Article citation info:
Hu C, Zhao Y. Mechanical fault diagnosis system based on genetic algorithm optimization and multi-sensor data fusion. Diagnostyka. 2025;26(4):2025410. https://doi.org/10.29354/diag/215723.

DIAGNOSTYKA, 2025, Vol. 26, No. 4

e-ISSN 2449-5220
DOI: 10.29354/diag/215723

MECHANICAL FAULT DIAGNOSIS SYSTEM BASED ON GENETIC ALGORITHM OPTIMIZATION AND MULTI-SENSOR DATA FUSION

Caodi HU [1,*], Yumeng ZHAO [2]

[1] College of Electromechanical Engineering, Jiaozuo University, Jiaozuo, Henan, 454000, China
[2] College of Artificial Intelligence, Jiaozuo University, Jiaozuo, Henan, 454000, China
Corresponding author's E-mail address: caodi_hu@hotmail.com

Abstract

Industrial machinery frequently experiences machine failures during operation. Promptly diagnosing these failures can improve industrial operational efficiency to a certain extent. Existing research has many shortcomings. This paper considers two issues: data noise and the weight of multi-sensor data fusion, and designs a Genetic Algorithm-based Multi-Sensor Data Fusion Diagnosis (GA-MSDFD) method. This method first uses manually defined indicators to filter noise and uses a custom method to eliminate it. Feature extraction is then performed on the evolved data, and a genetic algorithm is used for multi-objective feature selection. This algorithm has inherent advantages over other machinery because its design considers the impact of noisy data. Experimental results show that our model achieves a fault diagnosis accuracy of 96.8%, far exceeding several other machinery models. The model also far outperforms these models in noise robustness, noise resistance, and convergence performance. The proposed model is of great significance for the maintenance of industrial machinery operations.

Keywords: machine failure, fault diagnosis, multi-sensor data fusion, genetic algorithm, optimization

1. INTRODUCTION

With the further development of the industrial age, the manufacturing industry has continuously increased its requirements for the stability of industrial machines, the efficiency of fault diagnosis, and the speed of processing. The causes of machine failure are diverse. On the one hand, they are due to production problems such as material quality and design defects. On the other hand, there are also failures caused by long-term use of the machine during operation, including wear and breakage of the load-bearing shaft and gears. Therefore, an important goal of industrial development at this stage is to establish a low-error, high-response machine fault diagnosis system [1].

Traditional detection methods rely on human experience. Technicians with extensive fault diagnosis experience use comprehensive analysis to diagnose machine faults by listening, touching, and observing, as well as by reviewing machine maintenance records. This method has the advantages of low cost and high accuracy, but its disadvantages are that it cannot monitor machine faults in real time and handle them in a timely manner, and it is often highly dependent on experienced human experts [2]. At the same time, there are also detection methods that obtain detection results based on the analysis of physical signals. These physical signals include vibration signals, spectrum, temperature, etc. These physical signals come from physical sensors, and how to fuse these physical signals is a major problem.

Current research has used genetic algorithms to improve fault diagnosis algorithms. These approaches can be divided into three main areas: (1) Using genetic algorithms to optimize features extracted from sensor data and select effective features to improve the accuracy of multi-sensor machine fault diagnosis. (2) Using genetic algorithms to determine the weights of multi-sensor fusion to determine the importance of information provided by different sensors. (3) Using genetic algorithms to optimize parameters in neural networks to improve the accuracy of classifiers. This paper proposes an improved multi-sensor data fusion method based on the second approach.

GA is an intelligent optimization method that mimics the process of biological evolution. It treats possible solutions as "individuals" and iterates through "survival of the fittest" (selection), "gene exchange" (crossover), and "random mutation" (variation) to gradually find a better answer.

Traditional weight determination is mainly based on rules, through equal weight averaging or maintenance personnel determining weights based

Received 2025-08-08; Accepted 2025-05-17; Available online 2025-05-18

on experience. The main reason is that the operating state of the equipment is dynamically changing. Equal weight or experience-based methods are based on the weight value determined at a certain moment and cannot be perfectly adapted to other states. In order to improve the accuracy of machine fault detection using multi-sensor data fusion methods in complex industrial environments, this paper proposes a fusion method based on weight distribution optimization. The core idea of this method is to evaluate the noise level of each sensor data after obtaining data in real time, and on the other hand, combine human knowledge to determine the weight matrix of each data source and data type [3].

The model in this paper mainly consists of the following modules: (1) Data collection center that integrates multiple IoT sensors. (2) Data noise assessment module: evaluates the data types and noise levels of different sensors. (3) Data purification module: removes noisy data. (4) Genetic algorithm weight optimization module: derives the final weight matrix based on the purified data. (5) Machine fault analysis module: derives the final analysis results based on the results of the above five modules.

The significance of this study lies in the proposed adaptive multi-sensor weighting method, which promotes the transformation of data weighting from "static weighting" to "dynamic weighting." At the same time, considering data noise and real-time weighting, this further improves the accuracy, reliability, and robustness of machine fault analysis results.

2. RELATED WORKS

2.1. MSDA

Multi-sensor data fusion, abbreviated as MSDF (Multi-Sensor Data Fusion), can be defined as the merging, processing, and analysis of data from multiple sources to achieve the perception of the real state of physical entities. Hu et al. [4] applied the weighted statistical and filtering-based fusion method to the military and compared the results. They pointed out that the weighted average method is only applicable when the data generated by each sensor is independent and identically distributed. If there are errors in the data or the data is complex nonlinear data, there will be a huge deviation from the actual situation. Hua, L et al. [5] applied the Kalman filter (KF) to the field of machine fault detection and pointed out that the Kalman filter can be used as a data fusion device for linear data . It transforms nonlinear data in a certain way to obtain linear data. According to the idea of transformation, two new methods, EKF and UKF, were identified. Mo, C et al. [6] applied the improved Kalman filter (KF) method to a real machine fault detection scenario and conducted online tests to preliminarily diagnose early faults. In addition, some filter methods based on random sampling have gradually emerged. Ren, B et al. [7] applied it to nonlinear data that does not obey the Gaussian distribution and achieved good results. The core of this method is that the Monte Carlo method is used to simulate in the sampling space, thereby approximating the true probability distribution. Shi, G et al. [8] used a similar method to fuse various types of data in the hydraulic system, specifically the fusion of pressure, flow, temperature and other signals. The core of this method is to use the particle swarm algorithm to implement an adaptive method, which improves the accuracy of machine fault analysis in the hydraulic system. However, subsequent studies have pointed out that its computational space complexity is too high and there is an example degradation problem, which makes it unsuitable for real-time systems.

2.2. Application of Genetic Algorithms in Machine Fault Detection

Genetic Algorithm (GA) algorithm is widely used to optimize the hyperparameters of other intelligent algorithms. For example, Shi, R et al. [9] used support vector machines to analyze multi-sensor data to perform classification analysis of load-bearing shaft faults, and used GA(Genetic Algorithm) to select the kernel function parameters and penalty factors of SVM, which improved the accuracy of hyperparameter selection by 5% compared with the traditional grid search. Ugli, O et al. [10] used GA for feature selection, specifically to select the most representative feature subsets from high-latitude vibration features and use these subsets for fault analysis. The advantage of this method is that it improves the discrimination accuracy and has low time complexity. This research performs binary selection (select or not select) on manually extracted static features, while this paper performs continuous weight optimization on high-dimensional temporal deep features automatically extracted by RNN. The chromosome encoding uses real-valued weights instead of 0/1 masks. This paper jointly optimizes GA with deep learning models to achieve end-to-end collaboration between feature extraction and selection, GA runs independently of the classifier. This paper is geared towards multi-sensor dynamic fault diagnosis, emphasizing temporal modeling and noise robustness, while this paper focuses on single-source data under static working conditions. Therefore, the method in this paper is more adaptive and generalizable.

Yuan, Y et al. [11] designed a weighted average fusion strategy based on GA, taking the weight of each sensor as a "chromosome" and the final classification effect as the fitness function. The weight of each calculation is obtained by adjusting all previous results. However, it does not consider the data noise problem or the reliability of the data source. Experimental results show that the classification accuracy of this method is 8.3% higher than that of equal weight. Some studies have also used GA to optimize the structure of neural networks. For example, GA-DNN uses GA to optimize the selection of hidden layer nodes and the

adjustment of hyperparameters in deep networks. The results of this method show that this method has high generalization ability.

2.3. Multi-sensor data

Currently, single sensors are rarely used in the industrial world because multi-sensor data can provide more comprehensive information and is more helpful for fault detection in complex machine systems. However, the application of multi-sensor data still faces many challenges. First, due to differences in the settings of different sensors, the dimensions, location information, and sampling frequency of their data are different, making their alignment very difficult [12]. Second, due to the large number of sensors, the data they generate increases exponentially, and there is a large amount of duplicate data, invalid data, and interference data. The existence of these interference data greatly interferes with the accuracy of machine fault prediction. Finally, the information entropy of multi-sensor data is high, and its information entropy is determined by many factors, such as noise, missing data, and simultaneous sensor equipment failure. Therefore, the multi-sensor fusion algorithm must take into account data alignment and interference data [13].

2.4. Research Motivation

Genetic algorithms have been extensively studied in the field of multi-sensor data fusion for machine fault diagnosis, and many have achieved promising results. However, challenges remain, such as low diagnostic accuracy and a lack of data de-noising. Most studies in this area assume that the data provided by multiple sensors is reliable and clean, failing to consider the large amount of noisy and missing data generated under high-intensity industrial operating conditions. For example, study A directly extracts features from raw data using a CNN. These extracted features are then used directly for fault analysis without considering the weighting of different features. During the evaluation phase, the authors demonstrated that the presence of a large amount of redundant data resulted in low feature validity, which in turn affected accuracy.

In this field, genetic algorithms are primarily used for feature selection or optimizing classifier hyperparameters, but rarely for determining multi-sensor weights. Many models fail to incorporate specific device knowledge, such as sensor type, sensor location, and the physical meaning of connection parameters, during design. Incorporating this specific domain knowledge can improve model interpretability.

Therefore, this paper proposes a multi-sensor fault analysis and diagnosis method that combines data noise processing and dynamic fusion weighting. This paper introduces data noise processing and data purification to remove noise from the data, and uses an RNN to extract data features. The extracted features are then dynamically weighted using a genetic algorithm. Human experience, such as the more agile response of vibration sensors to machine impacts, is incorporated into the design of the genetic algorithm's fitness function. The core of this method lies in data quality perception and dynamic feature weight iteration. The proposed method aims to improve the accuracy of machine faults using multi-sensor data fusion methods in highly complex and nonlinear industrial environments, and is dedicated to building an intelligent, highly accurate, and robust machine fault diagnosis system.

3. MATERIALS AND METHODS

To reduce the impact of randomness in GA, all experiments are reported as the mean accuracy ± standard deviation of 10 independent runs. An elitism strategy and a fixed random seed are employed to ensure reproducibility.

GA is executed once during the offline training phase to generate the optimal feature subset and corresponding weights. This subset is used consistently during the online diagnostic phase and does not require real-time updates. If it is necessary to adapt to new working conditions, GA can be retrained periodically (e.g., monthly) using newly collected data.

3.1. System Architecture

The model architecture of this paper is shown in Figure 1.

The flowchart of the proposed method is shown. The human-shaped icon represents the initial feature set pre-selected by domain experts based on engineering experience (an optional step), and subsequent feature selection is completed automatically by a genetic algorithm.

First, real-time sensor data is acquired through sensors distributed at different locations on the machine. This data is fed into the data evaluation module for noise assessment and purification. The evolved data is then fed into a recurrent neural network for feature extraction. The diverse features output by the recurrent neural network are evaluated using a genetic algorithm to select the most important feature subset. This feature subset and its weights are then fed into a classifier for classification, ultimately yielding the results of the fault diagnosis analysis.

The model mainly consists of four blocks. The first layer is the data acquisition layer. It is responsible for acquiring real-time data from multiple sensors, including spectrum pressure, temperature, vibration and other types of real-time data . The second layer is the data evaluation layer. The second layer includes two steps: noise evaluation of the data and removal of noisy data. The third layer is the feature selection layer. The data output by the second layer is extracted through RNN, and the genetic algorithm is used to screen the features and obtain the weights.

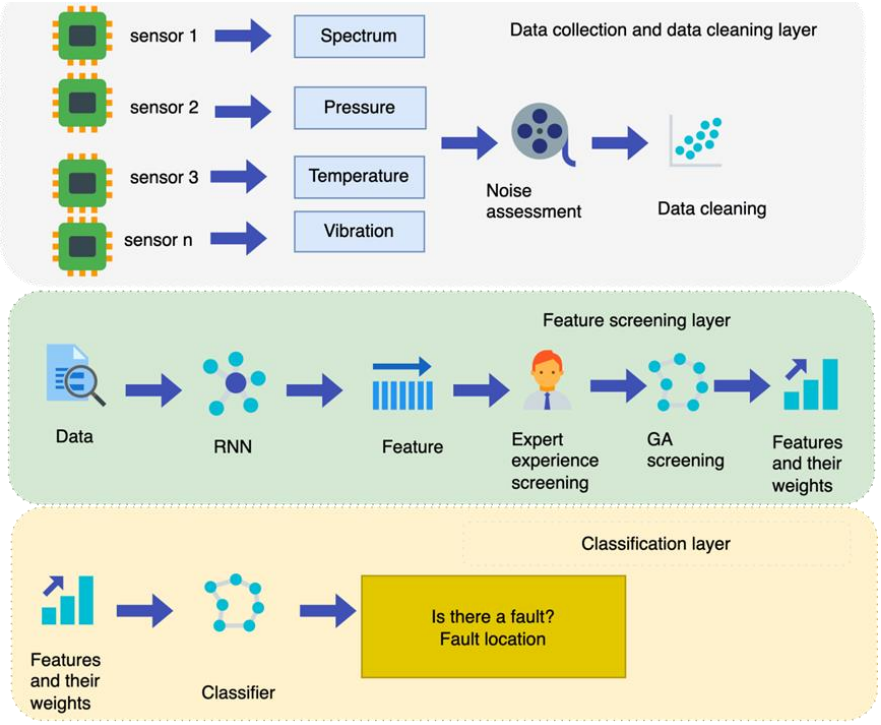

Fig. 1. Model architecture

The core features of the system proposed in this paper are dynamic and highly adaptable [14]. The weight matrix output by the third layer feature selection layer is dynamically updated. The third layer contains a manually defined rule database based on the experience of human experts. Before using the genetic algorithm for feature selection, the rules defined by human experts are used for preliminary screening. This can further improve the reliability and interpretability of the diagnosis results and help maintenance workers to repair the machine.

3.2. Data Noise Assessment Module

This model defines a series of metrics to quantify the noise problem in multi-sensor data. Since different sensors have different data types and different types of noise, we need to develop different evaluation metrics for different types of sensors.

Vibration sensors are primarily responsible for collecting periodic mechanical vibrations. In vibration sensors, high-frequency impact sounds may be mistaken for periodic vibration sounds. We use the CF indicator as an indicator for evaluating vibration sensor data. This indicator reflects the change in noise impact intensity of the vibration signal, as shown in Formula 1 [15].

$$CF = \frac{max|x_i|}{\sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}} \tag{1}$$

Where, is the maximum value of the signal and the denominator is RMS.

When a machine component fails, such as experiencing damaging wear and cracks, the vibration signal will show high-frequency peaks, and the CF value will be significantly higher than the normal range (2-3).

The K indicator is used to represent the change of the probability density function corresponding to the signal, as shown in Formula 2 [16].

$$K = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i-\mu}{\sigma}\right)^4 \tag{2}$$

Where is the mean of the vibration signal and is its standard deviation. When the data follows a normal distribution, K=3. When there is obvious noise in the data, such as gear meshing impact and bearing ball peeling, K will be greater than 3 [17].

The normal temperature change of the temperature sensor is slow, while the abnormal data change is manifested as an outlier. For the temperature sensor data, we use the TCE (Temperature Change Error)value to evaluate the noise. This indicator reflects the degree of deviation between the outlier point and the overall trend of the data, specifically as Equation 3 [18].

$$TCE = \frac{1}{N}\sum_{i=1}^{N}|x_i - \hat{x}_i| \,/\, max(x_i) \tag{3}$$

The figure shows a trend line when the sliding window is 30. When the temperature sensor is disturbed by the environment and anomalies occur, the TCE will increase.

The sensor data output by the pressure sensor consists of two parts: a dynamic fluctuation curve and a static baseline. If there is noise in the data, it will cause baseline drift. Therefore, we define an indicator to measure the degree of baseline drift, as shown in Formula 4 [19].

$$BDR = \frac{|\bar{x}_{\text{thelatter50\%}} - \bar{x}_{\text{thefirst50\%}}|}{\bar{x}_{\text{all}}} \tag{4}$$

In the formula, and are interpreted as the average of the first half and the last half of the data, and is the mean of all data [20].

Table 1. Noise level determination rules

Noise level	(Signal-to-noise ratio)SNR	(Coefficient of variation)CV	Wavelet Entropy
Class 1 (noiseless)	SNR≥30dB	CV≤0.1	Wavelet entropy ≤ 0.2
Level 2 (mild noise)	20dB≤SNR<30dB	0.1<CV≤0.3	0.2 < wavelet entropy ≤ 0.4
Level 3 (medium noise)	10dB≤SNR<20dB	0.3<CV≤0.5	0.4 < wavelet entropy ≤ 0.6
Level 4 (severe noise)	0dB≤SNR<10dB	0.5<CV≤0.7	0.6 < Wavelet entropy ≤ 0.8
Level 5 (extreme noise)	SNR<0dB	CV>0.7	Wavelet entropy > 0.8

Bubble interference is a common type of interference in pressure sensors. When this interference occurs, the Temperature Change Error (BDR) value suddenly increases, which can serve as a warning of noise in the data. We also define the SNR(Signal-to-Noise Ratio) as a universal reference metric. The SNR measures the degree of noise in the data and is applicable to all types of sensors. Its calculation formula is shown in Equation 5.

$$SNR = 10\,log_{10}\left(\frac{\sum_{i=1}^{N} x_i^2}{\sum_{i=1}^{N}(x_i-\hat{x}_i)^2}\right) \quad (5)$$

In Formula 5, the numerator is the sum of the effective signal energy, and the denominator is the sum of the noise energy. The output of the data noise assessment module is a noise level assessment matrix, which outputs the noise level of each sensor. For data segment a from sensor A, we define its noise level as. If data segment a satisfies the evaluation criteria, it can be considered noise-free. If any of the following conditions are met, or, it is considered noise-free. The specific noise level determination rules are shown in Table 1.

3.3. Data purification module

The purpose of the data purification module is to filter noisy data based on the noise level information provided by the data noise assessment module. We use differentiated approaches based on the specific noise level to maximize noise removal from multi-sensor data. In real industrial environments, the noise generated is complex and diverse, such as high-frequency impact, baseline drift, and random fluctuations, and the noise levels and data types vary. Therefore, we need to dynamically select purification methods based on the specific situation to avoid residual noise caused by a single treatment.

Level 1 noise processing. Level 1 noise is noise-free. Its primary purpose is to remove trace amounts of random noise that may be present in the data. To maximize the preservation of the original signal, we employ a moving average filter with a window size of 5. Specifically, for a 1 Hz signal, we retain a 5 ms time window. The calculation formula is Equation 6.

$$x_{\text{filtered}}(t) = \frac{1}{5}\sum_{i=0}^{4} x\,(t-i) \quad (6)$$

The principle of this method is to smooth the data by averaging, thereby suppressing noise. It can remove the noise in the signal without destroying the original signal. This method only requires a small amount of time complexity and is very fast, and can achieve real-time to a certain extent.

Level 2 noise is typically a glitch signal generated by point radiation, typically representing high-frequency interference. To address this issue, we first used the db4 wavelet basis to perform a three-layer wavelet decomposition of the data signal, obtaining wavelet coefficients for different frequency bands. We then set a threshold and applied soft thresholding to each wavelet coefficient. The specific formula is shown in Equation 7.

$$\lambda = \sigma\sqrt{2\,log\,N} \quad (7)$$

In this formula, is the standard deviation of a data segment for a particular sensor. This standard deviation is estimated by the median of the higher-frequency coefficients after DB4 decomposition. is the effective length of the signal. The processed data is reconstructed using wavelet transforms to obtain the final purified data. This method can separate noise from the effective signal in high-frequency data.

For data with a medium noise level, level 3, the noise exhibits highly nonlinear, time-varying characteristics. This type of noise may be pulsating noise, for example. We use an adaptive Kalman filter to handle this type of noise. Our new Kalman filter uses real-time noise observations and dynamically adjusts based on the variance. The specific formula is Equation 8.

$$\widehat{\mathrm{R}}_{\mathrm{k}} = (1-\alpha)\widehat{\mathrm{R}}_{\mathrm{k-1}} + \alpha(\mathrm{z_k} - \mathrm{H}\hat{\mathrm{x}}_{\mathrm{k}})^2 \quad (8)$$

In the formula, is the covariance update coefficient, is the specific observed value, is the observation matrix, and is the evaluation value of the state.

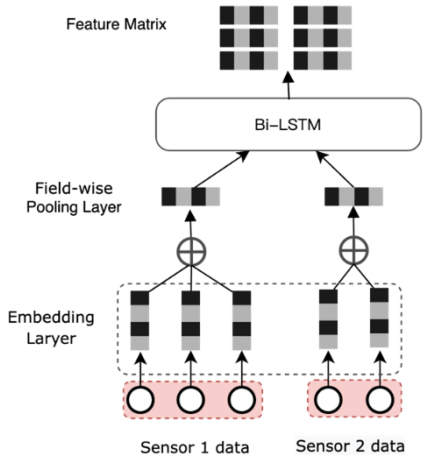

Fig. 2. Framework for feature extraction using Bi-LSTM

For the fourth level of noise, this type of noise is very serious, and this noise often means that the overall structure of the data has been seriously damaged. This type of noise may be a

long period of signal loss caused by poor sensor contact, or signal jump. To handle this type of noise, it is necessary to combine the relevant methods of anomaly detection in deep learning. In the model, our approach is to first use the random forest algorithm to identify anomalies in the data. We judge the degree of abnormality by calculating the sum of the path lengths of abnormal samples. We define a threshold, which is obtained through training of normal samples. If the calculated sum of the path lengths of the abnormal samples is greater than the threshold, the data point is judged to be severe noise, and we use linear interpolation to replace the abnormal data points with normal data points. The specific formula is shown in Formula 9.

$$\mathrm{x}_{\mathrm{repair}}(\mathrm{t}) = \mathrm{x}(\mathrm{t}-1) + \frac{\mathrm{x}(\mathrm{t}+1) - \mathrm{x}(\mathrm{t}-1)}{\mathrm{t}+1-(\mathrm{t}-1)}(\mathrm{t}-(\mathrm{t}-1)) \tag{9}$$

In Formula 9, we use the trend continuation method to reconstruct the sampling points with a length less than 10, instead of the linear interpolation method. This ensures that the signal is smooth to a certain extent.

Level 5 noise data, where the proportion of noise exceeds the proportion of normal signals, is often caused by sensor failure. To address this type of noise, redundancy mechanisms must be activated, replacing the abnormal data with redundant backup data. Alternatively, data from sensors of the same type and time period can be used, aligned using timestamps, to replace lost data caused by sensor failure. If backup data is unavailable, we use the LSTM deep learning model to leverage data from 100 or more past time points to predict the current time point.

The LSTM architecture used in this paper consists of two stacked hidden layers, each with 128 memory units, followed by a Dropout layer (dropout rate of 0.5) and a fully connected output layer, mapping the temporal input to a 64-dimensional feature vector. The input is a single-sensor vibration signal segment of length 2048.

The cleaned data requires rigorous quality assessment. We recalculate the signal-to-noise ratio to evaluate the results of data evolution. The specific formula is formula 10.

$$PSNR = 10\,log_{10}\left(\frac{(2^B-1)^2}{MSE}\right) \tag{10}$$

In this formula, represents the number of bits in the data. represents the mean squared error between the noisy original data and the reference data. We define a threshold of 25dB. If the Peak Signal-to-Noise Ratio (PSNR) of the cleaned data is greater than 25, the data meets the requirement. Otherwise, the data is discarded.

3.4. Feature extraction and selection

As shown in Figure 2, the cleansed data is characterized by high-dimensional, low-noise time series data. To improve machine fault diagnosis, we propose a feature screening method. This method requires a neural network to extract features from massive multi-sensor data. This paper uses a Bi-LSTM neural network architecture to extract data features.

Assume there is a sensor with a time window size of T. The data collected by a sensor within a time window can be recorded as Formula 11. Then, K sensors can be spliced into a matrix, as shown in Formula 12. The input of the Bi-LSTM receives the matrix shown in Formula 12.

$$\boldsymbol{x}^{(k)} = [x_{t-T+1}^{(k)}, \ldots, x_t^{(k)}] \in \mathbb{R}^T \tag{11}$$

$$\boldsymbol{X} \in \mathbb{R}^{T\times K} \tag{12}$$

Although previous studies have shown that features extracted by Bi-LSTMs have high discriminability for classification tasks, they can still be affected by redundant and noisy features. Therefore, this paper introduces a genetic algorithm to select a subset of features from the Bi-LSTM output and obtain their weights. We define the Bi-LSTM output features as Equation 13, and the weighted features as Equation 14.

$$\boldsymbol{w} = [w_1, w_2, \ldots, w_{2D}]^{\top} \in [0,1]^{2D} \tag{13}$$

$$\boldsymbol{f}_{\mathrm{weighted}} = \boldsymbol{f} \odot \boldsymbol{w} \tag{14}$$

In Formula 13 and Formula 14, each weight data is regarded as a chromosome in the GA algorithm, and the total population is defined as N. In this model, we use real number coding.

Each weight vector is considered a chromosome in the GA, and the total population size is defined as Np, which contains Np candidate solutions.

In terms of fitness function, we use a multi-objective fitness function and our classifier is SVM(Support Vector Machine). The specific fitness function is shown in Formula 15.

$$F(\boldsymbol{w}) = \alpha \cdot A(\boldsymbol{w}) + \beta \cdot S(\boldsymbol{w}) + \gamma \cdot R_e(\boldsymbol{w}) \tag{15}$$

Formula 15 represents the accuracy of machine fault diagnosis using SVM classification. This accuracy was validated using a 5-fold cross-validation. The specific formula is shown in Formula 16. This section, which measures feature sparsity and encourages the model to avoid selecting redundant features, is shown in Formula 17. Here , is the threshold and is the indicator function. The proposed features must satisfy predefined expert rules, as shown in Formula 18.

$$A(\boldsymbol{w}) = \frac{1}{5}\sum_{i=1}^{5}\frac{TP_i+TN_i}{TP_i+TN_i+FP_i+FN_i} \tag{16}$$

$$S(\boldsymbol{w}) = 1 - \frac{\|\boldsymbol{w}\|_0}{2D}, \quad \|\boldsymbol{w}\|_0 = \sum_{j=1}^{2D}\mathbb{I}\,(w_j > \tau) \tag{17}$$

$$R_e(\boldsymbol{w}) = \sum_{r=1}^{R}\lambda_r \cdot \mathbb{I}(\mathrm{Rule}_r \text{ satisfied}) \tag{18}$$

In addition, to prevent algorithm instability, the system re-optimizes the GA at regular time intervals to ensure real-time updates of weights, as shown in Formula 19.

$$\boldsymbol{w}^*(t) = \underset{\boldsymbol{w}}{argmax}\ F(\boldsymbol{w}; \mathcal{D}_{[t-N_u,t]}) \tag{19}$$

4. RESULTS AND DISCUSSION

In the experimental evaluation section, we designed several experiments to evaluate the proposed model from different perspectives. These

experiments focused on four dimensions: accuracy, noise resilience, convergence performance, and interpretability.

4.1. Experimental Setup

The experimental dataset used a public dataset from the Case Western Reserve University Bearing Data Center. This dataset records signal data for four types of faults under different load conditions and speeds. The data primarily consists of vibration signals. This type of dataset is essentially a four-category classification problem, including normal, inner race fault, outer race fault, and rolling element fault. To make the experimental data more realistic, we built a custom experimental platform to collect data such as temperature and current signals, which we combined with the original public data to form a multi-source dataset. We also added a new type of fault called a composite fault. The time window size set for each data sampling was 1000, resulting in a total of 5000 data samples.

To compare this study with the latest experimental results, we selected the five most powerful typical baselines in this field. They are: EW performs equal weighted averaging of all sensor data. PCA directly cools the extracted features and then performs classification. DNN directly extracts features from the data [21] and then directly classifies. CNN-LSTM combines CNN and LSTM to extract features and finally performs classification [22]. GA-SVM only uses genetic algorithms to optimize the hyperparameter selection of SVM and perform classification.

In this study, we used typical classification metrics such as accuracy, F1, and AUC for analysis. We also employed convergence algebra to compare the convergence speed of different models. Signal-to-noise ratio robustness was used to analyze the model's ability to handle varying levels of noise. This was achieved by artificially adding noise to a public dataset.

To ensure the reliability of the experiments, all our experiments were conducted in Python 3.8, and the genetic algorithm was implemented based on the DEAP library(https://github.com/DEAP). To ensure the reliability of the experimental results, we repeated each set of experimental results 10 times.

LSTM encodes the time series input from each sensor and outputs a 64-dimensional feature vector. If the system contains N sensors, the initial feature dimension after concatenation is $64 \times N$. In the experimental setup, GA ultimately retains 30 key features for subsequent fault classification. The number of features output by RNN depends on the number of sensors, while the number of features after GA selection is a fixed value of 30 dimensions.

4.2. Experimental Results

As shown in Figure 3, we use a heat map to illustrate the weight distribution of Bi-LSTM output features. Features in red areas are of high importance, while features in blue areas are of low importance. From the weight distribution of Bi-LSTM output features, we can identify features with high importance and their corresponding weight scores for subsequent classification in machine fault diagnosis. The figure shows that some features have a significant correlation with a specific fault, while other features have very low weights across all faults. This means they contribute little to fault diagnosis and are therefore considered unimportant data. These features often appear near white in the heat map.

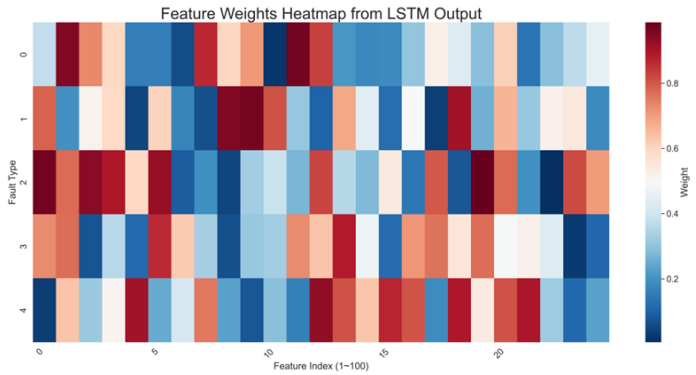

Fig. 3. Visualization of Bi-LSTM output feature importance

Table 2. Machine fault diagnosis results of 6 methods

method	Accuracy (%)	F1-score	AUC
EW	86.3	0.851	0.912
PCA + SVM	89.7	0.886	0.935
DNN	91.2	0.903	0.948
CNN-LSTM	93.5	0.928	0.963
GA-SVM	92.1	0.916	0.954
GA-MSDFD (Ours)	96.8	0.962	0.987

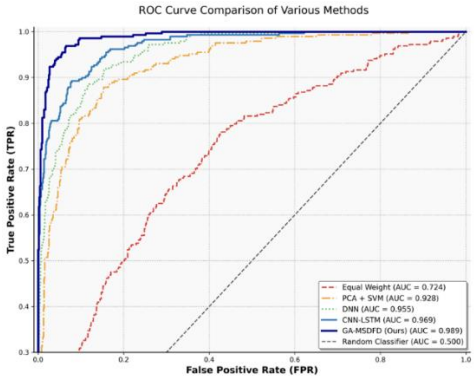

Fig. 4. Comparison of ROC curves of various methods

Table 2 shows the results of six diagnostic methods for machine fault diagnosis. As can be seen from the table, our method outperforms the other five baselines in all three metrics. In terms of accuracy, our method is 3 points higher than the best of the other methods. In terms of F1 score, our method is 3 points higher than the best of the other methods. This result demonstrates that our method has good generalization capabilities and performs well even on class-imbalanced datasets, significantly exceeding the other five baselines.

Figure 4 compares the ROC curves of our model with five other baselines for machine fault diagnosis. Our GA-MSDFD achieves an AUC of 0.989, significantly outperforming the other five machine fault diagnosis methods. The CNN-LSTM and DNN models also demonstrate excellent classification performance. The AUCs of these two machine fault diagnosis models reach 0.969 and 0.95, respectively, also excellent. The random classifier, on the other hand, achieves an AUC of only 0.5. This result demonstrates that our proposed model not only maximizes the true positive rate but also significantly reduces the false positive rate.

As shown in Table 3, when the population size is 30, the average number of convergence generations is 58.2, the optimal fitness score is 0.941, and the convergence time is 42.3 seconds. As the population size increases, the average number of convergence generations decreases, the optimal fitness score increases, and the runtime gradually increases. This demonstrates that our model has good convergence, the algorithm performs well, and is adaptable to populations of varying sizes.

As shown in Figure 5, we compared our model with five other machine fault diagnosis models. The specific results, shown in Figure 4, show that our model converged in nearly 100 generations with a population size of 200, and its fitness, or accuracy, was significantly higher than that of the other five models. This demonstrates that our model not only has strong convergence performance but also high accuracy.

Table 4. Noise immunity analysis of six machine fault detection methods

method	**SNR = ∞ (original)**	**SNR = 10 dB**	**SNR = 5 dB**	**Attenuation rate (5dB)**
DNN	91.2	83.4	75.6	17.1%
CNN-LSTM	93.5	86.7	79.3	15.2%
GA-MSDFD	96.8	94.1	91.5	5.5%

As shown in Table 4 , we artificially added varying proportions of noise to the dataset. In this case, we compared the performance of our method with several baselines . From the results, we can see that in the absence of noise, our model is already far superior to the other models. As the noise intensity increases, our model's performance slowly

Table 3. Convergence algebra of our model

Population size Np	**Average convergence algebra G∗**	**Optimal fitness Fmax**	**Run time (s)**
30	58.2	0.941	42.3
50	42.6	0.962	68.7
70	39.1	0.963	95.4

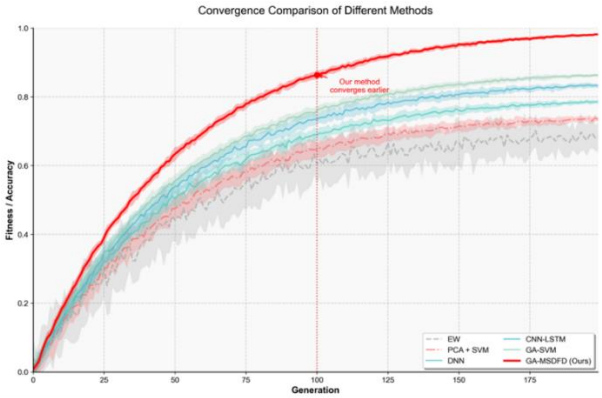

Fig. 5. Comparison of convergence time of six models

decreases, while the performance of the other models slowly increases. The final result is that our model's decay rate is 5.5%, while that of the other models exceeds 15%. This proves that our model is very robust in the face of noise interference. This also shows from another perspective that our data noise screening and data evolution strategies are effective, proving the rationality of the model design.

As shown in Figure 6, we conducted 100 experiments on six machine fault diagnosis algorithms and presented the results in a boxplot. The figure shows that our algorithm has the strongest noise resilience. Its upper and lower quartiles and average are the highest among the algorithms. However, our algorithm has a high fluctuation range, which may be a problem. Overall, however, our method has the strongest noise resilience.

As shown in Table 5, we conducted an ablation analysis to analyze the contribution of features provided by different types of sensors to the results.

Table 5. Comparison of average weights of modal features of each sensor

Sensor Type	Feature dimension range	Average weight	Expert expected weight ranking
Vibration (x3 axis × 2)	1–6	0.81	1 (highest)
Current	7–8	0.53	2
pressure	9–10	0.42	3
temperature	11	0.31	4 (lowest)

The results show that the vibration sensor contributes the most to the results, while the temperature sensor contributes the least to the results, which is also consistent with the expected weights of the experts.

4.7. Discussion

As shown in Table 6, we conducted an ablation analysis of the model across five sets of experiments. The results show that the complete model achieves the strongest performance. Removing the data noise assessment and data purification modules results in a 6.9% drop in performance. Removing the genetic algorithm weight optimization module results in a 4% drop in performance. Conversely, replacing the bidirectional LSTM implemented in this paper with a standard LSTM only results in a 1.3% drop in performance. Removing the expert rule results in a 1.8% drop in performance. This demonstrates that the core innovations of this paper lie in the data noise and evolution module and the genetic algorithm weight optimization module, which contribute most significantly to the results.

5. CONCLUSION

This paper proposes a machine fault diagnosis system based on multi-sensor data fusion based on genetic algorithms. This method can solve the noise problem of industrial sensor data to a certain extent and improve the accuracy of fault diagnosis. The experiments in this paper show that:

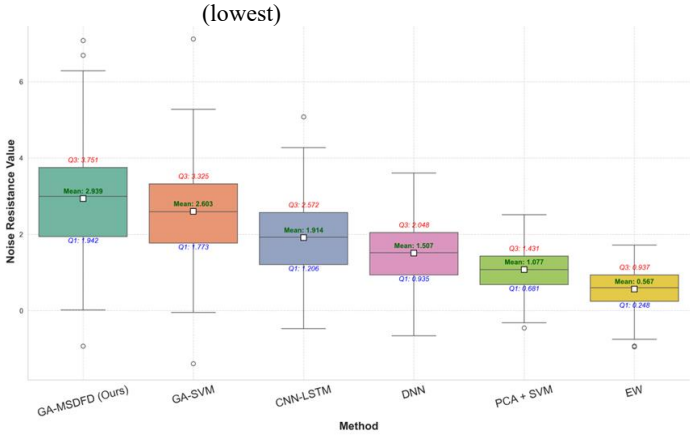

Fig. 6. Noise resistance performance of different algorithms in 100 experiments

Table 6. Ablation experiment results of 6 different model variants

Model variants	Accuracy (%)	F1-score	AUC	Performance degradation
Our Model (Full)	96.8	0.962	0.987	-
Remove data noise evaluation and purification module	90.1	0.893	0.932	6.9%
Removed the genetic algorithm weight optimization module	92.0	0.915	0.953	4.8%
Replace Bi-LSTM with LSTM	95.5	0.949	0.976	1.3%
Remove expert rule constraints	95.0	0.943	0.971	1.8%

(1) The dynamic weight mechanism proposed in this paper can effectively improve the accuracy of diagnosis and overcome the adaptability of static weights. The final result exceeds the optimal baseline by more than 3%.

(2) Noise processing of multi-sensor data can enhance the robustness of fault diagnosis. Experimental results demonstrate that the noise-processed method, proposed in this paper, significantly outperforms other methods in terms of noise robustness. Even with artificially added 5dB noise, its performance degraded by only 5.5%, a far lower degradation than that observed in several other models.

(3) Using genetic algorithms to optimize the feature extraction process of Bi-LSTM can greatly reduce the amount of data for effective features, thereby shortening the model's convergence generations and convergence time, and greatly improving the model's training and inference efficiency.

Source of funding: *This research received no external funding.*

Author contributions: *research concept and design, C.H., Y.Z.; Collection and assembly of data, C.H., Y.Z.; Data analysis and interpretation, C.H., Y.Z.; Writing the article, C.H, Y.Z.; Final approval of the article, C.H., Y.Z.*

Declaration of competing interest: *The author declares no conflict of interest.*

REFERENCES

1. Zhang WZ, Yan XA, Ye MY, Hua X, Jiang, D. Intelligent fault diagnosis based on multi-sensor data fusion and multi-scale dual attention enhanced network. Measurement Science and Technology. 2025;36(2):20. https://doi.org/10.1088/1361-6501/ada6ec.
2. Drosinska-Komor M, Gluch J, Brenkacz L, Ziólkowska N, Piotrowicz M, Ziólkowski P. Advanced genetic algorithm-based signal processing for multi-degradation detection in steam turbines. Mechanical Systems and Signal Processing. 2025; 224:30. https://doi.org/10.1016/j.ymssp.2024.112166.
3. Han XM, Xu J, Song SN, Zhou JW. Crack fault diagnosis of vibration exciter rolling bearing based on genetic algorithm-optimized Morlet wavelet filter and empirical mode decomposition. International Journal of Distributed Sensor Networks. 2022;18(8):12. https://doi.org/10.1177/15501329221114566.
4. Hu C, Xing FT, Pan SH, Yuan R, Lv Y. Fault diagnosis of rolling bearings based on variational mode decomposition and genetic algorithm-optimized wavelet threshold denoising. Machines. 2022; 10(8):19. https://doi.org/10.3390/machines10080649.
5. Hua LH, Zhang JF, Li DJ, Xi XB, Shah MA. Sensor fault diagnosis and fault tolerant control of quadrotor uav based on genetic algorithm. Journal of Sensors. 2022;8. https://doi.org/10.1155/2022/8626722.
6. Mo CC, Han HZ, Liu M, Zhang QH, Yang T, Zhang F. Research on SVM-based bearing fault diagnosis modeling and multiple swarm genetic algorithm parameter identification method. Mathematics. 2023;11(13):28. https://doi.org/10.3390/math11132864.
7. Ren B, Bai D, Xue ZP, Xie H, Zhang H. Method for fault feature selection for a baler gearbox based on an improved adaptive genetic algorithm. Chinese Journal of Mechanical Engineering. 2022;35(1):12. https://doi.org/10.1186/s10033-022-00728-x.
8. Shi GZ, Gu XH, Yang SP, Li HF, Liu ZC. Variable-length genetic algorithm-based blind deconvolution for railway bearing fault diagnosis. IEEE Sensors Journal. 2025;25(12):22041-22056. https://doi.org/10.1109/jsen.2025.3563763.
9. Shi RM, Wang BK, Wang ZY, Liu JQ, Feng XY, Dong L. Research on fault diagnosis of rolling bearings based on variational mode decomposition improved by the niche genetic algorithm. Entropy. 2022;24(6):14. https://doi.org/10.3390/e24060825.
10. Ugli OEM, Lee KH, Lee CH. Automatic optimization of one-dimensional CNN architecture for fault diagnosis of a hydraulic piston pump using genetic algorithm. IEEE Access. 2023;11:68462-68472. https://doi.org/10.1109/access.2023.3287879.
11. Yuan Y, Ma SL, Wu JW, Jia BW, Li, WX, Luo XW. Fault diagnosis in gas insulated switchgear based on genetic algorithm and densitybased spatial clustering of applications with noise. IEEE Sensors Journal. 2021;21(2):965-973. https://doi.org/10.1109/jsen.2019.2942618.
12. Chao Q, Gao HH, Tao JF, Liu CL, Wang YH, Zhou J. Fault diagnosis of axial piston pumps with multi-sensor data and convolutional neural network. Frontiers of Mechanical Engineering. 2022;17(3): 15. https://doi.org/:10.1007/s11465-022-0692-4.
13. Huo DY, Kang YY, Wang BY, Feng GF, Zhang JW, Zhang HR. Gear fault diagnosis method based on multi-sensor information fusion and VGG. Entropy. 2022;24(11):15. https://doi.org/10.3390/e24111618
14. Kibrete F, Woldemichael DE, Gebremedhen HS. Multi-Sensor data fusion in intelligent fault diagnosis of rotating machines: A comprehensive review. Measurement. 2024;232:17. https://doi.org/10.1016/j.measurement.2024.114658.
15. Kim S, Lee S, Lee J, Kim M, Kim SJ, Yoon H, Youn BD. Fault-relevance-based, multi-sensor information integration framework for fault diagnosis of rotating machineries. Mechanical Systems and Signal Processing. 2025;232:27. https://doi.org/10.1016/j.ymssp.2025.112742.
16. Li YZ, Luo X, Xie YH, Zhao WB. Multi-head spatio-temporal attention based parallel GRU architecture: a novel multi-sensor fusion method for mechanical fault diagnosis. Measurement Science and Technology. 2024;35(1):16. https://doi.org/10.1088/1361-6501/acfe29.
17. Liu C, Tong JY, Zheng JD, Pan HY, Bao JH. Rolling fault bearing diagnosis method based on multi-sensor two-stage fusion. Measurement Science and Technology. 2022;33(12):14. https://doi.org/10.1088/1361-6501/ac8894.
18. Sun KC, Yin AJ. Multi-sensor temporal-spatial graph network fusion empirical mode decomposition convolution for machine fault diagnosis. Information Fusion. 2025;114:12. https://doi.org/10.1016/j.inffus.2024.102708.
19. Xu ZR, Li QM, Qian LF, Wang, MY. Multi-sensor fault diagnosis based on time series in an intelligent mechanical system. Sensors. 2022;22(24):16. https://doi.org/10.3390/s22249973.

20. Yang ZK, Li G, Xue G, He B, Song Y, Li X. A novel multi-sensor local and global feature fusion architecture based on multi-sensor sparse Transformer for intelligent fault diagnosis. Mechanical Systems and Signal Processing. 2025;224:28. https://doi.org/10.1016/j.ymssp.2024.112188.
21. Yin AJ, Sun ZY, Zhou JL Hypergraph construction using Multi-Sensor for helicopter Tail-Drive system fault diagnosis. Measurement. 2024;231:13. https://doi.org/10.1016/j.measurement.2024.114586.
22. Zarchi M, Shahgholi M. An expert condition monitoring system via fusion of signal processing for vibration of industrial rotating machinery with unseen operational conditions. Journal of Vibration Engineering & Technologies. 2023;11(5): 2267-2295. https://doi.org/10.1007/s42417-022-00702-w.

Appendix A.

Experimental Details and System Configuration

A.1 Dataset Description

The primary dataset used in this study is derived from the Case Western Reserve University (CWRU) Bearing Data Center, which provides vibration signals under various fault conditions and operating loads. The original CWRU dataset includes four fault categories:

Normal condition (NC)

Inner race fault (IRF)

Outer race fault (ORF)

Rolling element fault (REF)

To simulate a realistic industrial scenario, we extended the dataset by integrating multi-modal sensor data collected from a custom-built test rig, including:

Three-axis vibration sensors (sampled at 12 kHz)

Current sensors (sampled at 1 kHz)

Pressure transducers (sampled at 500 Hz)

Thermocouples (sampled at 10 Hz)

Additionally, we introduced a composite fault class (CF), representing simultaneous occurrence of IRF and ORF, to evaluate the model's capability under complex failure modes. Each sample was segmented using a sliding window of length T = 1000, resulting in a total of 5,000 labeled samples (1,000 per class).

Sensor Deployment and Synchronization

All sensors were mounted on a rotating machinery test bench driven by a 3-hp induction motor. Time synchronization across heterogeneous sensors was achieved via hardware-triggered sampling using a National Instruments DAQ system (NI USB-6366). Timestamp alignment was further refined in software using linear interpolation to ensure temporal consistency before fusion.

GA-MSDFD Hyperparameters

Key hyperparameters of the proposed GA-MSDFD framework are summarized in table 7.

All experiments were implemented in Python 3.8, leveraging TensorFlow 2.10 for Bi-LSTM and DEAP v1.3.3 for genetic algorithm optimization. The source code and processed datasets are available upon reasonable request for academic validation.

Expert Rule Constraints (Re(w))

The expert rules embedded in the fitness function include:

Vibration features must contribute ≥60% of total weight.

Temperature features are penalized if their cumulative weight exceeds 10%.

If pressure sensor BDR > 0.3, its weight is capped at 0.2. These rules were formulated in consultation with three senior maintenance engineers from an automotive manufacturing plant.

Table 7. Hyperparameters

Component	Parameter	Value
Bi-LSTM	Hidden units per direction	64
	Dropout rate	0.5
	Output feature dimension	64 × K (K = number of sensors)
Genetic Algorithm	Population size (N)	50
	Crossover probability	0.8
	Mutation probability	0.1
	Selection method	Tournament (size = 3)
	Elitism	Top 10% preserved
	Maximum generations	100
	Chromosome encoding	Real-valued weights ∈ [0,1]
Fitness Function	α (accuracy weight)	0.6
	β (sparsity weight)	0.3
	γ (expert rule weight)	0.1
	Sparsity threshold τ	0.05
Noise Purification	SNR quality threshold (PSNR)	≥25 dB
	Wavelet basis (Level 2)	db4
	Kalman filter α (Level 3)	0.95

Caodi HU was born in Jiaozuo, Henan Province, China, in 1990. He received the B.Eng. degree in Mechanical Manufacturing and Automation from Henan University of Science and Technology, Luoyang, Henan Province, China, in 2013, the M.Eng. degree in Mechanical Engineering from the same university in 2016, and the Ph.D. degree in Information Technology from University of the East (Philippines), Manila, Philippines, in 2022.

From 2016 to 2018, he served as a Research and Development Engineer at Aeolus Tire Co., Ltd., focusing on the optimization of tire manufacturing equipment and process improvement. Since 2019, he has been working as a Lecturer in the School of Mechanical and Electrical Engineering, Jiaozuo University, Jiaozuo, Henan Province, China. He is mainly responsible for teaching undergraduate courses including Industrial Robotics, Mechanical Drawing, and 3D Printing Technology, and

has guided more than 20 students in their graduation design projects. His research interests include numerical simulation and optimization of mechanical equipment, intelligent manufacturing technology and its application, and intelligent control of industrial robots.

Mr. Hu has published several papers including "Numerical Simulation Study on the Wind Rotor of H-Type Vertical Axis Wind Turbine", "Research on the Development Path of Industrial Robot Industry Based on System Dynamics", "Optimization of Structural Parameters of Vacuum Magnetic Gear by Response Surface Methodology and Its Performance Analysis", and "Microstructure and Mechanical Properties of (TiN+SiC)/AlSi7Cu2Mg Composites Fabricated by SLM", etc.

e-mail: caodi_hu@hotmail.com

Yumeng ZHAO was born in Jiaozuo, Henan, China, on October 2, 1996. She received the B.S. degree in software engineering from Zhengzhou University, Zhengzhou, China, in 2019 and the M.S. degree in digital media technology from The University of Sydney, Sydney, Australia, in 2023.

In 2023, she served as a Project Manager at the Xiaomi Research and Development Center. She is currently a Teaching Assistant at Jiaozuo University, Jiaozuo, China. Her current research interest includes deep learning.

Ms. Zhao is dedicated to the application of cutting-edge technology in education and digital media fields.

e-mail: zhaoyumeng_zym@outlook.com