www.cad-journal.net Computer-Aided Design AND APPLICATIONS

GA-CNN-based Tool Wear State Identification Method

Yujun Wang[1] and Jian Mao[2]

[1] Shanghai University of Engineering Science, wyj17751829858sues@163.com
[2] Shanghai University of Engineering Science, jmao@sues.edu.cn

Corresponding author: Jian Mao, jmao@sues.edu.cn

Abstract. Accurate prediction of tool wear state on milling and grinding machines is critical to process safety and equipment maintenance in manufacturing. Aiming at the problems of low accuracy and poor efficiency of traditional tool wear state identification methods, an intelligent identification method of tool wear state based on a genetic algorithm (GA) optimized convolutional neural network (CNN) is proposed. Based on the data collected by the sensor during the tool cutting process, we analyze the characteristic information and influencing factors of tool wear, design the CNN structure applicable to wear state identification, and combine it with GA to obtain the optimal CNN model initialization parameters and network architecture. The model is then trained and validated using historical production data to predict the tool wear state for a new operating state. The results show that the CNN model based on GA optimization has strong iterative updating abilities and can improve the identification accuracy of the tool wear state.

Keywords: Genetic algorithm, Convolutional neural network, Tool wear, Milling and grinding machines
DOI: https://doi.org/10.14733/cadaps.2026.210-221

1 INTRODUCTION

Milling and grinding machine tool processing technology is an important advance in the intelligence of modern manufacturing. It can complete nearly 80% of parts processing and has significant applications, especially in the fields of aerospace, precision manufacturing, and mold processing. Milling and grinding machines can complete multiple machining tasks in a relatively short period of time, reducing the frequency of workpiece clamping and tool changes, reducing machining errors, improving machining quality, and shortening the production cycle [1-2]. With the continuous development of global technology, the digital, information-based and intelligent transformation of milling and grinding machine tools has become the mainstream trend. It is required that their numerical control systems be able to achieve self-prediction, so as to improve system reliability and production continuity. The wear status of tools is a key factor affecting the normal operation of machine tools. Therefore, the development of an efficient and accurate tool wear prediction technology is particularly urgent.

Currently, Li et al. [3] propose a tool wear prediction scheme based on feature transfer learning. Features related to tool wear are screened using GA, and the maximum mean square deviation (MMD) method is used to evaluate the similarity of the features. The particle swarm optimization support vector machine (PSO-SVM) model is used to predict the tool wear state during the machining of a new tool to achieve an accurate prediction of the tool wear state. Cheng et al. [4] proposed a new method for tool wear prediction based on whale optimization algorithm (WOA) to optimize support vector machine (SVM), based on extracting multi-domain features of cutting force and vibration signals in the time domain, frequency domain, and time-frequency domain. Cheng et al. [5] obtained the required one-dimensional signal by preprocessing the force and vibration signals and used the Gramian angle field (GAF) to process the one-dimensional signal into a data matrix. The deep residual transformation neural network (ResNext) automatically extracted the features of the data matrix to establish a multi-signal tool wear prediction model based on GAF and ResNext, which can achieve fast and accurate tool wear prediction. Zhang et al. [6] used image detection methods to calibrate tool wear parameters and compared the predicted values obtained from the model with the experimental values to establish a roughness model for simulating the milling topography after milling, which calculates tool wear based on the tool movement trajectory and roughness calculation principles. Natarajan et al. [7] constructed a balanced virtual instrument framework that is a perfect match for the physical system, collects data from the physical system, trains algorithms based on machine learning (ML) classification, and calculates with the help of a confusion matrix using a probabilistic neural network (PNN) to form a tool condition monitoring system through a digital twin model. Zhang et al. [8] proposed a new method for identifying tool wear based on converting force signals into two-dimensional images. A CNN model that considers both high-dimensional and low-dimensional image features was proposed to intelligently and accurately identify the degree of tool wear.

In summary, in order to achieve accurate prediction in tool condition monitoring, two main methods are used: physics-model-driven and data-driven. This paper proposes a GA-CNN-based tool wear state identification model based on data-driven.

2 THEORETICAL BASIS

2.1 CNN Model

Convolutional neural networks [9-10] are a class of feedforward neural networks with a deep structure that includes convolution or correlation calculations. Inspired by the mechanisms of the visual nerve, the design extracts the features of the input data by establishing multiple filters. Convolutional neural networks fuse feature extraction and feature classification into a single learning body, so that starting from the processing of raw data, optimization can be carried out simultaneously through backpropagation, extracting and classifying features to obtain more accurate results. A typical CNN network consists of an input layer, a convolutional layer, a pooling layer, a fully connected layer, and an output layer, as shown in Figure 1.

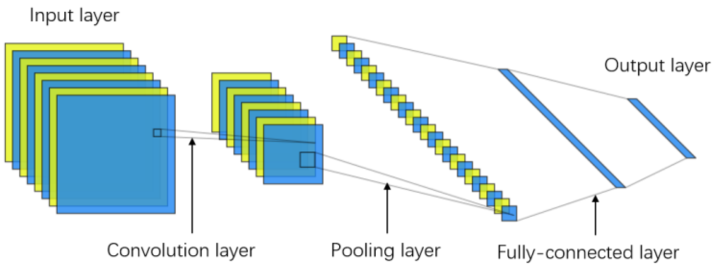

Figure 1: CNN model.

The convolutional process can effectively extract features. However, due to the single-channel method, the extracted features are not multiscale, and local features are integrated into global features in the lower network. The process is shown in Equation (2.1). where denotes input of feature vector N_j, $l-1$ is the network of l layer, x_i^{l-1} indicates the i input feature vector of the l layer, x_j^l indicates the vector obtained after the j convolution calculation on layer l, a_{ij}^l denotes the j convolution kernel that convolves the l layer with the i input feature vector, $f(\bullet)$ denotes an activation function, b_j^l denotes the j bias vector of the l layer, and $*$ denotes the j convolution operation. Pooling is a method of reducing the number of features and extracting features a second time. Pooling processes always follow the convolution process, and maximum pooling can extract discrete features. The mathematical expression of maximum pooling is shown in Equation (2.2). The activation value $\alpha_t^l(t)$ of the t neuron in the i feature vector of the l layer is indicative of the width of the pooling area w. The result of pooling corresponds $p_i^{l+1}(j)$ to the $l+1$ layer of neurons. The fully connected layer is used to implement the function of classification identification. The feature data after pooling and convolution is used as the input to the fully connected layer, and the classification result is used as the output. The fully connected layer usually uses the ReLU function as the activation function [11-12]. The ReLU function is shown in Equation (2.3).

$$x_j^l = f(\sum_{i\in N_j} x_i^{l-1} * a_{ij}^l + b_j^l) \tag{2.1}$$

$$p_i^{l+1}(j) = max_{(j-1)w+1\le t\le jw}\{\alpha_t^l(t)\} \tag{2.2}$$

$$ReLU: f(x) = max(0,x) = \begin{cases} x, x \ge 0 \\ x, x \le 0 \end{cases} \tag{2.3}$$

Since there is no power operation, the network can converge faster, solving the problem of gradient disappearance and achieving better performance. Since the output layer of the last layer of a convolutional neural network is usually a category with different labels, the last fully connected layer is combined with Softmax logistic regression to complete the classification task.

2.2 GA Optimized CNN Algorithm

The genetic algorithm (GA) [13-14] was first proposed by John Holland in the United States in the 1970s. The algorithm was designed based on the evolutionary laws of organisms in nature. It is a computational model that simulates the biological evolutionary process of natural selection and genetic mechanisms in Darwin's theory of evolution. It is a method that searches for the optimal solution by simulating the natural evolutionary process.

CNN training results highly depend on the network architecture and various hyperparameter settings, such as the convolution kernel size, the number of layers, the learning rate, the activation function, and the pooling method. Traditionally, these hyperparameters often need to be adjusted manually based on expert experience or through methods such as grid search or random search. However, the parameter space of this method is usually very large, and manual parameter adjustment or grid search calculations are extremely time-consuming, resulting in high tuning costs. Traditional methods may have difficulty escaping from local optima, which affects the final performance and can easily lead to local optima. The GA algorithm can optimise the structure of the feature extraction layer (such as the convolutional layer and pooling layer) in the CNN network, and select which features are the most relevant, thereby reducing the interference of irrelevant features on model training. Therefore, this paper combines the GA algorithm to search for the optimal combination of these hyperparameters. Each individual is represented as a network hyperparameter configuration, and after a series of genetic operations (selection, crossover, mutation), the hyperparameter combination that can maximize the performance of the CNN is found. This method is more efficient than traditional manual adjustment and can avoid local optima [15-16].

Therefore, introducing GA into the hyperparameter search of CNN and using the evolutionary process of GA to find the optimal hyperparameter configuration has become an attractive solution.

3 GA-CNN MODEL TRAINING FRAMEWORK

The complex force signal data of the milling and grinding machine tool is optimised based on the CNN combined with the GA algorithm, forming a GA-CNN tool wear state identification model, which provides a prediction method for the tool wear state of the milling and grinding machine tool. Figure 2 shows the flow chart of the GA-CNN model.

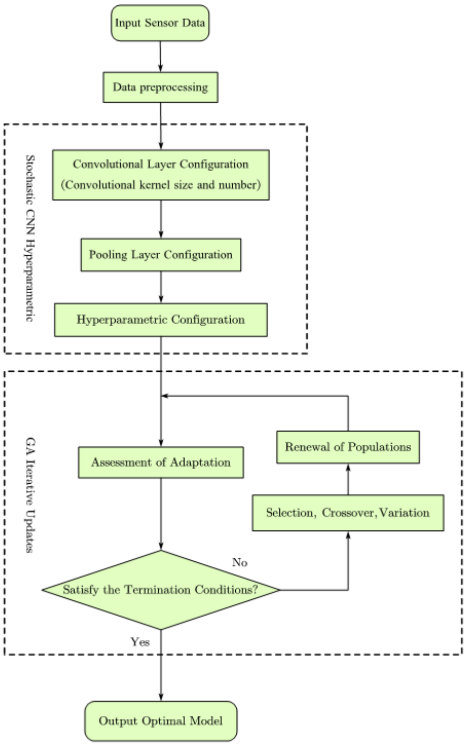

Figure 2: Flowchart of the GA-CNN model.

3.1 Force Signal Data Acquisition

During the grinding test on the milling and grinding machine, a model 9272 dynamometer was used to accurately measure the grinding force signal, and it was connected to a charge amplifier (HR-CA) to ensure the accuracy and stability of the signal acquisition. The measuring range of the dynamometer is -200N to 200N, and the data sampling frequency is set to 200Hz, as shown in Figure 3.

During the experiment, the dynamometer records the grinding force signal in real time and transmits the data to the dynamometer software for processing and analysis. However, after the recording was complete, it was found that the data fluctuated quite violently and with a high frequency, indicating that there may be strong high-frequency noise in the signal. Therefore, we regard this signal as a discrete signal and further analyze it to ensure the accuracy and reliability of the measurement data.

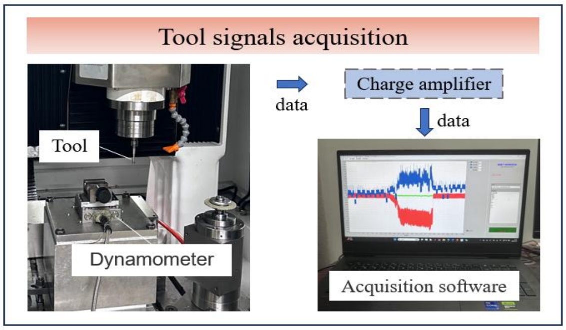

Figure 3: Tool signals acquisition.

During the analysis, special attention should be paid to the signal characteristics in the no-load state. Under no-load conditions, the dynamometer will still record a certain degree of high-frequency random vibration signal, which may be due to the background noise of the dynamometer itself, mechanical vibrations in the experimental environment, or electronic noise. These noise signals may interfere with the measurement of the grinding force, affecting the final analysis results. Therefore, after the data acquisition is complete, an appropriate data processing method needs to be applied in the dynamometer software to optimise the signal quality and eliminate or reduce the impact of high-frequency noise.

3.2 Initialise the Genetic Algorithm Population

In the process of using GA to optimise CNN, it is first necessary to define population individuals and represent key hyperparameters in the CNN architecture using genetic coding. These hyperparameters include the number of convolutional layers, the number and size of the convolutional kernels in each layer, the type of pooling layer and its configuration, the type of activation function, the learning rate, the batch size, etc. These hyperparameters directly determine the structure, training process, and regression performance of the CNN model.

In the initiation phase, we usually generate a certain number of individuals in a random way to form the initial population. Each individual corresponds to a CNN model, and its hyperparameters are determined by genetic coding. This random initiation method explores a diversity of search spaces and improves the possibility of the GA escaping from local optima during the optimization process.

3.3 Fitness Assessment

The adaptability function is used to measure the performance of the CNN model. During training, each individual represents the CNN architecture trained using machine force signal data.

During training, the fitness is evaluated based on the CNN model's performance on the validation set. The most common fitness function used in this paper is the loss function, which measures the error between the model's prediction and the true value. The smaller the loss value, the better the model performance. For machine tool force signal identification tasks, the fitness can directly reflect the model's performance in recognising different operating or fault states.

3.4 Iterative Updating

GA optimises the CNN architecture iteratively. Each iteration consists of the following steps:

- Train the CNN: Train each individual in the population and calculate the fitness value;
- Selection: Roulette selection or tournament selection is used to select individuals with higher fitness from the current population as parents of the next generation;
- Crossover: New individuals are generated through cross-breeding, so that they inherit superior genes;

- Mutate: randomly change some gene values to increase the diversity of the population and prevent it from falling into a local optimum;
- Population replacement: Replace individuals with lower fitness with newly generated individuals to form a new population.
- The process iterates continuously until the termination condition is met and the fitness converges.

3.5 Output Optimal CNN Architecture

When the algorithm converges or reaches the termination condition, the individual with the highest fitness in the final generation is selected as the optimal CNN architecture, which can more accurately identify the machine tool force signal. Subsequently, the final training is performed on the entire training set using this optimal architecture and hyperparameters to build the model with the best performance. Finally, the model's performance is evaluated on a test set to verify the classification accuracy and ensure good generalisation ability.

In this chapter, a training framework for a GA-CNN-based tool wear state identification model is designed for the identification task of milling and grinding machine force signals. This method uses the global search capability of GA to optimise the architecture and hyperparameters of the CNN, thereby finding the optimal CNN configuration for the machine tool force signal and ultimately improving the identification performance.

4 RESULTS AND DISCUSSION

The GA-CNN-based tool wear state identification model proposed in the text is used to predict the force signal of the milling and grinding machine tool. The data collected by the machine tool dynamometer is substituted into the above model, and the process is as shown in Chapter 3 to achieve tool force signal prediction.

4.1 Laboratory Equipment and Processing Materials

The grinding experiment in this paper was carried out on the ZCS-XM30 ultrasonic milling and grinding composite high-precision CNC machine tool, as shown in Figure 4. The machine tool is equipped with high-precision spindles such as a high-speed ultrasonic aerostatic spindle and an aerostatic wheel dressing spindle. It can achieve rapid movement in the X, Y, and Z directions, with a resolution of 0.1μm and a travel range of 200mm×200mm×180mm. The spindle speed can reach up to 80,000 rpm, and the whole machine is highly precise.

Figure 4: Milling and grinding machine.

The experimental processing material is a laminate composed of T-300 carbon fibre and thermosetting epoxy resin. The thickness of each ply is 0.1 mm, all the fibres in each ply are oriented at 0°, the initial number of plies is 100, and the total thickness is 10 mm, as shown in Figure 5. This paper uses a high-speed electroplated CBN grinding wheel with the model number 85412-BM as shown in Figure 6.

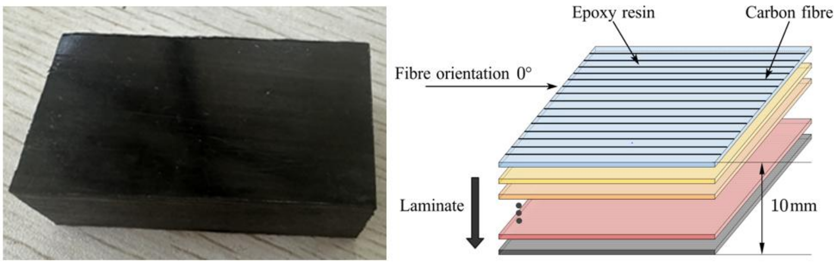

Figure 5: Carbon fibre composite laminate.

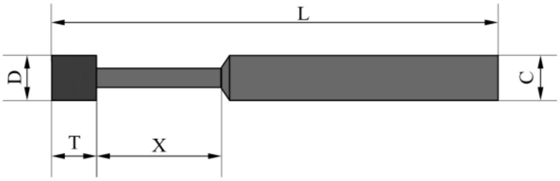

Figure 6: High-speed electroplated CBN grinding wheel.

In Figure 6, the diameter of the grinding wheel head D is 3 mm; the length of the grinding wheel head T is 4.8 mm; the length of the root cut X is 12.7 mm; the total length of the grinding wheel L is 31.7 mm; the diameter of the template handle Electroplated grinding wheel C is 3.17 mm, with 230 particles. Electroplated grinding wheels only have one layer of abrasive and cannot be used for resharpening. The experiment was carried out using a surface grinding method, dry grinding 40μm, and applying vibration along the spindle direction of the grinding wheel. The machining parameters used in the experiment were kept constant, as shown in Table 1.

Main parameters	Main shaft speed (rpm)	Grinding depth (μm)	Feed speed (mm/r)
Data	600	40	0.054mm/r

Table 1: Experimental parameters.

4.2 Force Signal Processing

During the grinding process, the cutting force is usually divided into three components: the horizontal cutting force in the X-axis direction, the cutting force in the Y-axis feed direction, and the axial force in the Z-axis vertical direction. Among these, the axial force is often much greater than the cutting forces in the other two directions, which is due to the characteristics of the grinding process. Therefore, in order to more accurately analyse the changing law of the grinding force and its impact on the surface quality of the workpiece and the wear of the tool, this experiment focuses mainly on the axial force in the vertical Z-axis direction. After low-pass filtering and de-drifting the signal, a stable signal of the axial force is obtained, as shown in Figure 7.

The processed axial force signal shows a significant change in characteristics, with smooth and stable characteristics, which not only improve the accuracy of the GA-CNN model in feature extraction but also provide reliable data support for predicting the wear status of the tool.

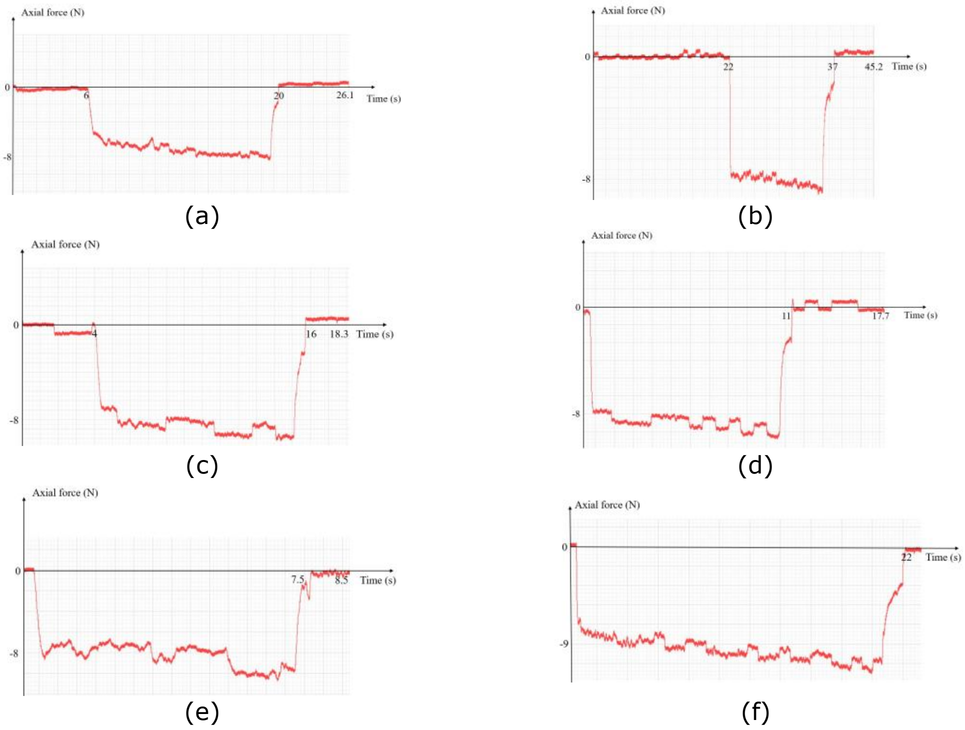

Figure 7: Recording of on-site mechanical signals during the grinding process:(a) The first group, (b) The second group, (c) The third group, (d) The fourth group, (e) The fifth group, and (f) The sixth group.

4.3 Training Effect

The CNN model for tool wear state identification is trained to generate CNN hyperparameters as shown in Table 2.

Layer	*Optimisation scope*	*Optimal parameter*
Number of convolutional layers	[1,2,3,4]	2
Number of pooling layers	[1,2,3,4]	2
Convolution kernel size	[1,2,3,4]	3
Number of convolution kernels	[16,32,64,128]	64
Pooling Window Size	[1,2,3,4]	2
Dropout	[0.2,0.5,0.7]	0.5
Learning rate	[0.01,0.001,0.0001]	0.001

Table 2: CNN hyperparameters.

The coded data we use consists of a convolutional layer, a pooling layer, a fully connected layer, and an optimiser where the convolutional layer filters=64, kernel_size=3, the fully connected layer units=112, and the optimiser is Adam.

As shown in Figure 8, the training loss and the test loss of the model rapidly decrease in the 5-15 iteration interval, quickly converging from their initial high values to close to 0. Subsequently, both curves remained within the range of lower loss values, and the gap between the training loss and the test loss was very small, indicating that the model did not show obvious overfitting in the later stages of training. Overall, the model converged effectively in a relatively short training period and generalised well.

In Figure 9, the overall trend of the predicted values of the GA-CNN model is consistent with the true values, indicating that the model has a good fitting effect in the tool wear state identification task. Through a calculation error, it is found that the fitting error between the true value and the predicted value is approximately ±0.1N, and the overall error is small, indicating that the GA-CNN model has learned the mapping relationship between the tool wear state and the force signal well, and can predict the wear state more accurately.

We carried out ablation experiments by replacing the original model with different models, replacing the GA-CNN model with three models: SVM, BP, and CNN. We define MAE, RMSE, R2 as the performance evaluation metrics of the model, and we evaluate the performance results of the model according to the performance evaluation metrics, as shown in Table 3. We demonstrate that the GA-CNN model used is the optimal choice, improving the R2 metric by about 2% compared to the other three models.

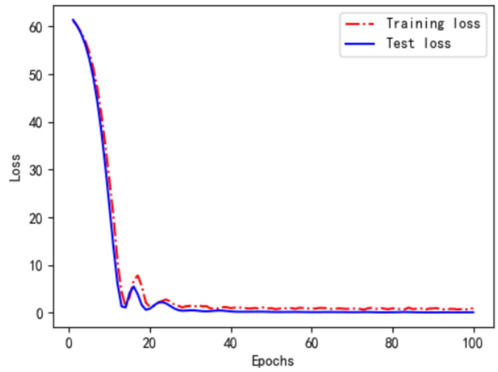

Figure 8: Training loss and testing loss.

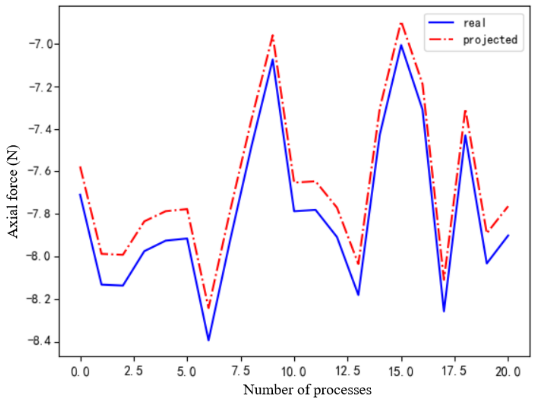

Figure 9: GA-CNN model prediction graph.

Algorithms	*MAE*	*RMSE*	*R2*
SVM	0.260	0.268	0.855
BP	0.276	0.287	0.845
CNN	0.228	0.235	0.863
GA-CNN	0.160	0.164	0.884

Table 3: Mean absolute error and root mean square error of different prediction models.

We deployed the model to the computer, and the data was transferred to the computer in real time with a transmission interval of 1s. The model predicts the magnitude of the axial force on the machine in real time from the input data, as shown in Figure 10.

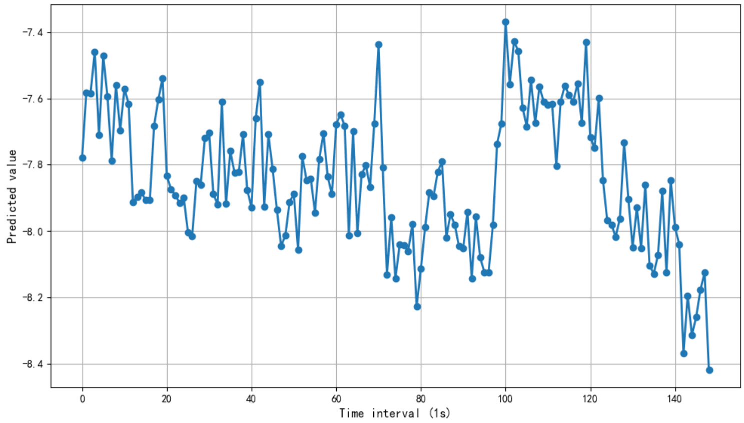

Figure 10: GA-CNN model prediction graph.

5 CONCLUSIONS

Traditional methods for identifying the wear status of tools on milling-grinding machines have problems with low accuracy and efficiency, leading to a decline in machining quality and limited production efficiency. This paper proposes an intelligent identification method based on GA-optimised CNN to improve the accuracy and stability of tool wear state prediction. The following conclusions were drawn from the research:

- The GA algorithm optimises the hyperparameter configuration of the CNN through a global search, such as the number of convolutional layers, the size of the convolutional kernel, the parameters of the pooling layer, etc., which improves the CNN's feature extraction capability for complex force signal data. The optimised CNN model can more accurately identify tool wear on milling-grinding machines, thereby improving the accuracy and stability of force signal prediction.
- The experimental results show that the GA-optimised CNN model converges quickly in the tool wear state identification task, with both the training loss and test loss remaining at a low level. There is no obvious overfitting, indicating that the model has good generalisation performance and can accurately learn the mapping relationship between the tool wear state and the force signal. Compared with traditional CNN, the GA-optimised CNN model has better identification accuracy and stability, providing an efficient prediction method for the wear status of milling and grinding machine tools.

The GA-CNN tool wear state identification model can effectively improve the prediction accuracy of tool wear state. However, in practical applications, limited by the difficulty of obtaining high-quality experimental data, how to use limited working data for efficient training remains a key challenge. Therefore, future research will focus on exploring how to fully tap and utilise the working data of existing equipment to improve the adaptability and generalisation ability of the model, so as to achieve more accurate identification of tool wear state.

ACKNOWLEDGMENTS

This research is sponsored by the Natural Science Foundation of Sichuan Province (2023 NSFSC0366).

Yujun Wang, https://orcid.org/0009-0000-0941-9965
Jian Mao, https://orcid.org/0009-0006-2061-8395

REFERENCES

[1] Liu, Z.-Y.; Guo, Y.-B.; Sealy, M.-P.; Liu, Z.-Q.: Energy consumption and process sustainability of hard milling with tool wear progression, Journal of Materials Processing Technology, 229, 2016, 305-312. https://doi.org/10.1016/j.jmatprotec.2015.09.032
[2] Li, K.; Li, Z.; Jia, X.; Liu, L.; Chen, M.: A domain adversarial graph convolutional network for intelligent monitoring of tool wear in machine tools, Computers & Industrial Engineering 187, 2024, 109795. https://doi.org/10.1016/j.cie.2023.109795
[3] Li, J.; Lu, J.; Chen, C.; Ma, J.; Liao, X.: Tool wear state prediction based on feature-based transfer learning, The International Journal of Advanced Manufacturing Technology, 113, 2021, 3283-3301. https://doi.org/10.1007/s00170-021-06780-6
[4] Cheng, Y.; Gai, X.; Jin, Y.; Guan, R.; Lu, M.; Ding, Y.: A new method based on a WOA-optimized support vector machine to predict the tool wear, The International Journal of Advanced Manufacturing Technology, 121, 2022, 6439-6452. https://doi.org/10.1007/s00170-022-09746-4
[5] Cheng, Y.; Lu, M.; Gai, X.; Guan, R.; Zhou, S.; Xue, J.: Research on multi-signal milling tool wear prediction method based on GAF-ResNext, Robotics and Computer-Integrated Manufacturing, 85, 2024, 102634. https://doi.org/10.1016/j.rcim.2023.102634
[6] Zhang, L.; Zheng, M.; Zhang, W.; Li, K.: Influence and prediction of tool wear on workpiece surface roughness based on milling topography analysis, The International Journal of Advanced Manufacturing Technology, 122, 2022, 1883-1896. https://doi.org/10.1007/s00170-022-09939-x
[7] Natarajan, S.; Thangamuthu, M.; Gnanasekaran, S.; Rakkiyannan, J.: Digital Twin-Driven Tool Condition Monitoring for the Milling Process, Sensors, 23, 2023, 5431. https://doi.org/10.3390/s23125431
[8] Zhang, Y.; Qi, X.; Wang, T.; He, Y.: Tool Wear Condition Monitoring Method Based on Deep Learning with Force Signals, Sensors, 23, 2023, 4595. https://doi.org/10.3390/s23104595
[9] He, Y.; Xiang, H.; Zhou, H.; Chen, J.: In-situ fault detection for the spindle motor of CNC machines via multi-stage residual fusion convolution neural networks, Computers in Industry, 145, 2023, 103810. https://doi.org/https://doi.org/10.1016/j.compind.2022.103810
[10] Liao, L.; Zhao, Y.; Wei, S.; Wei, Y.; Wang, J.: Parameter Distribution Balanced CNNs, IEEE Transactions on Neural Networks and Learning Systems, 31(11), 2020, 4600-4609. https://doi.org/10.1109/TNNLS.2019.2956390
[11] Xu, Y.; Zhang, H.: Convergence of deep ReLU networks, Neurocomputing, 571, 2024, 127174. https://doi.org/https://doi.org/10.1016/j.neucom.2023.127174
[12] He, J.; Li, L.; Xu, J.: Approximation properties of deep ReLU CNNs, Research in the Mathematical Sciences, 9(3), 2022, 38. https://doi.org/10.1007/s40687-022-00336-0
[13] Dao, S. D.; Abhary, K.; Marian, R.: An innovative framework for designing genetic algorithm structures, Expert Systems with Applications, 90, 2017, 196-208. https://doi.org/https://doi.org/10.1016/j.eswa.2017.08.018

[14] Katoch, S.; Chauhan, S. S.; Kumar, V.: A review on genetic algorithm: past, present, and future, Multimedia Tools and Applications, 80(5), 2021, 8091-8126. https://doi.org/10.1007/s11042-020-10139-6
[15] Fan, Z.; Bai, K.; Zheng, X.: Hybrid GA and Improved CNN Algorithm for Power Plant Transformer Condition Monitoring Model, IEEE Access, 12, 2024, 60255-60263. https://doi.org/10.1109/ACCESS.2023.3316251
[16] Liang, J.; Cao, H.; Lu, Y.; Su, M.: Architecture search of accurate and lightweight CNNs using genetic algorithm, Genetic Programming and Evolvable Machines, 25(1), 2024, 13. https://doi.org/10.1007/s10710-024-09484-4