feature extraction, attention

NOMENCLATURE

MSACNN	multi-scale attention convolutional neural network
1D CNN	one-dimensional convolutional neural networks
SENet	squeeze-and-excitation networks
PReLU	parametric rectified linear unit
BN	batch normalization
MSFE	multi-scale feature extraction
FU	fully connected

1. INTRODUCTION

Different from rigid contact in wheel grinding, abrasive belt grinding is widely used for material removal of complex curved surfaces due to the advantages of elastic contact between the abrasive belt and the workpiece [1]-[3]. Due to the random distribution of the grain size, height, and direction of the abrasive belt cutting edge, abrasive belt grinding is a grinding method with an indeterminate amount of removal [4]. In addition, due to the influence of abrasive belt wear and errors, it is difficult to achieve ideal results of material removal in actual grinding, so the material removal detection of workpieces is particularly important [5][6].

The detection of material removal state is usually divided into direct measurement and indirect measurement. The main methods of direct measurement are three-coordinate measuring machines and surface profiler, but all have cumbersome measurement procedures, which greatly increases the time cost and affects the processing efficiency[7][8].

Indirect measurement can be divided into mechanism model prediction and learning model prediction. It is difficult to predict the material removal state by establishing a data model or a simulation model [9][10]. Artificial intelligence algorithms can establish complex nonlinear mapping relationships between sensor signals and targets, so it can be used for target prediction under complex conditions in the industry [11]. Benefiting from the rapid development of big data intelligent algorithms, indirect prediction methods for material removal prediction based on analysis of collected signal data have been extensively studied.

Ren et al. proposed a method for monitoring the material removal rate of belt grinding by spark field measurement, the correct rate of grinding depth recognition can reach 95%, which is an effective method for monitoring the material removal rate of belt grinding [12]. Zhong et al. Zhong et al. proposed a back-propagation artificial neural network. This method improves the accuracy of predicting the material removal rate of ultrasonic machining [13]. Shen et al. applied Support Vector Fuzzy Adaptive Network as a parameter-free nonlinear regression technique to model material removal rate. The algorithm retains the advantages of fuzzy adaptive network and support vector machine, and it is a more effective modeling algorithm for complex manufacturing processes [14]. A material removal model for Inconel 718 robotic belt grinding based on acoustic sensing and machine learning is proposed [15]. A method utilizing discrete wavelet decomposition and fast Fourier transform is used to identify system idle operating cycles and eliminate noise. The results show that the method based on acoustic signal and ensemble learning model can effectively predict the material removal rate in the complex grinding environment. To obtain practical material removal, Wang et al. proposed a new method for monitoring material removal using multi-sensors and a two-dimensional convolutional neural network learning algorithm [16]. To obtain an accurate force control model and realize the uniformity of material removal, a new method was proposed to monitor the material removal rate and corresponding wear state of the abrasive belt online only by using the grinding sound signal [17]. This method can provide a good basis for the monitoring of material removal rate and abrasive belt wear during abrasive belt grinding. These methods all use deep learning models to establish a nonlinear relationship between signal data and material removal, to predict material removal. The superiority of deep learning for predicting material removal is proved.

Although the above research has achieved certain results, the collected signals are susceptible to interference and the cost is high, which limits its industrial application. Displacement sensors have the advantages of easy layout and low cost. Therefore, this paper uses the displacement data collected by the eddy current displacement sensor to predict the removal state of the workpiece. Based on the analysis of displacement data, a multi-scale attention convolutional neural network (MACNN) is proposed for the prediction of workpiece removal status. This method requires relatively little data volume and can achieve accurate prediction of material removal status. The main contributions of this paper are as follows:

(1) The method of predicting the state of material removal by displacement data is proven feasible. Not only the feasibility of predicting material removal using displacement data is demonstrated but also the feasibility of data augmentation. The first-order difference and overlapping sliding window enhancement processing are performed on the displacement data, which increases the amount of data and solves the contradiction between the difficulty of collecting a large amount of data and the large demand for model data.

(2) The combination of multi-scale convolutional neural network and attention mechanism achieves accurate prediction of material removal state. Using parallel convolution kernels of different sizes can extract the feature data of different frequency bands of the data, and combine the attention network to assign different weights to the features to achieve accurate extraction of important features and improve the prediction accuracy of the model.

The rest chapters of this paper are organized as follows. The second section introduces the use of related theories in this paper, including one-dimensional convolutional neural networks (1D-CNN) and Squeeze-and-Excitation Networks (SENet). Section 3 introduces the data preprocessing method. It mainly includes first-order difference, mean normalization processing, and data enhancement methods, which are also the premise for the deep learning model to achieve good performance. Section 4 is experiment and analysis, which introduces the built experimental platform and the prediction results of the proposed prediction method. Section 5 is the conclusion, which illustrates the achievements of this paper.

2. RELATED THEORY APPLICATIONS

For the feature extraction target of displacement data, this paper adopts 1D CNN, attention network Senet, and fully connected neural network respectively. 1D CNN is mainly used for noise reduction and feature extraction, and SENet is used to assign weights to the extracted multi-scale features so that the model pays more attention to important features and ignores secondary features. The fully connected neural network mainly classifies the extracted features and realizes the judgment of whether the material removal is qualified.

2.1 1D CNN

CNN is widely used in fault diagnosis, tool wear, and other applications, and has achieved good results [18][19]. 1D CNN is mainly used to process one-dimensional signal data, and its convolution kernel and pooling kernel are both one-dimensional. It has the same structure and properties as CNN.

CNN has excellent feature extraction ability and can extract key features in signal data. The CNN network includes convolutional layers and pooling layers. CNN networks can be freely combined to build models with better performance [20].

The convolution layer is the core layer of CNN, which mainly performs convolution operations on the data. The sliding window convolution operation is performed on the data through convolution checks of different lengths to extract key features in the data. Unlike the fully connected neural network, the neurons in the convolution operation have the characteristics of parameter sharing, which effectively reduces the complexity of the network on the premise of ensuring feature extraction. The calculation announcement is as follows.

$$y_l = k_l * y_{l-1} + b_l \tag{1}$$

$$y_l' = f(y_l) = \begin{cases} y_l & if\ y_l > 0 \\ ay_l & if\ y_l \le 0 \end{cases} \tag{2}$$

Where y_{l-1} is the output of the *l-1*-th layer, which is the input of the *l*th convolutional layer. y_l is the result of the convolution calculation, k_l is the size of the convolution kernel, and b_l is the bias. *f* is the parametric rectified linear unit (PReLU) activation function, and *a* is a fixed value.

Compared with the rectified linear unit (ReLU), the PReLU improves model fitting ability by considering negative values, at the same time, almost no additional computational cost is added, and the risk of overfitting is smaller [21].

It is worth noting that BatchNorm (BN) processing is required after the convolution operation, that is, before the activation function. The role of the BN layer is to keep the input of each layer of the network, that is, y_l in Equation 1, to maintain the same distribution. If BN is not used, when training data with different distributions in each batch, the network must adapt to different distributions, which will greatly increase the learning difficulty of the model and reduce the network training speed. The role of BN is equivalent to whitening in image processing, that is, to transform the input data distribution to a normal distribution with 0 mean and unit variance [22]. BN has the advantages of improving model training speed and preventing overfitting, and it is widely used in intelligent network models [23].

The pooling layer is connected to the PReLU layer, and by down-sampling the extracted features, the computational complexity is greatly reduced on the premise of ensuring that no features are lost. The pooling layer does not change the number of channels in the network, but only reduces the initial characteristics of the data. The pooling layer is not a necessary network layer. When the number of data features is significantly small, the pooling layer is usually chosen to be canceled. It is divided into average pooling and max pooling. Its calculation formula is as follows:

$$y_{l+1}^{P_a} = \frac{1}{k}\sum_{i=1}^{n} y_{l,i}^{'} \tag{3}$$

$$y_{l+1}^{P_m} = max\{y_{l,1}^{'}, y_{l,2}^{'} \ldots y_{l,k}^{'}\} \tag{4}$$

Where *l+1* represents the *l+1*-th layer. $y_{l+1}^{P_a}$ and $y_{l+1}^{P_m}$ are the outputs of average pooling and max pooling, respectively. *k* is the kernel of the pooling layer.

The application of CNN in this paper includes denoising and multi-scale feature extraction.

2.1.1 DENOISING LAYER

Because the actual processing situation is complicated, the collected signals often inevitably contain noise. Combined with the excellent feature extraction ability of convolution, the use of a wide-kernel convolution kernel can effectively denoise the data [24]. The average pooling process can effectively smooth and denoise the data, reduce the influence of noise. The parameters used in this paper are that the convolution kernel is 32, the stride is 9, and the pooling kernel is 4. The output channel of the denoising layer is 8. Its structure is shown in FIGURE 1. The preprocessed data is processed by the denoising layer, which can not only reduce the interference factors of the data but also map the data to a multi-dimensional feature space, which is convenient for the subsequent network to further extract features.

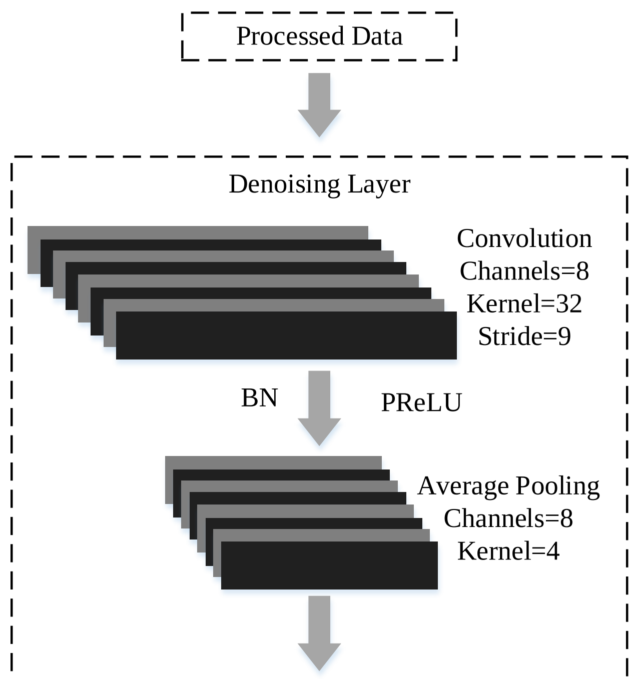

FIGURE 1: THE FRAMEWORK OF DENOISING LAER

2.1.2 MULTI-SCALE FEATURE EXTRACTION LAYER

The multi-scale feature extraction (MSFE) layer consists of one layer of 1D CNN and three parallel 1D CNNs. Due to the complexity of the collected data signals and the use of sliding window data enhancement processing in this paper, simple and fixed single-scale convolution processing cannot extract diverse features. This paper adopts three multi-scale parallel convolutional networks. A variety of effective feature layers can be extracted. The convolution kernels are 4, 6, and 12 respectively. Each convolutional layer is followed by BN, PReLU, and max pooling. where the pooling kernel is 2. Its network structure is shown in FIGURE 2.

First, a convolutional layer with a channel of 4 is used to reduce the number of channels in the noise reduction layer, and further filter the features to remove the initially unimportant common features. There is no pooling layer here. Then three parallel convolution layers with different convolution kernels can filter the features at different scales and extract features in different frequency bands. Finally, three parallel pooling layers are used to highlight the extracted multi-scale features. Similarly, the BN layer and PReLU are sequentially connected after each convolutional layer.

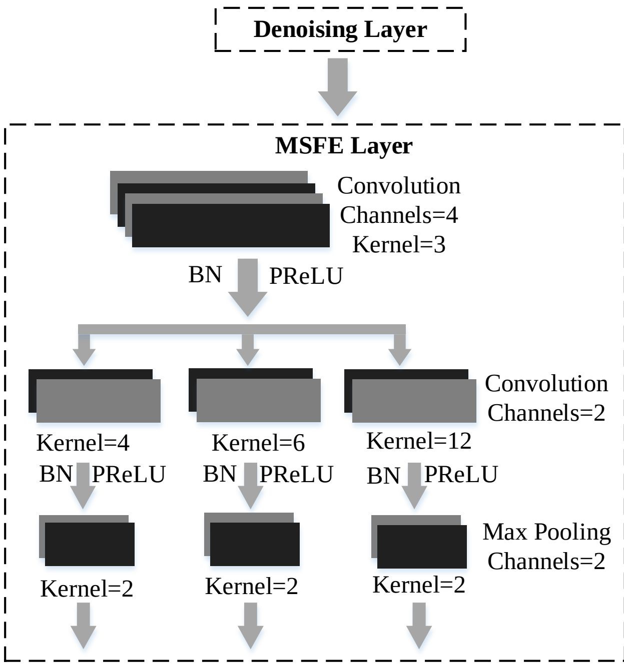

FIGURE 2: THE FRAMEWORK OF THE MSFE LAYER

2.2 ATTENTION LAYER

Although the multi-scale network can extract features at different levels, it also brings the diversity and complexity of features. The importance of these features is different, so these extracted features need to be screened. The attention layer can help the network pay attention to those important features. This is similar to when people identify the content of a painting, they will focus on representative features and ignore those common secondary features. Thus, this paper builds an attention network layer to identify those important features. The attention layer consists of a one-layer output 6-channel 1D CNN and a SENet. Because the extracted features are more complex, a layer of the convolutional neural network is used for initial attention extraction, and then the SENet network is used for final attention. The framework of the attention layer is shown in FIGURE 3. Unlike the previous convolutional layers, both the input and output channels of the convolutional layer in the attention layer are 6. The convolution does not change the channels of the original multi-scale layer. This is because the convolution operation here is only to integrate the original multi-scale features. The same is that the convolution is followed by BN, ReLU, and pooling. The decisive role in the attention layer is SENet.

Unlike CNN, SENet is a functional sub-network that can be combined with other networks to achieve better results. SENet was first proposed by Hu et al, and it has been used in combination with various networks and achieved good results [25]. The main purpose of SENet is to learn the weights under different features and assign different importance, which is automatically realized according to the loss value. The application of SENet is also very simple and flexible, and it can be directly applied to various existing network structures. In this paper, SENet is used in conjunction with CNN, applied to the multi-scale feature extraction layer, and it is followed by a classification fully connected (FC) layer. SENet itself also includes a two-layer FU neural network. First, perform the global average pool operation on each pass of the input c channels to obtain c scalars, and obtain c scalars of 0-1 through FC-ReLU-FC-Sigmoid, which is the weighted value of the channel. It is then weighted corresponding to the original input channel, and the feature importance of different channels is learned. The action principle of Senet is shown in FIGURE 4. Where y represents the features extracted by the multi-scale layers integrated by the convolutional neural network. $\tilde{y}$ represents the weighted result of y and the importance calculated by SENet. It is the final feature value extracted, and it is also the result of the combined action of the denoising layer, the MFFE layer, and the attention layer. It is the goal of these networks.

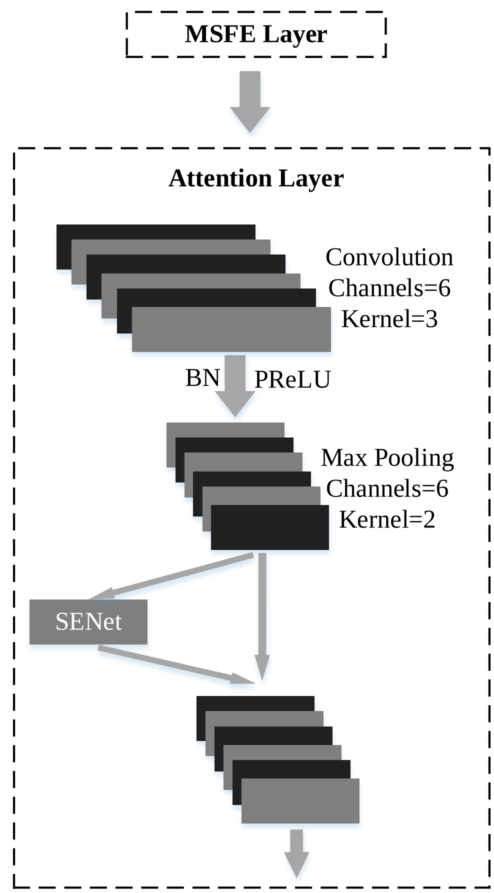

FIGURE 3: THE FRAMEWORK OF ATTENTION LAYER

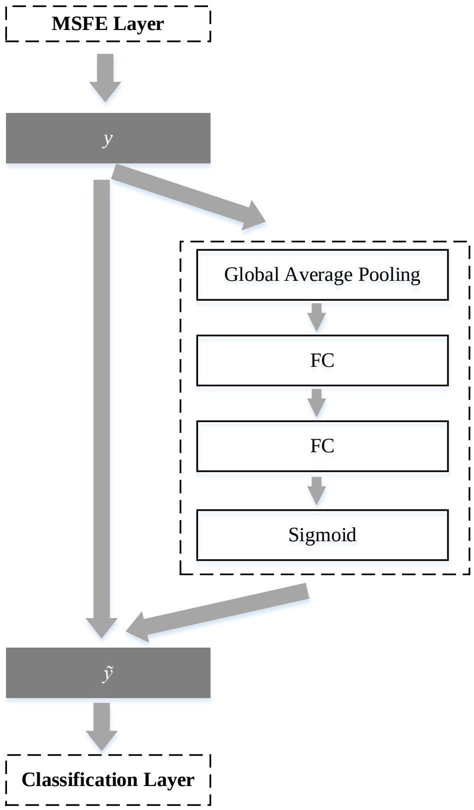

FIGURE 4: THE PRINCIPLE OF SENet

2.3 CLASSIFICATION LAYER

The classification layer is actually a layer of FC neural network. The main function of the above layers is to extract the corresponding features of the target, but the classification based on these features is the function of the classification layer [26]. In this paper, the material removal status is divided into two categories: qualified and unqualified. A single input of 800 data points is extracted to 24 key features through the denoising layer, multi-scale layer, and attention layer. And the purpose of this paper is to predict two states of material removal. Therefore, a FC neural network with 24 input neurons and 2 output neurons is used to output the final goal of this paper. The structure of the classification layer is shown in FIGURE 5. To ensure that the data output features of each batch obey the same distribution, similarly, the feature data is first processed by BN before connecting to the FC neural network.

3. DATA PREPROCESSING

The original data usually has interference factors such as noise and errors, so it is necessary to preprocess the data first. In this paper, the material removal criteria for different positions of the workpiece are the same, and the position of the displacement sensor is fixed, resulting in different displacements at different positions. The displacement data used in this paper not only contains interference factors but also error factors. Direct analysis of the displacement data is bound to fail to obtain ideal results. This is because the material removal criterion is the same for different positions of the workpiece, and the position of the displacement sensor is fixed. Thus, the displacement data of different positions of the workpiece is different even under the same material removal state.

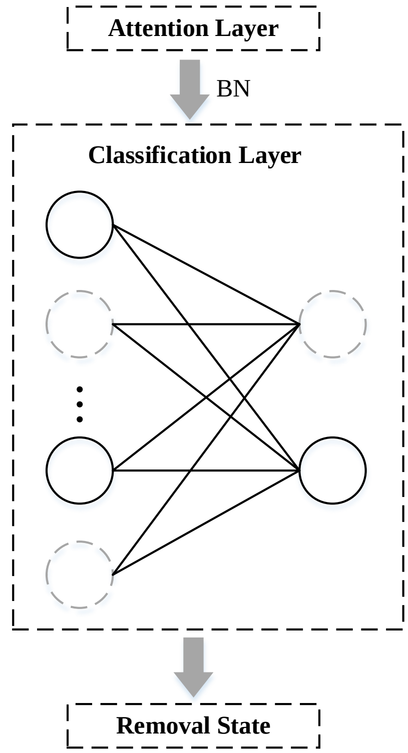

FIGURE 5: THE FRAMEWORK OF THE CLASSIFICATION LAYER

To eliminate this error, the change data of displacement is used as the analysis object, because although the displacement data of different positions are different, the changing trend of the displacement is similar. In this paper, first-order difference processing is performed on the displacement data, and the change of displacement is taken as the object of analysis. The first-order difference is to subtract the previous data from the next data of the time series displacement data in turn to form new displacement change data. Its calculation formula is as follows:

$$\begin{cases} y_i = y'_{i+1} - y'_i \quad i = 1,2,3\ldots n \\ y = [y_1, y_2, y_3 \ldots y_n] \end{cases} n = len(y') - 1 \tag{5}$$

where y'_i represents the i-th data of the original data y'. y is the displacement change data, n represents the total number of data of y.

It is well known that the performance of deep learning networks is proportional to the size of the dataset [27]. In actual processing, it is time-consuming and labor-intensive to collect a large amount of processing data, which is difficult to achieve. To solve this contradiction, this paper uses sampling and sliding window enhancement to expand the dataset, which is often used in the fault diagnosis of rotating machinery [28]. In actual processing, the rotation of the abrasive belt is periodic, which is also the reason for using data expansion. Different amounts of sampled data were compared experimentally. The experimental results show that when a single sample has 800 data points and the number of overlapping data is 20, a good prediction effect can be obtained. The data augmentation method is shown in FIGURE 6. With this data augmentation method, the source data set is augmented from 168 to 29520. Considering that the acquisition time of displacement data during actual processing is longer than the actual processing time, it is necessary to intercept the data first. In this paper, 4901 data from 100 to 5000 were extracted.

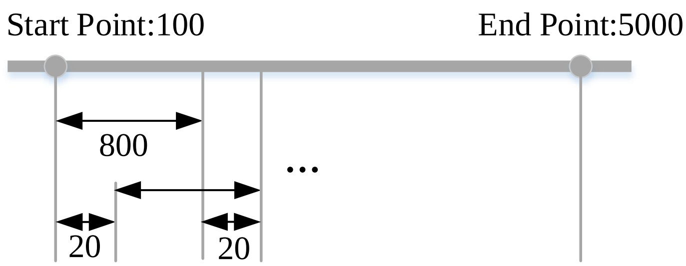

FIGURE 6: DATA AUGMENTATION METHODS

Differential preprocessing and denoising layer also have limitations on noise control. If the differential data is directly analyzed, the local noise features contained in the data will also be concerned by the model, which will cause interference. To prevent the local features of the extracted sample data from being too obvious, mean normalization is performed on the data before data enhancement. And normalization processing can not only speed up the convergence of the model but also improve the model accuracy. Its calculation formula is as follows:

$$\begin{cases} x_i = \frac{y_i - \mu}{\sigma} \\ x = [x_1, x_2 \dots x_n] \end{cases} \tag{6}$$

Where μ and σ are the mean and standard deviation of the data y, respectively. x is the normalized data and is the final data used in this paper to predict the material removal status

The part of data after difference and normalization successively are shown in FIGURE 7. It can be seen that the data before and after the difference are quite different. It also illustrates the necessity of differential processing. Although the data trends before and after normalization in FIGURE 7 do not change, the value of the data has changed. The minimum and maximum values of unnormalized data are -0.01868 and 0.015503, respectively. The minimum and maximum values of normalized data are -1.78883 and 1.485467, respectively. That is, the value of the data has changed by normalization. Although this change is not visible, these changes can be recognized by the model and have a huge impact on the convergence effect of the model. It helps to improve model convergence and generalization when experimentally validated.

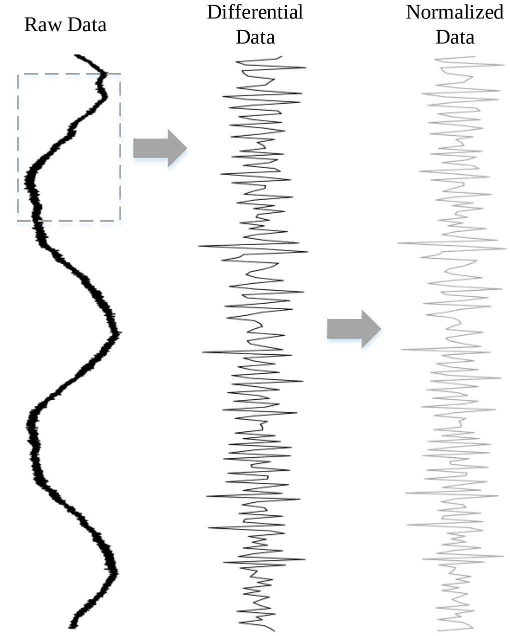

FIGURE 7: DATA PREPROCESSING EXAMPLE

4. EXPERIMENT AND VALIDATION

In this section, the prediction method MSACNN proposed in this paper is used to predict the material removal state, and the experimental results show the feasibility of using displacement data for prediction. The model training hardware is a high-performance computer with an NVIDIA GeForce 3060 graphics processing unit.

4.1 DATA COLLECTION

An experimental platform was built using NI-CRIO data acquisition system and 2M554D CNC abrasive belt grinder, as shown in FIGURE 8. The 2M554D CNC abrasive belt grinder is used to grind the circular workpiece, and the eddy current displacement sensor is responsible for collecting the displacement data of the workpiece during the machining process.

FIGURE 8: EXPERIMENT PLATFORM

The experimental processing position and the position of the displacement sensor are shown in FIGURE 9: SCHEMATIC DIAGRAM OF THE EXPERIMENTAL LAYOUT It can be seen that the abrasive belt is directly above the workpiece, and the displacement sensor is directly below the workpiece.

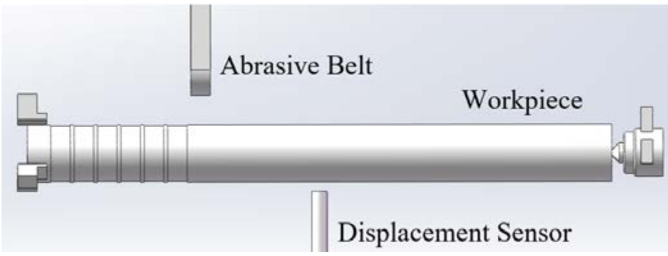

FIGURE 9: SCHEMATIC DIAGRAM OF THE EXPERIMENTAL LAYOUT

The experimental processing object is 20CrMnTi alloy steel with a length of 30cm and a diameter of 24mm. A displacement sensor is used to collect the displacement change data of the workpiece during the processing. To increase the amount of machining data for a single workpiece and improve the practicality of the experiment, the workpiece was repeatedly machined 11 times with an interval of 20 mm. Since this paper explores the relationship between workpiece removal rate and displacement data, an alumina abrasive belt with a larger removal rate is used, and its properties are shown in TABLE 1.

TABLE 1: BELT PROPERTIES

Belt model	Sharpen	granularity	width
VSM-XK870F	Alumina oxide	120#	10 mm

In this paper, the removal state of the material is studied, and the workpiece is a round bar, so Δd is used to represent. The material removal calculation formula is as follows.

$$\Delta d = d_1 - d_2$$

Where d_1 and d_2 are the data before and after processing, respectively.

In this paper, the workpiece with Δd less than 0.5mm is regarded as unqualified. Based on this, the collected displacement data samples can be divided into 70 qualified and 89 unqualified, a total of 159 pieces of data. 159 sets of data are expanded to 29520 sets of data by the expansion method. Due to the randomness of the data division, the prediction results of one time are not representative. The traditional way of dividing the data into training datasets and test datasets according to a certain proportion cannot accurately reflect the performance of the model. K-fold cross-validation is a way of dividing data sets. Different from traditional data division, it can avoid the limitations and particularities of fixed division data sets, and is widely used in modeling applications [29]. K-fold cross-validation means that the data is divided into K sub-samples, K-1 samples are used as the training dataset, and another sample is used as the test dataset.

In order to further prove the reliability of the method, the K6 cross-validation method is used. The dataset is divided into 6 parts K1, K2, K3, K4, K5, K6. The division of the dataset is shown in FIGURE 10. The average of the six prediction results is used as the final prediction result.

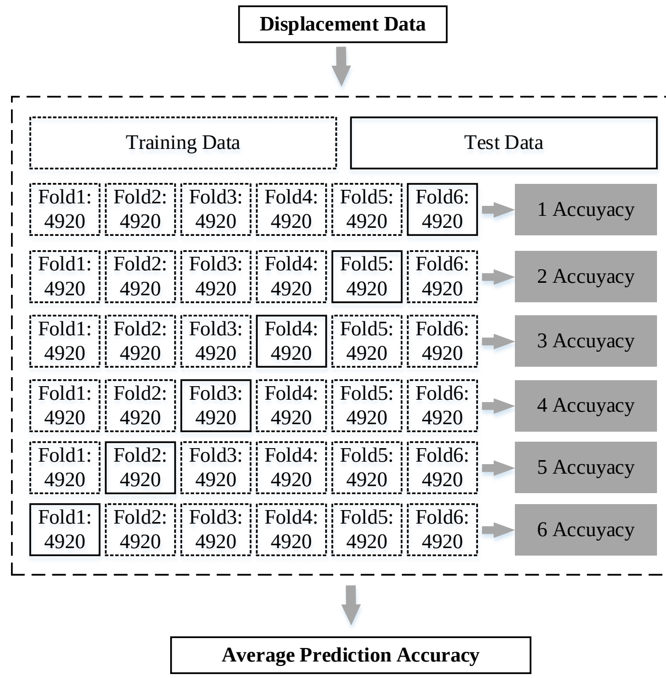

FIGURE 10: DATA DIVISION ACCORDING TO 6-FOLD CROSS VALIDATION

4.2 METHOD VALIDATION

The material removal state prediction method MSACNN proposed in this paper consists of five parts. The first part is the data preprocessing part. Through the first-order difference, mean normalization, and data enhancement processing of the data, it is the premise that the data can be used for deep learning prediction. The second part is the noise reduction part, which uses a wide-kernel convolutional neural network to perform noise reduction processing on the data. The third part is the multi-scale feature extraction part, which is mainly composed of three parallel convolutional neural networks with different convolution kernels, which can propose the features of different frequency bands of the data. The fourth part is the attention mechanism part, which is mainly composed of a convolutional neural network and senet. A convolutional neural network pre-extracts multi-scale features, and then SENet assigns the importance of the extracted features to help the model pay more attention. important features. The fifth part is the classification part. There is a single-layer neural network to classify the extracted features and realize the prediction and judgment of the removal state. It is the joint action of these five parts that ensures the predictive ability of the model.

The hyperparameter learning rate and batch size of the model are 0.0001 and 768, respectively. And Adam optimizer is used to optimize training[] . The datasets were trained 100 times separately under k6 cross-validation. The loss function for model training is shown in FIGURE 11. The loss function curve is only to verify whether the model converges, so this paper only draws the loss function curve for one verification, instead of drawing all the six loss function curves. It can be seen from the figure that the loss function converges to around 0.21, and does not converge to 0. As can be seen from the figure, although the loss function does not converge to 0, it converges to around 0.21. This not only proves that the model has learned specific features but also avoids overfitting the model to the training dataset. This is provable in the accuracy of the test dataset.

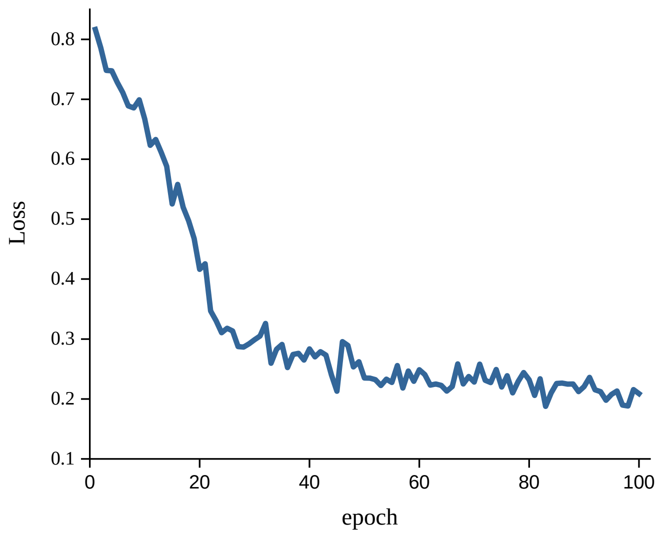

FIGURE 11: THE LOSS CURVES OF THE MSACNN

It can be seen from the loss function curve of the model that the prediction model has fitted the nonlinear relationship between the displacement data and the material removal state. Therefore, the learned model is used for validation on the test dataset. To eliminate the chance of prediction results and validate the robustness of the model, prediction validation was performed on the K1, K2, K3, K4, K5, and K6 datasets of K6 cross-validation respectively. The predicted curves of K1 to K6 are shown in FIGURE 12**:** It can be seen that although the final prediction accuracy is different, the accuracy curve from k1 to k6 converges well. There are slight fluctuations in the prediction accuracy of different datasets. The main reason is the problem with the dataset. Due to the dataset being small, it cannot be directly used for the establishment of deep learning models. Therefore, this paper expands the dataset. This results in a certain imbalance in the dataset. However, the overall prediction accuracy of the model is still very well. The average of the 6 predictions is used as the final material removal state prediction accuracy.

The average precision curve is shown in FIGURE 13. In this paper, the average prediction curve is taken as the final model performance curve. It can be seen that the average prediction accuracy curve is very smooth, indicating that the model has good prediction performance and can accurately identify the features of different states of material removal.

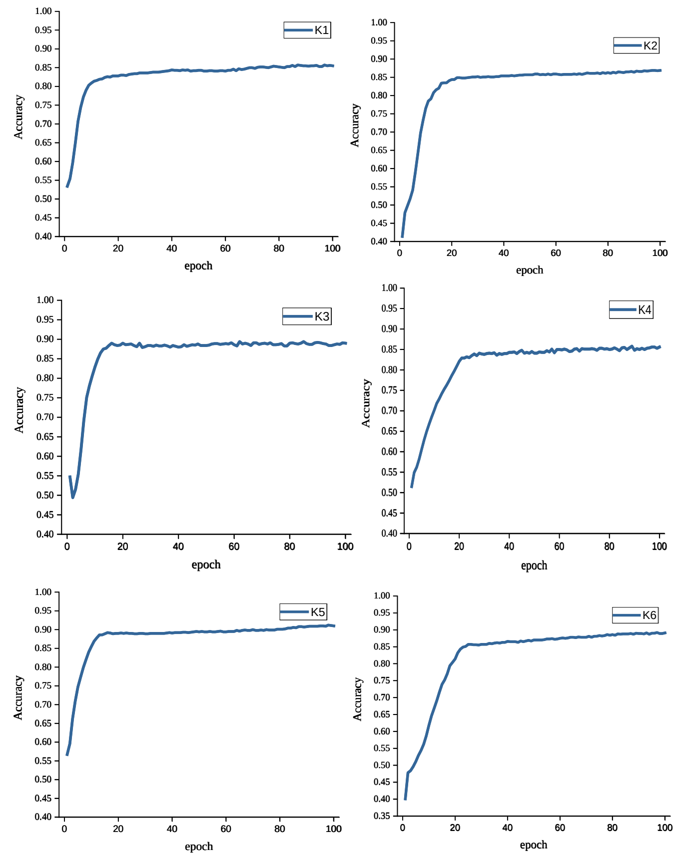

FIGURE 12: PREDICTION CURVES FROM K1 TO K6

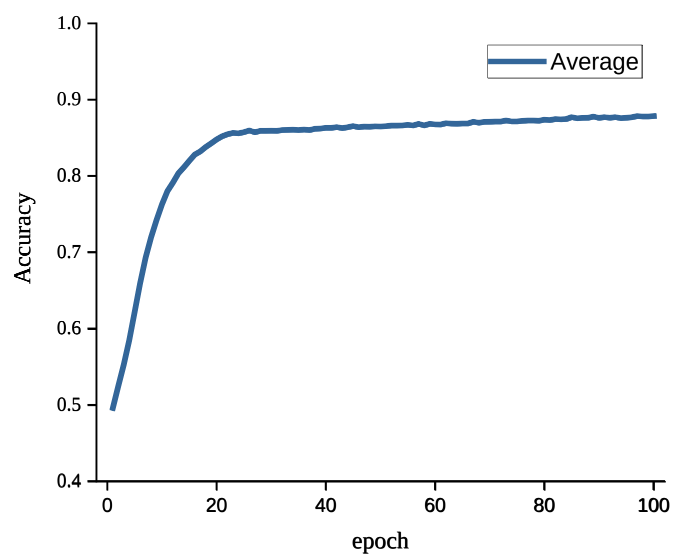

FIGURE 13: THE ACCURACY CURVE OF THE MSACNN

The prediction accuracies of k1 to k6 are shown in TABLE 2. The prediction accuracy of the material removal state is a minimum of 85.5% and a maximum of 91.0%. The average prediction accuracy of material removal status is 87.9%. It can be seen that the model prediction precision of K1, K2, and K4 data sets is basically the same, while the model prediction precision of K2, K5, and K6 data sets has little difference. That is to say, the prediction accuracy of models under different data sets is different, which means that the single accidental prediction accuracy cannot be used as the final prediction result. It proves the necessity and rationality of k6 cross validation in this paper.

TABLE 2: PREDICTION ACCURACY OF EACH PART OF THE DATASETS

Datasets	Accuracy
k1	85.5%
k2	86.9%
k3	89.0%
k4	85.6%
k5	91.0%
k6	89.1%
Average	87.9%

The average prediction results curve of the 6 predictions is shown in FIGURE 14.

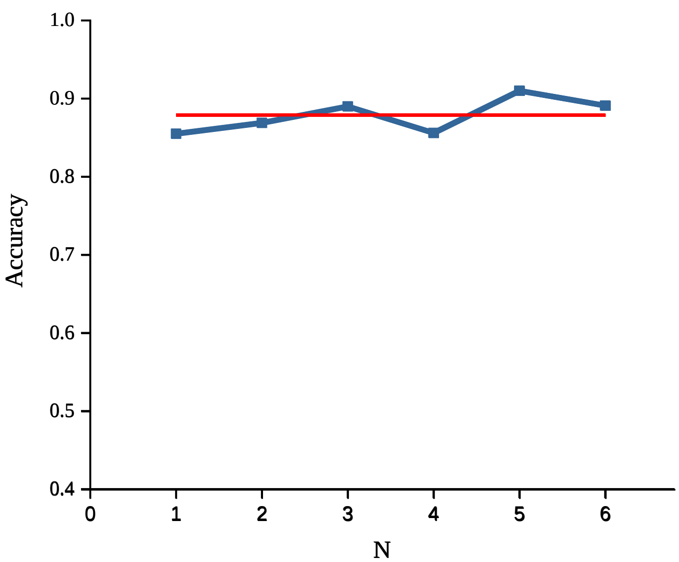

FIGURE 14: THE AVERAGE PREDICTION RESULTS

It can be seen that although the prediction result of the removal rate prediction method proposed in this paper has fluctuations, the overall fluctuation of the curve is not obvious and can be accepted. There should be two reasons for the fluctuation. On the one hand, interference errors are inevitably introduced when data expansion is adopted. For example, there are differences between single samples in the same removal state. On the other hand, the material removal state is defined according to the diameter difference of the workpiece. Human error is inevitable in measurement. For example, because the workpiece surface is not smooth after grinding, a slight deviation of the measuring point will lead to the difference in measured diameter. These factors are hard to avoid entirely. Therefore, the prediction method proposed in this paper can solve the contradiction between the small amount of data and the complex model parameters, and it can predict the removal rate to a certain extent, which can be applied to the actual processing prediction, and effectively reduce the defective rate of products.

5. CONCLUSION

This paper proposes a new method to remove state prediction, named MACNN. The experimental results show that the average prediction accuracy of the prediction method is 87.9%, it can be used to determine the material removal state during processing, which is beneficial to the online monitoring of the processing quality.

In this paper, using displacement data to predict material removal status proved feasible. The difference and sliding window processing of displacement data solves the contradiction between the large amount of data required by the deep learning model and the difficulty of collecting a large amount of data. Then, the combination of multi-scale CNN and SENet improves the attention of important features, which improves the prediction accuracy of the model.

ACKNOWLEDGEMENTS

This work is supported by the National Key Research and Development Program of China under grant 2018YFB1701203.

REFERENCES

[1]. Xiao G, Huang Y. Equivalent self-adaptive belt grinding for the real-R edge of an aero-engine precision-forged blade. Int J Adv Manuf Technol, 2016,83(9-12):1697–1706.

[2]. Tian Y. B., Liu F, Wang Y, et al. Development of portable power monitoring system and grinding analytical tool. Journal of Manufacturing Processes, 2017,27: 188–197.

[3]. Uhlmann E, Lypovka P, Hochschild L, et al. Influence of rail grinding process parameters on rail surface roughness and surface layer hardness. WEAR, 2016,366–367:287–293.

[4]. Wang W, Salvatore F, Rech J, et al. Comprehensive investigation on mechanisms of dry belt grinding on AISI52100 hardened steel. TRIBOLOGY INTERNATIONAL, 2018,121:310–320.

[5]. Zhou K, Xiao G, Xu J, et al. Material removal behavior of C-f/SiC ceramic matrix composites as a function of abrasive wear during diamond abrasive belt grinding. WEAR, 2021,486.

[6]. Jamshidi H, Budak E. An analytical grinding force model based on individual grit interaction. JOURNAL

OF MATERIALS PROCESSING TECHNOLOGY, 2020,283.

[7]. Kurada S, Bradley C. A review of machine vision sensors for tool condition monitoring. COMPUTERS IN INDUSTRY, 1997,34(1): 55–72.

[8]. Zou L, Huang Y, Zhang G, et al. Feasibility study of a flexible grinding method for precision machining of the TiAl-based alloy. MATERIALS AND MANUFACTURING PROCESSES, 2019,34(10): 1160–1168.

[9]. Pandiyan V, Caesarendra W, Tjahjowidodo T, et al. Predictive Modelling and Analysis of Process Parameters on Material Removal Characteristics in Abrasive Belt Grinding Process. APPLIED SCIENCES-BASEL, 2017,7(4).

[10]. Ren X, Cabaravdic M, Zhang X, et al. A local process model for simulation of robotic belt grinding. Int J Mach Tool Manu, 2007,47(6):962–970.

[11]. Pandiyan V, Caesarendra W, Glowacz A, et al. Modelling of Material Removal in Abrasive Belt Grinding Process: A Regression Approach. SYMMETRY-BASEL, 2020,12(1).

[12]. Ren L, Zhang G, Wang Y, Zhang Q, et al. 2021,“A new in-process material removal rate monitoring approach in abrasive belt grinding. Int J Adv Manuf Technol, 2021,116(9-10): 3367.

[13]. Zhong G, Kang M, Yang S. Precision Obtained Using an Artificial Neural Network for Predicting the Material Removal Rate in Ultrasonic Machining. APPLIED SCIENCES-BASEL, 2017, 7(12).

[14]. Shen J, Pei Z. J, Lee E. S. Support vector fuzzy adaptive network in the modeling of material removal rate in rotary ultrasonic machining. JOURNAL OF MANUFACTURING SCIENCE AND ENGINEERING-TRANSACTIONS OF THE ASME, 2008, 130(4).

[15]. Gao K, Chen H, Zhang X, et al. A novel material removal prediction method based on acoustic sensing and ensemble XGBoost learning algorithm for robotic belt grinding of Inconel 718. Int J Adv Manuf Technol, 2019,105(1-4): 217–232.

[16]. Wang N, Zhang G, Ren L, et al. In-process material removal rate monitoring for abrasive belt grinding using multisensor fusion and 2D CNN algorithm. Int J Adv Manuf Technol.2022.

[17]. Wang N, Zhang G, Pang W, et al. Novel monitoring method for material removal rate considering quantitative wear of abrasive belts based on LightGBM learning algorithm. Int J Adv Manuf Technol, 2021,114(11-12): 3241–3253.

[18]. Eren L, Ince T, Kiranyaz S. A Generic Intelligent Bearing Fault Diagnosis System Using Compact Adaptive 1D CNN Classifier. JOURNAL OF SIGNAL PROCESSING SYSTEMS FOR SIGNAL IMAGE AND VIDEO TECHNOLOGY, 2019,91(2):179–189.

[19]. Caesarendra W, Triwiyanto T, Pandiyan V, et al. A CNN Prediction Method for Belt Grinding Tool Wear in a Polishing Process Utilizing 3-Axes Force and Vibration Data. ELECTRONICS, 2021,10(12).

[20]. Zhang J, Sun Y, Guo L, et al. A new bearing fault diagnosis method based on modified convolutional neural networks. CHINESE JOURNAL OF AERONAUTICS, 2020,33(2): 439–447.

[21]. He K, Zhang X, Ren S, et al. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification,” 2015 IEEE INTERNATIONAL CONFERENCE ON COMPUTER VISION (ICCV), 1026–1034.

[22]. Bourennane S, Fossati C, Lin T. Noise Removal Based on Tensor Modelling for Hyperspectral Image Classification. REMOTE SENSING, 2018,10(9).

[23]. Ioffe S, Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. INTERNATIONAL CONFERENCE ON MACHINE LEARNING, 2015, 37:448–456.

[24]. Zhang W, Peng G, Li C, et al. A New Deep Learning Model for Fault Diagnosis with Good Anti-Noise and Domain Adaptation Ability on Raw Vibration Signals. SENSORS, 2017,17(2).

[25]. Hu J, Shen L, Sun G, et al. Squeeze-and-Excitation Networks. 2018 IEEE/CVF CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR), 7132–7141.

[26]. Langer S. Analysis of the rate of convergence of fully connected deep neural network regression estimates with smooth activation function. JOURNAL OF MULTIVARIATE ANALYSIS, 2021,182.

[27]. Wu D, Yan P, Pei J, et al. Quality estimation method for gear hobbing based on attention and adversarial transfer learning. Measurement, 2022,188.

[28]. Jin T, Yan C, Chen C, et al. Light neural network with fewer parameters based on CNN for fault diagnosis of rotating machinery. Measurement, 2021, 181.

[29]. Kucheryavskiy S, Zhilin S, Rodionova O, et al. Procrustes Cross-Validation-A Bridge between Cross-Validation and Independent Validation Sets. ANALYTICAL CHEMISTRY, 2020, 92(17):11842–11850.

[30]. Barakat A, Bianchi P. CONVERGENCE AND DYNAMICAL BEHAVIOR OF THE ADAM ALGORITHM FOR NONCONVEX STOCHASTIC OPTIMIZATION. SIAM JOURNAL ON OPTIMIZATION, 2021,31(1): 244–274.