IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 25 August 2023, accepted 26 October 2023, date of publication 2 November 2023, date of current version 8 November 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3329830

RESEARCH ARTICLE

Saw Blade Wear Identification Based on Data Enhancement and Feature Fusion

CHENGCHAO WANG[1], XIANGJIANG WANG[1], AND CHAO ZENG[2]
[1]College of Mechanical Engineering, University of South China, Hengyang 421001, China
[2]School of Nuclear Science and Technology, University of South China, Hengyang 421001, China

Corresponding author: Xiangjiang Wang (wangxiangjiang72@163.com)

This work was supported by the Natural Science Foundation of Hunan Province, China, under Grant 2023JJ50131.

ABSTRACT In order to solve the problem of low accuracy of tool wear detection due to the poor quality of generated data under small sample problems, a deep learning model based on data enhancement and feature fusion is proposed. Firstly, in order to solve the problem that there is no quality evaluation standard in the training process of the traditional generative adversarial network (GAN), the K nearest neighbor algorithm is proposed to test the data generated by the GAN model for the second time. The improved GAN model can be automatically trained to get the optimal model according to the second test results. Secondly, in order to enhance the anti-interference effect of the model, a double-path parallel convolutional neural network (DPCNN) which combines with the characteristics of frequency domain and time-frequency domain is constructed to analyze the wear data. Furthermore, the hyperparameters of the model are optimized by Bayesian optimization algorithm (BOA). Finally, the effectiveness of this method is verified in the saw blade wear detection experiment. The results show that the performance of this model is better than other models, and the accuracy rate in the experimental detection reaches 100%.

INDEX TERMS Saw blade wear identification, feature fusion, Bayesian optimization algorithm, double-path parallel convolution neural network, generative adversarial networks.

I. INTRODUCTION

Since the development of manufacturing, the effective working process of cutting tools as an important tool has been paid more and more attention by researchers. Furthermore, with the continual improvement of the processing environment, materials, precision and other requirements, tool wear as the main failure form has been widely studied by scholars. At present, although a large number of wear monitoring systems have been established for single-edge, double-edge or triple-edge cutting tools such as lathe tools [1] or milling cutters [2], and achieved some progress. However, there is still a lack of research on wear monitoring of ultra-multi-edge tools. As a representative of this kind of cutting tool, the saw blade plays an important role in the cutting process of materials. Therefore, to ensure the security of its operation, a targeted wear monitoring system is established with necessity.

Current research on tool wear monitoring is mainly based on optical measurement and machine vision, or sensing technology and artificial intelligence. Machine vision as one branch of the whole tech category uses imaging equipment to collect image data to detect tool wear [3]. Such as Peng et al. [4] collect milling cutter wear pictures through charge-coupled device cameras, and analyze the images by gray level co-occurrence matrix, thus establishing a milling cutter wear detection system. In addition, Bagga et al. [5] use artificial neural networks to analyze the tool wear images captured by industrial cameras regularly and establish a tool wear detection system. These mentioned methods above are feasible without doubt; but considering that the imaging equipment is expensive and easily affected by the environment, such methods based on optical measurement and machine vision are not suitable for practical production. However, the combination of sensing technology and

The associate editor coordinating the review of this manuscript and approving it for publication was Liudong Xing.

artificial intelligence can effectively break through the above limitations and has been wildly employed nowadays [6]. Such as, Zhou et al. [7] build a milling tool wear monitoring system that combines a kernel extreme learning machine and a current sensor. Ying et al. [8] and Ge et al. [9] analyses signals collected by vibration sensor and realizes wear monitoring of broach and milling cutter via optimizing support vector machine. Twardowski et al. [10] compared the performance of different classification trees in milling cutter wear prediction and built a monitoring model with optional algorithms. While these approaches combining artificial intelligence shallow learning models with sensing technology provide a solution for tool wear prediction, the detection accuracy does not meet the requirement as expected due to the limited feature extraction ability of the shallow learning model. To this end, Dong et al. [11] commit to research the performance of a deep learning model (backpropagation neural network) in woodworking milling tool wear identification, and the finding shows that the model works well. Yang et al. [12] build a self-coding model based on a backpropagation neural network, which indicates that the model has a better performance than other models in milling experiments. But even though this kind of model is able to extract signal depth feature, it is its poor robustness that cannot be ignored because the model only takes single signal as input. The study showing that constructing prediction model with multi-type signals as input has a stronger anti-interference capacity [6]. For instance, the deep learning model with multiple input parameters such as cutting force, vibration, acoustic emission, etc., performs well in tool wear prediction [13], [14]. However, in the practical equipment operation, it is common for the sensors' installation to be limited because of insufficient space. Therefore, Wu et al. [15] integrates these characteristics of vibration signals, including "time" and "frequency", to realize the purpose that the wear of cutting tool in computer numerical control machine can be detected. Likewise, for the milling wear process, Zheng et al. [16] transform the feature space of the original signal to build a high-accuracy recognizing model by means of empirical mode decomposition, variational mode decomposition, and Fourier transforms.

These studies show that the tool wear monitoring model based on the deep learning model and multi-domain feature fusion gets advantages in high recognition accuracy and anti-interference ability. However, there is still a problem that obtaining a large amount of data for deep learning model training is difficult in practical application. The research in [17] shows that this problem can be effectively solved by expanding the data. At present, in mechanical fault diagnosis, Generative Adversarial Networks (GAN) are widely used to address the small sample problem owing to its powerful data generation capability [18]. For example, [19] combines finite element simulation technology with GAN model and realizes the aim, expansion of the cutting force signal of the tool. Molitor et al. [20] use different types of generative adversarial networks to enhance data, which improves the accuracy of tool wear detection by about 18%. In addition, Shah et al. [21] constructed a singular generative adversarial network, which expanded the scale diagram of acoustic emission and vibration signals of tool wear signal. The above data enhancement method based on the GAN model expands the experimental data of tool wear, but the authenticity of the extended data cannot be guaranteed because there is no evaluation standard for the training performance of the GAN model.

Therefore, according to the above research, this paper constructs a saw blade wear detection model, which is based on data enhancement and feature fusion, to improve the accuracy of saw blade wear detection and secure the process of the saw blade-cutting. The training performance of the GAN model is tested by the k-nearest neighbor algorithm, and the validity of generated data can be controlled by the test results. At the same time, a double-path parallel convolution neural network (DPCNN) is constructed to fuse the multi-domain features of saw blade wear vibration data to enhance the robustness of the model. Generally speaking, the main chapters of this paper are organized as follows: Section II introduces the requisite theoretical basis of building the model; Section III builds an optimized model based on the theory mentioned in Section II; Section IV verifies the advantages of the model performance through the cutting wear experiment of circular saw; section V is the conclusion of this paper.

II. THEORETICAL BACKGROUND

A. GENERATIVE ADVERSARIAL NETWORK

The GAN network based on the adversarial game theory, consists of a generator and a discriminator. The training process is shown in Figure 1. According to the actual data, the discriminator checks the generated data to continuously optimize the discrimination indexes. The generator constantly improves the data generation ability to pass the test of the discriminator. Through the confrontation game, the model finally reaches the dynamic Nash equilibrium.

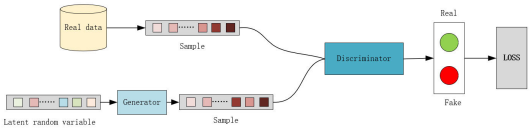


FIGURE 1. GAN.

The training loss can be calculated as follows:

$$\min_G \max_D V(D, G) = E_{x\sim P_{data}(x)}\left[\log D(x)\right] + E_{Z\sim P_{noise}(Z)}\left[\log\left(1 - D\left(G\left(z\right)\right)\right)\right] \quad (1)$$

where $V(D, G)$ is the generator loss, $V(D, G)$ is the discriminator loss, x and z are samples respectively obtained from the true distribution P_{data} and the noise distribution P_{noise}. E is the expectation of the requested type. $D(\cdot)$ is the discriminator model, $G(\cdot)$ is the generator model.

The generator model obtained by training of GAN has good data generation ability. This paper refines the model to

expand the scale of experimental data, as instructed in section III-A for details.

B. CONVOLUTIONAL NEURAL NETWORK

Convolutional neural network (CNN), as a classic deep learning model, has been widely used in identifying signals such as data and images. The basic structure of CNN is shown in Figure 2. In addition, its main functions are realized by the convolution layer, pooling layer, and fully connected layer.

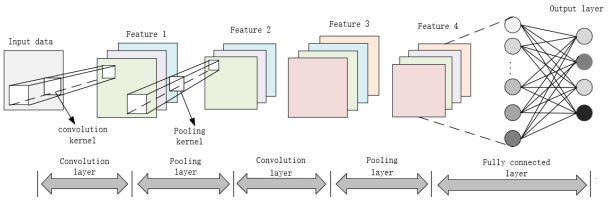


FIGURE 2. CNN structure.

a: CONVOLUTION LAYER

Convolutional layers perform inner products between learnable convolutional kernels and specific regions of the input data. By controlling the movement stride, the convolutional kernels traverse the entire input data to extract signal features. The output of the convolution layer can be calculated as follows:

$$X_j^l = f\left(\sum_{i \in M_j} X_i^{l-1} * \omega_{ij}^l + b_j^l\right) \tag{2}$$

where X_j^l is the l-th characteristic graph output by the j-th convolution layer, M_j is the data set input by this layer, ω_{ij}^l is the convolution kernel of this layer which is the weight matrix corresponding to the input X_i^{l-1}, b_j^l is the corresponding offset value, $f(\cdot)$ is the activation function, $*$ is convolution operation.

b: POOLING lAYER

The pooling layer plays a role in downsampling the convolution features with the main features of the data retained. In this way, the pooling layer can help avoiding the overfitting phenomenon of the model by reducing the amounts of parameters and lessening the quantities of feature dimensions. The common pooling operations are maximum pooling and average pooling, and the maximum pooling operation is selected in this paper. The calculation formula is as follows:

$$a_{ij}^k = \max_{u \in D_j^{k-1}} \left[x_i^{k-1}(u)\right] \tag{3}$$

where a_{ij}^k is the feature map of the j-th pooled region corresponding to the upper convolution feature, D_j^{k-1} is the pool domain of layer k-1, x_i^{k-1} is the i-th characteristic map output by the k-th convolution layer, u is the mapping area of the pooled core in the pooled domain.

c: FULLY CONNECTED LAYER

The function of the fully connected layer is to spread and connect all the feature graphs and provide the required data form for the softmax layer (output layer) by integrating and reducing the dimensions of the extracted features. The calculation process is given below:

$$P_j = \frac{\exp(O_j)}{\sum_{i=1}^{K} \exp(O_i)} \tag{4}$$

where P_j represents the proportion of the output value of the j-th neuron, O_j is the output value of the j-th neuron and K is the number of types that the model needs to classify.

C. BAYESIAN OPTIMIZATION ALGORITHM

Bayesian optimization algorithm (BOA) uses the Bayesian theorem to search for the optimal value of the objective function. The optimization strategy is to build a proxy model for the first, then select the optimal interval updating model according to prior knowledge, and find the optimal solution to the problem finally. The detailed process is shown in Figure 3.

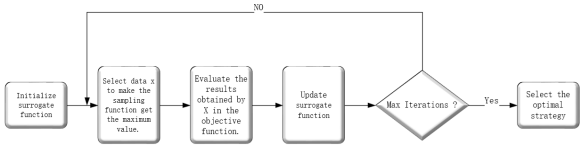


FIGURE 3. Bayesian optimization.

As shown above, it is obvious that the surrogate function and sampling function are the core elements of the BOA optimization process. BOA updates the optimal range of prior distribution through the sampling function and approximates the surrogate function to the target curve. The surrogate function is as follows:

$$y = f\,(x) \sim G_p[\mu\,(x)\,, k(x, x')] \tag{5}$$

Actually, the above formula represents Gaussian process regression. y is the output of the surrogate function under the prior distribution x, $f\,(x)$ is the output of the objective function, $\mu\,(x)$ and $k\,(x, x')$ respectively represent the mean function and covariance function of the real process, G_p is a Gaussian process.

In this paper, the author uses the expected improvement function to update the prior distribution x. In other words, the expected improvement function is used as the sampling function. Calculation process reference formula (6), as shown at the bottom of the next page. $\mu_t(x)$ and $\sigma_t(x)$ are the mean and standard deviation of the actual distribution, $f(x^+)$ is the optimal value corresponding to the optimal position x^+, ε is the control parameter, $\Phi(\cdot)$ is a distribution function, and $\phi(\cdot)$ is a probability density function. $EI(\cdot)$ is the expected improvement function.

III. ALGORITHM OPTIMIZATION

A. DATA EXPANSION MODEL

For the small sample problem, although the GAN described in section II-A can be used for data expansion through training, there are still some problems, there is no performance evaluation standard in the training process of the traditional generative adversarial networks. It is impossible to make the network performance optimal by adjusting the network parameters accurately and artificially.

For this problem, Xu et al. [22] studied the influence of different evaluation methods on the training performance of the GAN network. And the results show that the 1-Nearest Neighbor test is an effective method to evaluate the training performance of the GAN network. On this basis, a GAN model optimized by the k nearest neighbor algorithm (K-GAN) is constructed to expand the data of saw blade wear vibration. The process is shown in Figure 4.

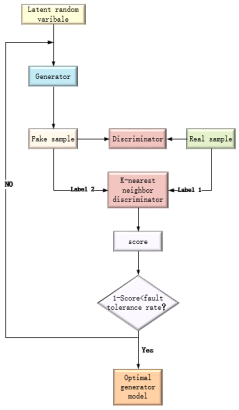


FIGURE 4. K-GAN.

K-GAN model tests the generated data and real data by the K-nearest neighbor algorithm, calculates the Euclidean distance between each data and other data by using the leave one out cross-validation method, Select the class with the highest frequency among the top k nearest types as the data type. When the K- nearest neighbor recognition rate reaches 50%, it shows that the generated data has the same distribution as the real data, at this point, the network training stops, and the optimal model is obtained. The Euclidean distance calculation formula is given as follows:

$$D = \sqrt{\sum_{i=1}^{N} (X_i - Y_i)^2} \tag{7}$$

where N is the data dimension, X_i is the discriminated data and Y_i is the training data set, D is the calculated Euclidean distance.

In this paper, k=11, and the criterion of the K-GAN network is the model score value in the training process. The calculation process is given below.

$$C = 1 - \left| \frac{\sum_{i=1}^{M} K(x_i)}{M} - 0.5 \right| \tag{8}$$

where C is the model score. The formula $\left| \frac{\sum_{i=1}^{M} K(x_i)}{M} - 0.5 \right|$ indicates the model fault tolerance rate, which is set to be less than or equal to 0.01 in this paper, M is the sample size, $K(*)$ is the K-nearest neighbor discriminator, and x_i is the discriminant sample.

B. IDENTIFICATION MODEL

1) DOUBLE-PATH PARALLEL CONVOLUTION NEURAL NETWORK

Due to the complex working environment of a circular saw, the vibration acceleration signal is easily disturbed by environmental noise, and because of the continuity of the vibration signal in the time domain, the time series signal is often complicated. Therefore, in order to enhance the anti-interference ability of the recognition model and improve the recognition accuracy of the model, the author constructs a double-path parallel convolution neural network which is based on feature fusion of frequency domain and time-frequency domain. The frequency domain and time-frequency domain transformations are as follows.

a: FAST FOURIER TRANSFORM

The Fast Fourier Transform is based on the Discrete Fourier transform (DFT), but it effectively avoids the repeated calculation of DFT. The DFT calculation is given below.

$$X(k) = \sum_{n=0}^{N-1} x(n) W_N^{kn} \tag{9}$$

where N is the volume of the discrete signal $x(n)$, $W_N = e^{-j\frac{2\pi}{N}}$, $k = 1, 2, 3 \ldots\ldots N - 1$. $X(k)$ is the value obtained by the DFT of point k. Fast Fourier Transform utilizes the symmetry and periodicity of the W_N term to decompose the DFT of N points into DFTs of N/2 points, thereby reducing its complexity. The decomposition process is as follows:

$$X(k) = \sum_{n=0}^{N-1} x(2n) W_N^{2kn} + W_N^k \sum_{n=0}^{N-1} x(2n+1) W_N^{2kn} \tag{10}$$

$$EI(x) = \begin{cases} \left(\mu_t(x) - f\left(x^+\right) - \varepsilon\right) \Phi\left(\dfrac{\mu_t(x) - f\left(x^+\right) - \varepsilon}{\sigma_t(x)}\right) \\ +\sigma_t(x)\phi\left(\dfrac{\mu_t(x) - f(x^+) - \varepsilon}{\sigma_t(x)}\right) \sigma_t(x) > 0 \\ 0 \ \sigma_t(x) = 0 \end{cases} \tag{6}$$

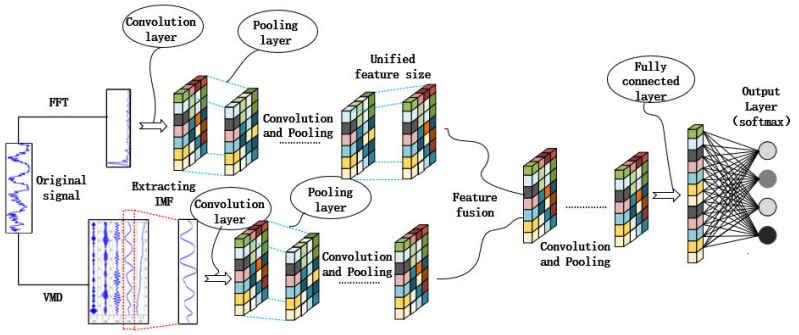


FIGURE 5. DPCNN.

The first term on the right part of the equation is the DFT transform of the even-numbered sequence $x(n)$, and the second term is the DFT transform of the odd-numbered sequence.

b: VARIATIONAL MODE DECOMPOSITION

Variational Mode Decomposition decomposes the original signal into several subsequences of different frequency scales and relative stability by constructing a variational problem. This approach effectively mitigates the influence of interfering signals on the analysis. The constrained variational model is as follows.

$$\min_{\{u_k\}:\{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[(\delta(t) + j/\pi t) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\}$$
$$s.t. \sum_{k=1}^{K} u_k = f \tag{11}$$

where K is that number of decomposition subsequences, the $\{u_k$ and ω_k are the k-th subsequence and its central frequency respectively, $\delta(t)$ is a Dirac function, ∂_t stands for derivative, f is the unknown signal, and $j^2 = -1$.

By introducing the Lagrange multiplication operator into the above formula, the constrained variational problem is transformed into an unconstrained variational problem. Calculation process reference formula (12), as shown at the bottom of the next page. In the formula (12), for reducing the interference of Gaussian noise, α is added as a secondary penalty factor and λ is Lagrangian multipliers, $f(t)$ is the real part of the original signal. By searching the saddle point of the augmented Lagrange function, the variational model iteratively optimizes u_k, ω_k and λ, and finally obtains $\{u_k$ and ω_k.

The signal is transformed by Fast Fourier Transform and variational mode decomposition so that the Intrinsic Mode Function (IMF) and the frequency domain signal, both of them are with highest energy ratio, are selected as the inputs of the DPCNN. Then, the two signals are convolved and pooled twice to extract features. Furthermore, the inner product of the two features is calculated to obtain the fusion features. Feature fusion is calculated as follows.

$$Fu_i^j = F_i^j \cdot Tf_i^j \tag{13}$$

In which Fu_i^j represents the j-th eigenvalues in the i-th feature matrix of the fusion feature, F_i^j and Tf_i^j denote the extracted frequency domain features and time-frequency domain features, respectively.

Finally, main features from the fused features are extracted again through convolution and pooling, and the results are obtained through the output layer. The structure of the Double-path Parallel Convolution Neural Network is shown in Figure 5.

2) BOA IMPROVES DPCNN

As a variant structure of the deep learning algorithm, the model in this paper exhibits significant sensitivity to network performance through the adjustment of various hyperparameters. Due to the diversity and wide range of hyperparameter settings, manual selection becomes challenging. it is difficult to achieve manual selection. Therefore, this paper optimizes the important hyperparameters of the DPCNN through BOA. The types and the optimization range of hyperparameters are shown in Table 1.

TABLE 1. BOA optimization hyperparameter types.

Hyperparameter type	Optimization range
Initial Learn Rate	[0.0001 0.1]
Max epochs	[200 500]
Dropout Rate	[0.4 0.7]
Learn Rate Drop Factor	[0.00001 0.001]
L2 Regularization	[0.0001 0.01]

The optimization process of BOA is shown in Figure 6. The author divides the data into training sets and test sets, and takes the test accuracy of the test set as the optimization objective function.

By setting the maximum number of iterations required for optimization, BOA is ensured to produce an acceptable value after each iteration. Upon reaching the maximum iteration count, the optimal target is selected from all acceptable values.

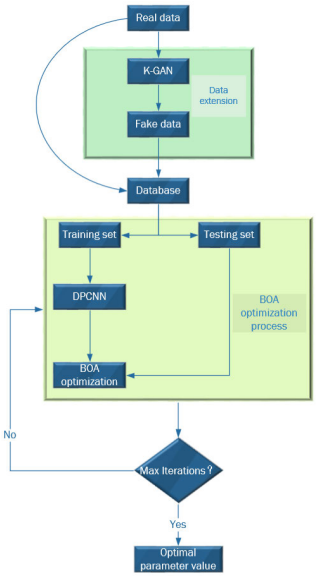


FIGURE 6. Optimization Process.

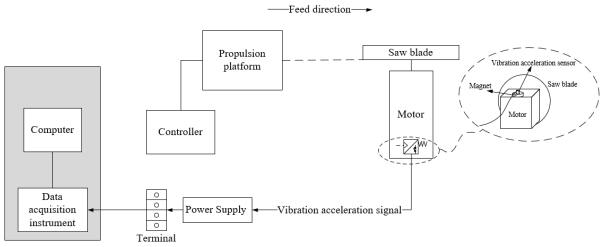


FIGURE 7. schematic layout of experimental equipment.

IV. EXPERIMENTAL VERIFICATION AND PERFORMANCE ANALYSIS

Windows 11 system and MATLAB2022b are used to build the algorithm model in this paper. The hardware facilities are AMD R7-4800H CPU, Nvidia GTX 1050Ti GPU and 16G memory.

A. SAW BLADE WEAR EXPERIMENT AND DATA INTRODUCTION

The experimental equipment for saw blade wear utilizes a tungsten steel inlaid alloy saw blade, the size of which is 405mm×60T. The saw blade is mounted on a "YM90L5-2" main spindle motor with a power of 6.5KW/380V. In the experiment, vibration data is collected by the IEPE vibration acceleration sensor and Advantech PCI-1710U data acquisition car. The collected data were saved by MATLAB software in the desktop host. The schematic layout of the experimental equipment is shown in Figure 7.

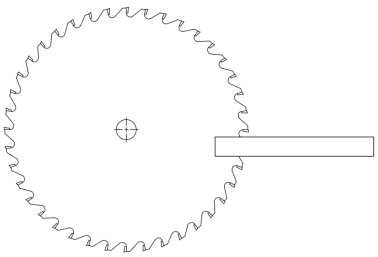


FIGURE 8. Saw blade cutting condition.

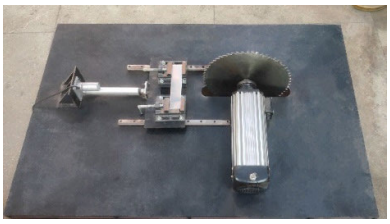


FIGURE 9. Experimental condition.

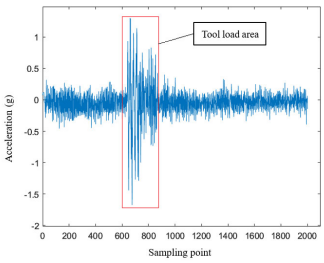


FIGURE 10. Data samples.

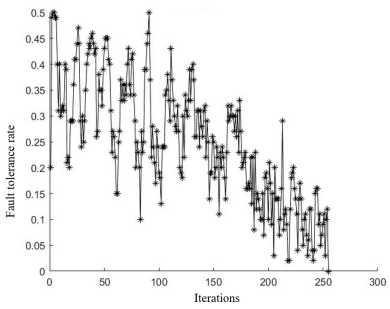


FIGURE 11. Fault tolerance rate of the K-GAN model.

The sensor's frequency response range is from 1 to 10 kHz. In order to stabilize the installation and reduce the influence of the vibration from the test bench, the sensor are

$$L(\{u_k\},\{\omega_k\},\lambda) = \alpha \sum_k \left\| \partial_t \left[(\delta(t) + j/\pi t) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| f(t) - \sum_k u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_k u_k(t) \right\rangle \tag{12}$$

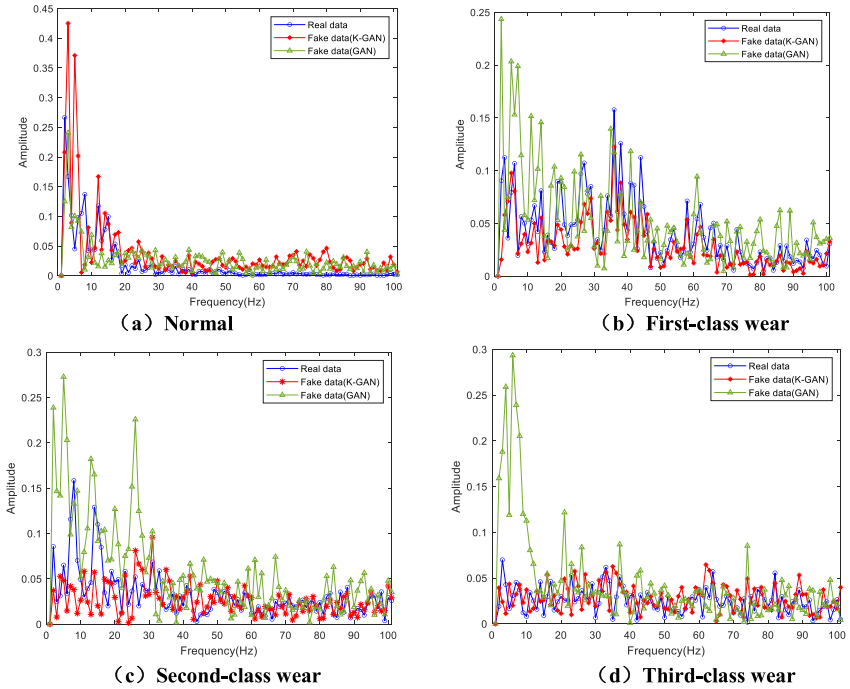


FIGURE 12. The spectrum of Real Data and Fake Data.

securely installed on the rear of the motor through a powerful magnet. MATLAB is configured with a sampling frequency of 200 Hz. The cutting edge of the circular saw blade is polished by an angle grinder and measured by a digital vernier caliper, which provides a measurement accuracy of up to 0.01 mm. Each wear amount and its label are shown in Table 2.

TABLE 2. Wear value and label of the saw blade.

Wearing value /mm	Tool state	Label
0	Normal	1
[0.1,0.2）	First-class wear	2
[0.2,0.3）	Second-class wear	3
[0.3,04）	Third-class wear	4

The wear value in the table is the wearing band of the tool flank. According to the tool wear standard, the wear value above 0.1mm is defined first-class wear, above 0.2mm as second-class wear, and above 0.3mm as third-class wear.

Firstly, a cutting experiment was conducted on the saw blade under normal conditions. The material being cut is an aluminum plate with a thickness of 4mm, and its cutting state is depicted in Figure 8. Due to the small thickness of the cutting plate, a one-time cutting approach was employed. The upper surface of the cutting plate is positioned 2mm away from the center of the saw blade, and the cutting length produced by the saw blade is 30mm.

Due to the soft cutting material and a limited number of cutting plates, the saw blade does not experience significant wear. Therefore, after conducting cutting experiments under normal conditions, the negligible wear of the saw blade in various cutting states has been deliberately ignored. The same saw blade was subsequently ground three times consecutively by an angle grinder, with each grinding removing approximately 0.1mm of material. And vibration acceleration signals have been collected in turn for the saw blade under these three levels of wear. The experimental condition is depicted in Figure 9.

TABLE 3. Cutting parameters.

Plate thickness/mm	Feed speed/(mm • s-1)	Cutter head rotation speed/(r • s-1)	Cutting length/mm
4	30	48	30

In which the cutting parameters are shown in Table 3.

The vibration acceleration signals have been collected in the above four states, and so does 50 sets of original data in each state, resulting in a total of 200 sets of data. Each data set consists of 2000 sampling points, including 200 vibration data under the tool loading condition with the loading duration 1s. The data sample is shown in Figure 10.

B. DATA EXPANSION EXPERIMENT AND ANALYSIS

In this section, the author extracts the data to be used as training data from the original experimental data, specifically from the tool loading region. These data are expanded after normalization by using the K-GAN model constructed in Section III-A. The generator and discriminator structures of the K-GAN model are shown in Table 4.

The model continuously updates the weights of the generator and discriminator, and generates data after fully

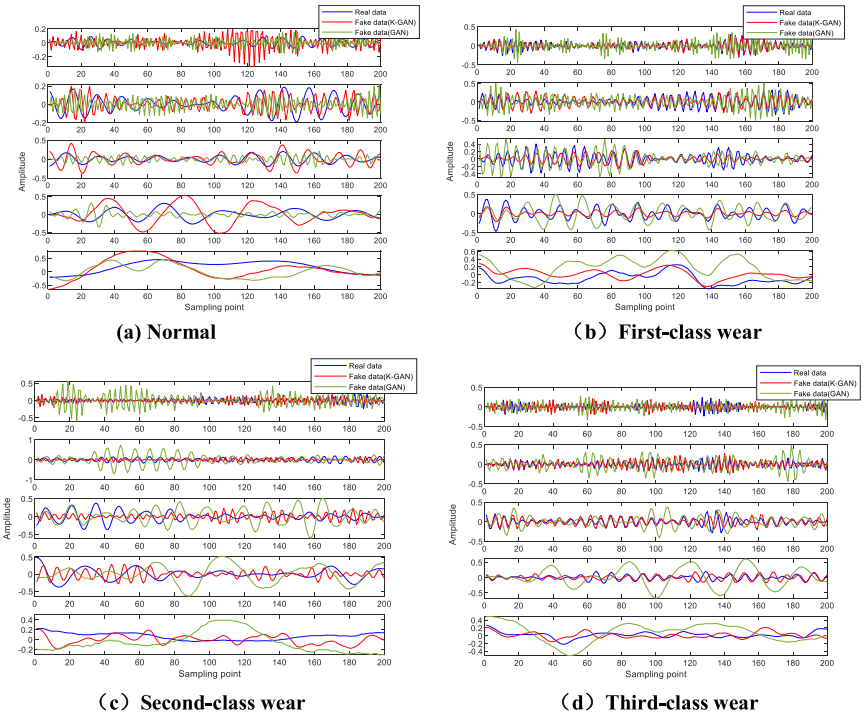


FIGURE 13. IMF of real data and fake data.

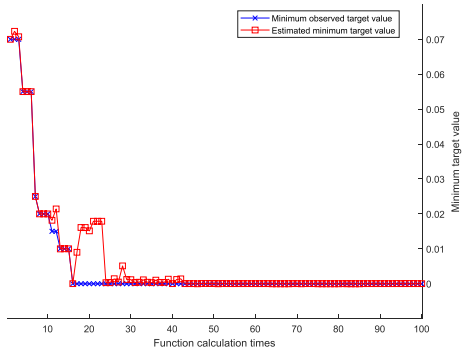


FIGURE 14. Optimization process.

TABLE 4. The generator and discriminator structures of the K-GAN model.

Generator		Discriminator	
Layer name	Number of layers	Layer name	Number of layers
Fully connected layer	4	Fully connected layer	4
Leakyrelu	4	Leakyrelu	3
		Dropout layer	3
		Sigmoid	1

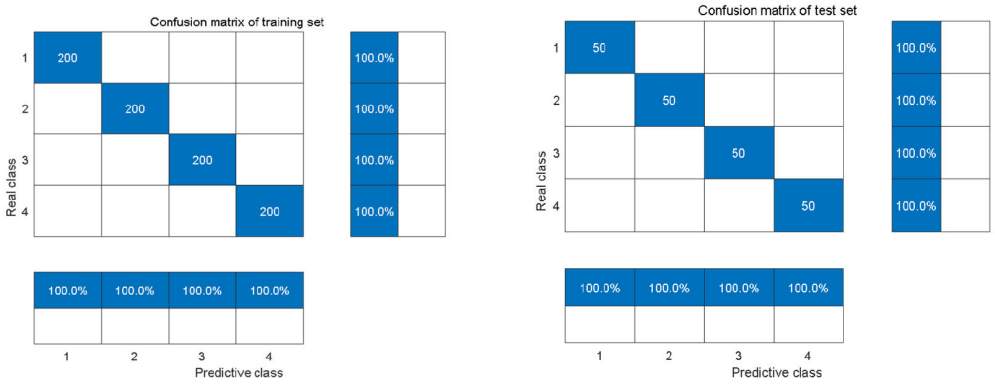
connection. Each batch of group consists of 200 sets of data, and the total data dataset comprises 1000 sets. In K-GAN model training, the change of fault tolerance rate is shown in Figure 11.

Clearly, the tolerance rate of the K-GAN model shows a decreasing trend during the training process. After approximately 260 iterations, the tolerance rate of the model drops below the set threshold of 0.001. In other words, the generated data matches the distribution of real data at this point, indicating optimal model performance. Following this, a comparison was made between the generated data and the real data through Fast Fourier Transform to obtain frequency domain signals. Additionally, a comparison was conducted between the IMF obtained from Variational Mode Decomposition, as shown in Figure 12 and Figure 13.

Upon comparative observation, the frequency distribution of the generated data in the frequency domain closely resembles that of real data. The trend of IMF is largely consistent, and the generated data lags slightly in phase and exhibits amplitude fluctuations. Considering the variations introduced by random data conforming to orthogonal distribution and the inherent distinctions between actual experimental data, this divergence falls within an acceptable tolerance range.

To further illustrate the advantages of the data generated by the optimized GAN network, The author calculated the similarity index between the data generated by the two GAN models and the real data. As shown in Table 5.

According to the criteria, a higher value of Pearson correlation coefficient indicates greater similarity between the data, while a smaller Euclidean distance indicates higher data similarity. The data presented in Table 5 demonstrates that, in all states, the similarity between the data generated by the optimized GAN model and the real data is higher than before optimization. Therefore, the comparison results show that the data generated by the K-GAN model has high similarity with the real data, which makes the generated data have reliable classification and recognition function.



（a）Confusion matrix of the training set　（b）Confusion matrix of the test set

FIGURE 15. Confusion matrix.

C. BOA OPTIMIZATION EXPERIMENT AND MODEL IDENTIFICATION PERFORMANCE ANALYSIS

In this section, the author employs 800 sets of generated data as training data and 200 sets of original data as testing data, and optimizes the model through BOA. As described in section III-B, channel 1 and channel 2 of the DPCNN model are mainly composed of two convolution layers and two pool layers. The fusion features of the model are completed by a convolution layer and a pooling layer. The BOA optimization is set with 100 iterations, leading to the acquisition of the optimal model parameters, as follows.

TABLE 5. Pearson correlation coefficient and Euclidean distance between generated data and real data.

Data type	Pearson Correlation Coefficient	Euclidean Distance
Normal data(K-GAN)	0.4392	6.7513
Normal data(GAN)	0.1383	8.4910
First-class wear data (K-GAN)	0.8814	3.0463
First-class wear data (GAN)	0.3100	8.3081
Second-class wear data(K-GAN)	0.5607	3.6787
Second-class wear data(GAN)	-0.0780	9.4421
Third-class wear data (K-GAN)	0.4156	3.3002
Third-class wear data (GAN)	0.1042	7.2652

1) OPTIMIZATION EXPERIMENT

The author uses BOA to optimize the hyperparameter of the model described in section III-B2, and the optimization process is shown in Figure 14.

The figure above shows that the observed minimum value coincides with the estimated minimum value after about 43 iterations. Furthermore, no other minimum values have been observed during the optimization process, indicating that the network has found the optimal parameters at this moment. The optimal parameter values are shown in Table 6.

TABLE 6. Optimal hyperparameters.

Hyperparameter type	The optimal value
Initial Learn Rate	0.001
Max epochs	392
Dropout Rate	0.53
Learn Rate Drop Factor	1.0163e-04
L2 Regularization	9.8 e-03

2) IDENTIFICATION PERFORMANCE ANALYSIS

The prediction and identification performance of the optimal model obtained in the previous section is analyzed through utilizing the training set and test set. Also, the comparison with other wear prediction models which use the same data has been conducted. The results are as follows.

a: MODEL PERFORMANCE ANALYSIS

The confusion matrix of the training set and test set in the optimal network prediction is constructed, as shown in Figure 15.

Obviously, both the test set and the training set have achieved high accuracy in model prediction. There are 800 groups in the training set, and all kinds of predictions have reached 100%. Furthermore, there is no over-fitting phenomenon in 200 groups of data in the test set, and the prediction rate of the model for different states also reaches 100%. Therefore, it shows that the optimized model possesses a strong classification performance.

b: COMPARATIVE ANALYSIS OF PERFORMANCE

In order to further verify the advancements of the model proposed in this paper, a comparison was made in terms of prediction accuracy between the proposed model and a CNN based on self-attention mechanism, as well as the DPCNN. The comparison result is shown in Figure 16.

The graphic data show that the average prediction accuracy of CNN based on the attention mechanism is 97.5%, that of DPCNN is 98.5%, and DPCNN optimized by BOA is 100%. The above data show that the performance of the model

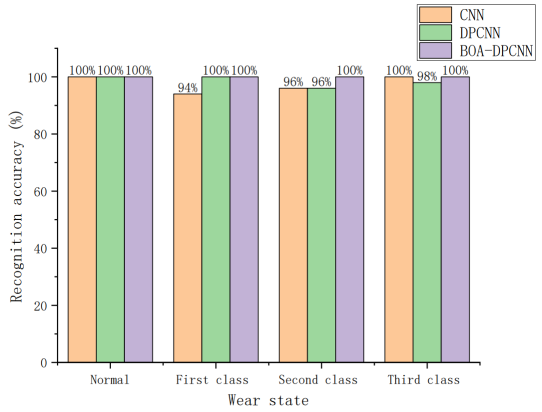


FIGURE 16. Model performance comparison.

based on feature fusion is obviously better than that of the convolutional neural network with single domain input, and BOA optimization can effectively improve the recognition accuracy of the model. After BOA optimization, the model performance is better than the other two networks in four states of saw blade wear, which shows that the double-path parallel convolution neural network based on BOA optimization can effectively solve the problem of saw blade wear identification.

V. CONCLUSION

In this paper, a method for saw blade wear state identification is proposed. According to the three different wear states of the circular saw blade in the cutting process, a custom-made experimental setup is developed to observe the effect of the wear, and the acceleration vibration signals of different wear degrees have been collected. The signal is analyzed by the established model. The main conclusions can be drawn as follows:

(1) To solve the problem of difficult data acquisition of saw blade wear, a GAN model optimized by the K-nearest neighbor algorithm is proposed. Experiments on data expansion show that this method can generate signals which has highly similar to the real data, and can effectively solve the problem in saw blade wear identification of small samples.

(2) In addition, a Double-path Parallel Convolution Neural Network is constructed for wear state identification. Fast Fourier transform and variational modal decomposition are used to transform the feature space, and the two signals are fused in the model. The performance of the model is verified in the prediction of the saw blade wear state. The results show that compared with other models, the performance of the model is obviously improved after feature fusion.

(3) For the recognition of saw blade wear state, the recognition accuracy of the model constructed in this study has reached 100% in 200 groups of data. It provides a reference for the wear prediction of other types of tools.

REFERENCES

[1] S. S. Aralikatti, K. N. Ravikumar, H. Kumar, H. S. Nayaka, and V. Sugumaran, "Comparative study on tool fault diagnosis methods using vibration signals and cutting force signals by machine learning technique," *Structural Durability Health Monitor.*, vol. 14, no. 2, pp. 127–145, 2020.

[2] W. Hou, H. Guo, L. Luo, and M. Jin, "Tool wear prediction based on domain adversarial adaptation and channel attention multiscale convolutional long short-term memory network," *J. Manuf. Processes*, vol. 84, pp. 1339–1361, Dec. 2022.

[3] R. G. Lins, B. Guerreiro, P. R. M. de Araujo, and R. Schmitt, "In-process tool wear measurement system based on image analysis for CNC drilling machines," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 8, pp. 5579–5588, Aug. 2020.

[4] R. Peng, J. Liu, X. Fu, C. Liu, and L. Zhao, "Application of machine vision method in tool wear monitoring," *Int. J. Adv. Manuf. Technol.*, vol. 116, nos. 3–4, pp. 1357–1372, Sep. 2021.

[5] P. J. Bagga, M. A. Makhesana, P. P. Darji, K. M. Patel, D. Y. Pimenov, K. Giasin, and N. Khanna, "Tool life prognostics in CNC turning of AISI 4140 steel using neural network based on computer vision," *Int. J. Adv. Manuf. Technol.*, vol. 123, nos. 9–10, pp. 3553–3570, Dec. 2022.

[6] K. Gao, X. Xu, and S. Jiao, "Measurement and prediction of wear volume of the tool in nonlinear degradation process based on multi-sensor information fusion," *Eng. Failure Anal.*, vol. 136, Jun. 2022, Art. no. 106164.

[7] Y. Zhou and W. Sun, "Tool wear condition monitoring in milling process based on current sensors," *IEEE Access*, vol. 8, pp. 95491–95502, 2020.

[8] S. Ying, Y. Sun, C. Fu, L. Lin, and S. Zhang, "Grey wolf optimization based support vector machine model for tool wear recognition in fir-tree slot broaching of aircraft turbine discs," *J. Mech. Sci. Technol.*, vol. 36, no. 12, pp. 6261–6273, Dec. 2022.

[9] Y. Ge, J. Zhang, G. Song, and K. Zhu, "An effective LSSVM-based approach for milling tool wear prediction," *Int. J. Adv. Manuf. Technol.*, vol. 126, nos. 9–10, pp. 4555–4571, Jun. 2023.

[10] P. Twardowski, J. Czyżycki, A. Felusiak-Czyryca, M. Tabaszewski, and M. Wiciak-Pikuła, "Monitoring and forecasting of tool wear based on measurements of vibration accelerations during cast iron milling," *J. Manuf. Processes*, vol. 95, pp. 342–350, Jun. 2023.

[11] W. Dong, X. Xiong, Y. Ma, and X. Yue, "Woodworking tool wear condition monitoring during milling based on power signals and a particle swarm optimization-back propagation neural network," *Appl. Sci.*, vol. 11, no. 19, p. 9026, Sep. 2021.

[12] Y. Yang, X. Zhao, and L. Zhao, "Research on asymmetrical edge tool wear prediction in milling TC4 titanium alloy using deep learning," *Measurement*, vol. 203, Nov. 2022, Art. no. 111814.

[13] X. Zhou, T. Yu, G. Wang, R. Guo, Y. Fu, Y. Sun, and M. Chen, "Tool wear classification based on convolutional neural network and time series images during high precision turning of copper," *Wear*, vol. 522, Jun. 2023, Art. no. 204692.

[14] G. Kucukyildiz and H. G. Demir, "A multistage cutting tool fault diagnosis algorithm for the involute form cutter using cutting force and vibration signals spectrum imaging and convolutional neural networks," *Arabian J. Sci. Eng.*, vol. 46, no. 12, pp. 11819–11833, Dec. 2021.

[15] J. Wu, Y. Su, Y. Cheng, X. Shao, C. Deng, and C. Liu, "Multi-sensor information fusion for remaining useful life prediction of machining tools by adaptive network based fuzzy inference system," *Appl. Soft Comput.*, vol. 68, pp. 13–23, Jul. 2018.

[16] G. Zheng, W. Sun, H. Zhang, Y. Zhou, and C. Gao, "Tool wear condition monitoring in milling process based on data fusion enhanced long short-term memory network under different cutting conditions," *Eksploatacja Niezawodność-Maintenance Rel.*, vol. 23, no. 4, pp. 612–618, Dec. 2021.

[17] Y. Zhou, G. Zhi, W. Chen, Q. Qian, D. He, B. Sun, and W. Sun, "A new tool wear condition monitoring method based on deep learning under small samples," *Measurement*, vol. 189, Feb. 2022, Art. no. 110622.

[18] T. Pan, J. Chen, T. Zhang, S. Liu, S. He, and H. Lv, "Generative adversarial network in mechanical fault diagnosis under small sample: A systematic review on applications and future perspectives," *ISA Trans.*, vol. 128, pp. 1–10, Sep. 2022.

[19] Q. Zhu, B. Sun, Y. Zhou, W. Sun, and J. Xiang, "Sample augmentation for intelligent milling tool wear condition monitoring using numerical simulation and generative adversarial network," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, 2021.
[20] D. A. Molitor, C. Kubik, M. Becker, R. H. Hetfleisch, F. Lyu, and P. Groche, "Towards high-performance deep learning models in tool wear classification with generative adversarial networks," *J. Mater. Process. Technol.*, vol. 302, Apr. 2022, Art. no. 117484.
[21] M. Shah, V. Vakharia, R. Chaudhari, J. Vora, D. Y. Pimenov, and K. Giasin, "Tool wear prediction in face milling of stainless steel using singular generative adversarial network and LSTM deep learning models," *Int. J. Adv. Manuf. Technol.*, vol. 121, nos. 1–2, pp. 723–736, 2022.
[22] Q. Xu, G. Huang, Y. Yuan, C. Guo, Y. Sun, F. Wu, and K. Weinberger, "An empirical study on evaluation metrics of generative adversarial networks," 2018, *arXiv:1806.07755*.

CHENGCHAO WANG received the B.Sc. degree in engineering from Jishou University, Hunan, China, in 2021. He is currently pursuing the M.Sc. degree with the College of Mechanical Engineering, University of South China, Hunan. His research interests include fault diagnosis and vibration control.

XIANGJIANG WANG received the B.Sc. degree in engineering from the Harbin Institute of Technology, Heilongjiang, China, the M.Sc. degree in mining machinery from the University of South China, Hunan, China, and the Ph.D. degree in mechanical design from Southeast University, Jiangsu, China. He is currently a Professor with the University of South China, presiding over a number of scientific research projects and receiving support from the National Natural Science Foundation of China and the Natural Science Foundation of Hunan Province, China. His research interests include intelligent mechanism and technology, mechatronics technology, and numerical control technology.

CHAO ZENG received the B.Sc. degree in engineering and the M.Sc. degree in mechanical engineering from the University of South China, Hunan, China, in 2014 and 2018, respectively, where he is currently pursuing the Ph.D. degree with the School of Nuclear Science and Technology. His research interest includes equipment automation.