

# Intelligent Fault Diagnosis of Delta 3D Printers using Attitude Sensors based on Extreme Learning Machines

Xiaoyan Li<sup>a,b</sup>, Jianwen Guo<sup>a,\*</sup>, Xuejun Jia<sup>c</sup>, Shaohui Zhang<sup>a</sup>, and Zhiyuan Liu<sup>b</sup>

<sup>a</sup>*School of Mechanical Engineering, Dongguan University of Technology, Dongguan, 523808, China*

<sup>b</sup>*College of Mechatronics and Control Engineering, Shenzhen University, Shenzhen, 518060, China*

<sup>c</sup>*Dongguan Neutron Science Center, Dongguan, 523890, China*

---

## Abstract

The influence of intelligent fault diagnosis on industrial development is becoming more and more important. In order to study the fault diagnosis technique of delta 3D printers using extreme learning machine (ELM), a low-cost attitude sensor was used in our designed machine. In the research process, the cross validation method was used to train ELM to obtain the optimal model. Through the analysis of different activation functions, we found that the correct recognition rates corresponding to the same activation function are different, and there are great differences among training samples and fault categories. The sin function, mexihat function, and tribas function recognition effects were better. The analysis of different activation functions revealed that the correct recognition rates corresponding to the same activation function are different, and there are great differences in different training samples and fault categories.

**Keywords:** fault diagnosis; extreme learning machine; activation function; delta 3D printer

(Submitted on August 7, 2019; Revised on November 13, 2019; Accepted on November 17, 2019)

© 2019 Totem Publisher, Inc. All rights reserved.

---

## 1. Introduction

3D printers have emerged due to the developments of manufacturing techniques. They have largely reduced the cost of products, benefiting enterprises economically and improving efficiency of resources [1]. As 3D printers become increasingly popular in industry, their intelligent fault diagnosis becomes significant.

Intelligent fault diagnosis technology is different from traditional fault diagnosis technology, which observes changes or trends of a certain state detection index [2]. Intelligent fault diagnosis is effective analysis using a large amount of collected data that automatically provides reliable diagnosis results [3-4]. At present, intelligent fault diagnosis is mostly applied to rotating machinery. Qian et al. [5] proposed a three-stage fault diagnosis network based on supervised regularized spars filtering faulting diagnosis of a bearing case and a gearbox case. Adrián et al. [6] presented an online detection of novel faults and automatic learning using fuzzy clustering techniques. Zhang et al. [7] proposed a subset-based deep auto-encoder for bearing fault diagnosis. Yan et al. [8] developed a new scheme based on improved variational mode decomposition and instantaneous energy distribution permutation entropy. Chen et al. [9] utilized weak transient signals based on a convolutional restricted Boltzmann machine model for bearing fault diagnosis. Long et al. [10] proposed a Pareto-based evolutionary algorithm featuring a genetic algorithm combined with a local search strategy to tackle the multi-objective. Ma and Wu [11] proposed discriminant sparse and collaborative preserving embedding for bearing defect classification. Lu et al. [12] proposed a new method involving transient signal analysis-based angular resampling to realize variable-speed motor bearing fault diagnosis. Zhao et al. [13] presented a novel fault diagnosis approach of rolling bearing based on feature reduction using global-local margin Fisher analysis. Li et al. [14] presented a model of statistical feature learning from vibration measurements of rotating machinery.

The aforementioned intelligent fault diagnosis techniques are applied to the bearings of rotating machinery, and most intelligent diagnostic algorithms require separate feature extraction of data [15]. When referring to 3D printers, there are

\* Corresponding author.

E-mail address: [guojw@dgut.edu.cn](mailto:guojw@dgut.edu.cn)

many benefits such as freedom of design, mass customization, waste minimization, the ability to manufacture complex structures, and fast prototyping [16]. Currently, 3D printers are commonly used in the field of medicine [17] and manufacturing [18]. With the increasing use of 3D printers, the health of intelligent diagnostic 3D printers has become very important. He et al. [19] proposed an attitude monitoring approach that adopts support vector machines (SVM) to diagnose the faults of delta 3D printers. Zhang et al. [20] developed an error fusion of multiple sparse auto-encoders (EFMSAE) to monitor the condition of 3D printers dynamically. Although they both diagnosed delta 3D printers, the high-precision sensors they used are costly. High-precision sensors have high maintenance costs and complicated methods of use in engineering practice. When the precision of the sensors is poor, the recognition effect of the model will become worse, and there are few diagnostic studies on 3D printers at present. In this paper, a fault diagnosis method for delta 3D printers using a low precision attitude sensor based on ELM is proposed. The nine-channel attitude sensor is used to obtain data, and then the data is classified by the extreme learning machine (ELM), so as to achieve fault diagnosis for 3D printers.

In this paper, we proposed intelligent fault diagnosis of delta 3D printers by adopting attitude sensors based on extreme learning machines (ELM). Compared with SVM, ELM is more superior in terms of generalization performance and convergence speed [21]. At present, ELM is extensively applied in the engineering field. Liu et al. [22] proposed a novel failure diagnosis method combining extreme learning machine and Dempster-Shafer evidence theory to solve the data-driven failure problems. Yang et al. [23] proposed a dual-ELM-based fault diagnostic and applied it to multistage gearbox systems to quickly and accurately detect faults. Hu et al. [21] proposed an effective method based on a weighted extreme learning machine (WELM) with wavelet packet decomposition (WPD) and kernel principal component analysis (KPCA).

Extreme learning machine (ELM) structure is a single hidden layer feedforward neural network whose training speed is quicker than other structures. It has a global optimum characteristic and optimal approximation performance [24].

At present, there are many studies on intelligent fault diagnosis using ELM, but their objects of fault diagnosis are not 3D printers, and most of them are improved algorithms to boost the recognition accuracy. In this paper, the fault diagnosis and analysis of delta 3D printers are carried out. Data collection is performed using a low cost attitude sensor, and then the data is analyzed for failure using different activation functions in the extreme learning machine. Though experimental analysis, it is shown that the sinusoidal function, mexihat function, and tribas function recognition effects are better. Regarding stability and convergence speed, the sinusoidal function performs the best among them.

## 2. Methodology

### 2.1. Base of Extreme Learning Machine (ELM)

Extreme learning machine (ELM) is a kind of progressive learning machine algorithm designed for single layer feedforward neuron networks (SLFN) [25]. Compared to support vector machines (SVMs) and neural networks [26], ELM eliminates drawbacks such as slow learning speed, negligible human intervention, and poor scalability.

ELM has an output layer, a hidden layer, and an input layer. The structure diagram is shown in Figure 1. The parameters are randomly generated between the input layer and the hidden layer. The only parameter of output weight that needs to be optimized is that between the hidden layer and the output layer.

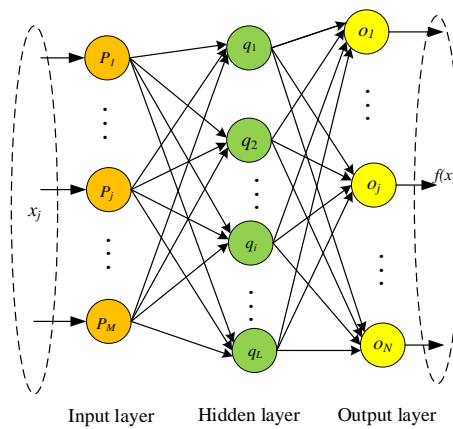


Figure 1. ELM structure diagram

Assume that there is a total of  $L$  hidden layer nodes, the quantity of input layer nodes is  $M$ , the output function of ELM is  $N$ , and the quantity of training sample is  $Q$ .

$$f_L(\mathbf{x}) = \sum_{i=1}^L \delta_i(\mathbf{x}) \cdot \gamma_i = \sum_{i=1}^L \gamma_i \zeta(\alpha_i, \tau_i, \mathbf{x}), \quad \mathbf{x} \in R^M, \quad \gamma_i \in R^L \quad (1)$$

Where  $x_i$  is the  $i^{\text{th}}$  input node, and  $\gamma_i$  connects the weight vector of the hidden layer and the output node.  $\delta_i(\mathbf{x})$  represents the output function  $\zeta(\alpha_i, \tau_i, \mathbf{x})$  of the  $i^{\text{th}}$  hidden node. ELM trains the SLFN in two main phases: stochastic feature mapping and linear parameter solving.

For additive nodes with activation function  $\xi$ ,  $\delta_i$  is defined as

$$\delta_i = \zeta(\alpha_i, \tau_i, \mathbf{x}) = \xi(\alpha_i \cdot \mathbf{x} + \tau_i), \quad \alpha_i \in R^M, \quad \tau_i \in R \quad (2)$$

Where  $\alpha_i$  is the weight vector connecting the input layer node and the  $i^{\text{th}}$  hidden layer node.  $\tau_i$  is the node deviation of the  $i^{\text{th}}$  hidden layer. The hidden node parameters  $(\alpha_i, \tau_i)$  of ELM can be randomly generated.  $\zeta(\alpha_i, \tau_i, \mathbf{x})$  is a nonlinear piecewise continuous function that satisfies the general approximation ability theorem of ELM.

The only thing that needs to be solved is the output weight in the extreme learning machine, and the interpolation theorem is usually used to solve weights. The approximation error is minimized to solve  $\gamma$ .

$$\min_{\gamma \in R^{L \times N}} \|\mathbf{Y}\gamma - \mathbf{G}\|^2 \quad (3)$$

Equation (1) can be written as

$$\mathbf{Y}\gamma = \mathbf{G} \quad (4)$$

$$\mathbf{Y} = \begin{bmatrix} \xi(\alpha_1 \cdot x_1 + \tau_1) & \cdots & \xi(\alpha_L \cdot x_1 + \tau_L) \\ \vdots & \ddots & \vdots \\ \xi(\alpha_1 \cdot x_Q + \tau_1) & \cdots & \xi(\alpha_L \cdot x_Q + \tau_L) \end{bmatrix}_{Q \times L} \quad (5)$$

$$\gamma = \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_L \end{bmatrix}_{L \times 1} \quad (6)$$

$$\mathbf{G} = \begin{bmatrix} g_1 \\ \vdots \\ g_Q \end{bmatrix}_{Q \times 1} \quad (7)$$

$$\gamma = \mathbf{Y}^\dagger \mathbf{G} \quad (8)$$

Where  $\mathbf{Y}$  is the hidden layer output matrix, and the  $i^{\text{th}}$  column of matrix  $\mathbf{Y}$  represents the  $i^{\text{th}}$  hidden layer unit output of the  $Q$  training samples.  $\mathbf{G}$  is the target matrix, and  $g_i$  is the label corresponding to the sample data.  $\mathbf{Y}^\dagger$  is the Moore–Penrose generalized inverse of matrix  $\mathbf{Y}$  [27–28]. To calculate the Moore–Penrose generalized inverse of a matrix, different methods can be used, such as the orthogonalization method, orthogonal projection method, iterative method, and singular value decomposition (SVD).

## 2.2. Activation Function

Activation functions are important for neural network models to learn and understand complex and nonlinear functions. The nonlinearity of a neural network will enhance the performance of the activation function. In addition, it can reinforce neural

network function approximation.

Commonly used activation functions in practical applications are shown in Table 1. The geometric construction is displayed in Figure 2.

Table 1. Failure mode table

Pattern no.	Failure location	Pattern no.	Failure location
1	Normal mode	9	Bearing failure h
2	Bearing failure a	10	Bearing failure i
3	Bearing failure b	11	Bearing failure g
4	Bearing failure c	12	Bearing failure k
5	Bearing failure d	13	Bearing failure l
6	Bearing failure e	14	Synchro belt failure o
7	Bearing failure f	15	Synchro belt failure p
8	Bearing failure g	16	Synchro belt failure q

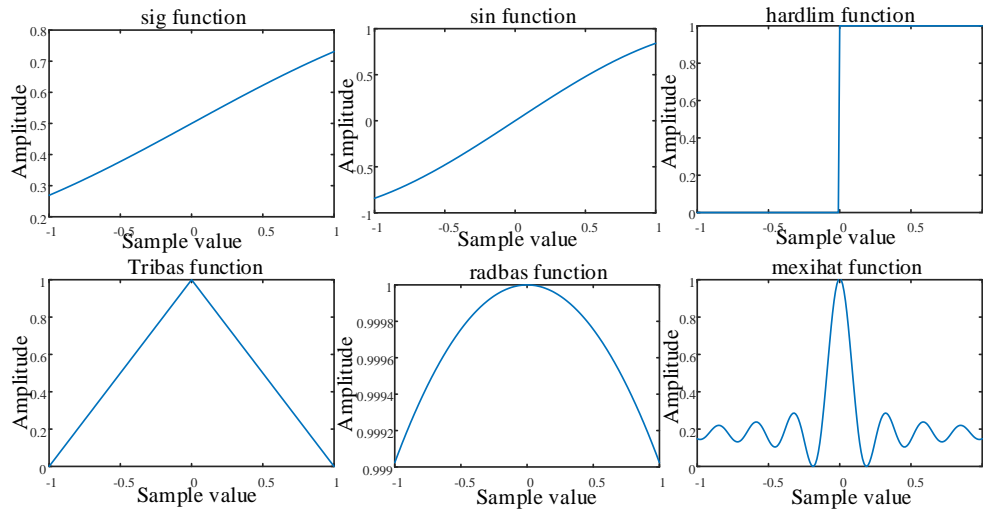


Figure 2. Geometric construction of activation function

Sinusoidal function:

$$\zeta(x) = \sin x \quad (9)$$

Tribas function:

$$\zeta(x) = \begin{cases} 1 - |x|, & \text{if } -1 \leq 0 \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

Radial basis function:

$$\zeta(x) = \exp(-b \cdot \|x - a\|) \quad (11)$$

Sigmoid function:

$$\zeta(x) = \frac{1}{1 + \exp(a \cdot x + b)} \quad (12)$$

Hard limit function:

$$\zeta(x) = \begin{cases} 1, & \text{if } a \cdot x + b \leq 0 \\ 0, & \text{otherwise} \end{cases} \quad (13)$$

Mexihat function:

$$\zeta(x) = (1 - x^2) e^{-\frac{x^2}{2}} \quad (14)$$

It is necessary that the activation function is in a larger amount of the neural network. Its function is to add nonlinear factors to solve the defect of insufficient expression ability of the linear model. Commonly used activation functions include the sinusoidal function, tribas function, radial basis function, sigmoid function, hard limit function, and mexihat function.

In Figure 2, the abscissa indicates the sample value, and the ordinate is the amplitude in a six activation function. The sig function was used frequently in the past, and it can be seen that its curve approximates a straight line and is monotonically increasing in one cycle. Its amplitude is greater than zero in the case of sample values between -1 and 1. The peak value is around 0.7. However, as its frequency of usage increases, gradient back-transfer in deep neural networks may cause gradient explosions and the disappearance of gradients. It can be seen that in the sample value [-1, 1], the sin function is similar to the sig function curve, but its curve fluctuation will be slightly larger than that of sig. The amplitude of the sin function is less than zero within [-1, 0), greater than zero within (0, 1], and equal to 0 when the sample value is equal to zero. The function is monotonically increasing in [-1, 1]. The main function of the activation function is to add nonlinear factors to solve the defect of insufficient linear model expression ability. However, the hardlim function curve shows that when the sample value is less than 0, the amplitude is zero, and when the sample value is higher than zero, the amplitude suddenly becomes 1. It can be speculated that this function is not well recognized in the ELM model. The tribas function is named because its shape resembles a triangle. Its amplitude is always higher than zero in the interval [-1, 1] with 0-symmetry. On the interval [-1, 0], it monotonically increases, and it monotonously decreases on the interval [0, 1]. The peak value is obtained at  $x = 0$ , which is 1. In the hidden layer, the nodes of the radial basis function (RBF) have their own centroids and impact factors. Their output is given by a radially symmetric function of the distance between the input and the center. From the curve of RBF, it can be seen that the curve resembles an arch. In the interval [-1, 1], the amplitude fluctuations are not large. The mexihat function, as the name suggests, is shaped like a Mexican hat. Its amplitude is higher than zero in [-1, 1] and equal to zero at  $x = 0.2$  and  $x = -0.2$ . The amplitude reaches a maximum of 1 when  $x = 0$ . At the same time, we can see that the image fluctuations are relatively large and have symmetry about  $x = 0$ . Moreover, the curve fluctuations are relatively large.

### 3. Experiment

#### 3.1. The Experimental Steps

As mentioned in previous section, we will use ELM for the fault detection of delta 3D printers. An attitude sensor of model BWT901 is used to collect data. This is an economically affordable attitude sensor that meets our experimental requirements, while also highlighting the advantages of ELM.

The attitude sensor BWT901 has nine channels. The output of the sensor is time, acceleration, angular velocity, angle, and magnetic field. In this paper, we use acceleration, angular velocity, and magnetic field. The  $x$ ,  $y$ , and  $z$  dimensions are measured separately. The voltage is 3.5-5V, the current requirement is less than 40mA, the data output frequency is 0.1-200Hz, and the data interface serial port is TTL level. In this paper, 200Hz is selected as the output frequency. The introduced method flowchart is shown in Figure 3. The general steps are as follows:

**Step 1** Install an attitude sensor on the moving platform of the delta 3D printer.

**Step 2** Connect the sensor to the computer through the network cable for data collection.

**Step 3** Without performing any signal preprocessing or feature extraction, divide the raw data into training samples and test samples, and then classify and label them.

**Step 4** Separate the training samples into training data and test data.

**Step 5** Test the trained model with a set of data to estimate its accuracy.

**Step 6** Verify the validity of the method by testing the sample.

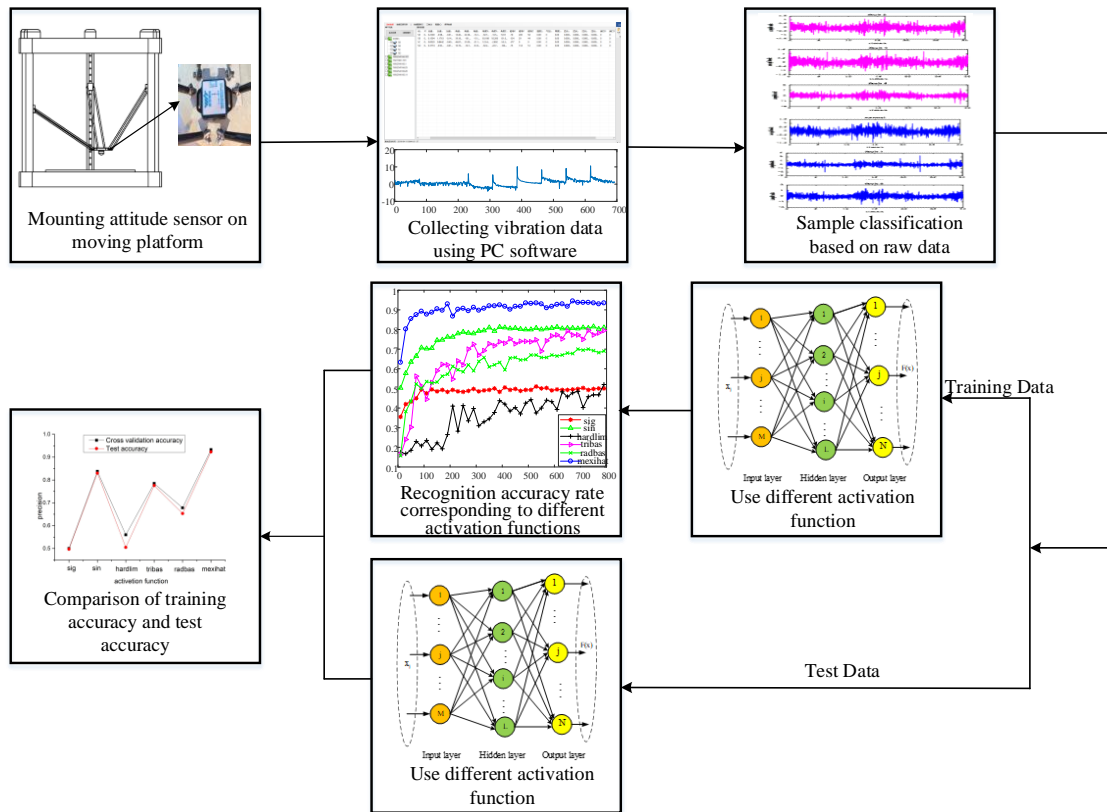


Figure 3. The flowchart of the method

### 3.2. Experimental Data of a Delta 3D Printer

The experimental setup is composed of a laptop, a delta 3D printer, and the model sensors Dell-N4110, SLD-BL600-6, and BWT901.

We verified the effectiveness of the introduced method using a delta 3D printer. For the experiment, we selected the delta 3D printer model SLD-BL600-6 from SHILEIDI of Dongguan, China, the BWT901 attitude sensor from RION of Shenzhen, China, and the Dell laptop model Inspiron N4110.

The attitude sensor was installed on the moving platform and connected to the computer through a network cable. The upper computer software was used to collect the output data from the attitude sensor and monitor the status data under different fault conditions. During the experiment, the sampling frequency of the attitude sensor was set to 100Hz.

The experiment was carried out using a delta 3D printer, which is shown in Figure 4. The mechanical components are mainly composed of a motor, a bracket, a synchronous wheel, a conveyor belt, a moving platform, and an extrusion nozzle. The delta 3D printer uses the fused deposition modeling (FDM) principle to convert a Slice file into a G code. It fuses the filaments of the material into heated base nozzles that are attached to the mobile platform and deposits them on the constructed platform. On the basis of code control, under the motor drive and slider on the vertical rail of the 3D printer, the joint bearing and the synchro belt move together, the nozzle of the extruder moves in the X-Y direction, the plane moves the printed product, and the slider moves along the Z-axis direction, completing the slice thickness.

There are three linear transmission motor drives on the delta 3D printer. In these parallel kinematic chains, each chain is a parallelogram, ensuring that the motion platform is always parallel to the ground and parallel to the platform being built. In other words, when the vibration is within a certain range, the moving platform and the built platform are not parallel, which will result in poor quality of printed products. If it is more serious, the printing task will fail.

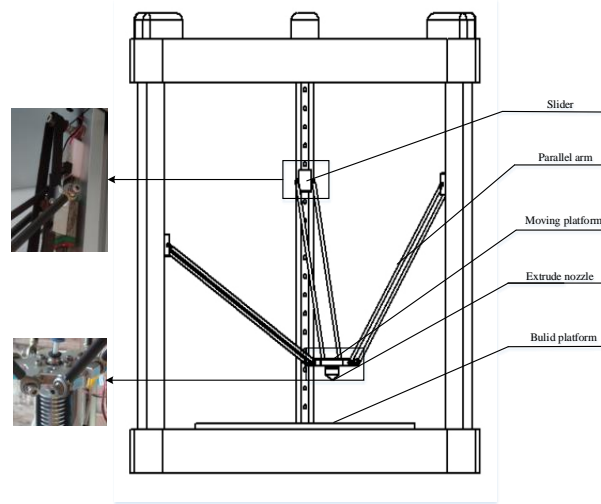


Figure 4. Structure of delta 3D printer

From the kinematics, dynamics, friction, and wear models, it can be seen that a direct solution to the fault problems of delta 3D printers is complicated and difficult. Failures may occur due to wear of joint bearings. Wear is difficult to detect, but it can affect the quality of printed products. Therefore, this paper establishes a data-driven model for the diagnosis of delta 3D printers. To avoid creating and solving complex kinematics, dynamics, and frictional wear models, we tried to monitor the health of delta 3D printers using only one sensor, while considering kinematics, dynamics, and friction wear.

### 3.3. Fault Category Setting

Typically, a delta 3D printer has 12 bearing joints and 3 synchro belts. In the experiment, in order to simulate different failure modes under different working conditions, loose joint bearing and the synchro belt were selected to configure different joint bearing wear mode modes. We divided the health status of the delta 3D printer into 16 modes, one for a normal mode and 15 for failure modes. Among them, 12 failure modes were loose joints of the joint bearing two turns. The other three failure modes let the synchro belt loosen three teeth and loosen one belt at a time. Table 1 lists the health statuses of the delta 3D printer under all conditions.

The created model was a cylindrical shell with a radius of 75mm and a height of 0.3mm. The model code was saved as an STL format file. The file was imported into the replication host to generate the G code of the moving platform, and the attitude sensor was installed on the printer's moving platform to operate and collect data. Here, we set the sensor frequency as 100Hz and the attitude sensor as nine channels. Each channel had 77 samples, for a total of 693. Each fault was set to 600 samples, recorded as  $M$ . The training samples were separated into four groups: 50, 100, 150, and 200. The fault modes were separated into three categories for experimentation. The first category was eight faults, the second category was 12 faults, and the third category was 16 faults, as shown in Table 2.

Table 2. Experimental data

$M$ (Total samples)	600			
$m$ (Training samples)	50	100	150	200
Fault category	8	8	8	8
	12	12	12	12
	6	6	6	6

## 4. Experimental Data Analysis

In order to obtain a more reliable and stable model, we chose to use cross-validation methods for training. In actual training, the degree of fit of the training results to the training set is usually very good (initial conditions are sensitive), but the degree of fit to the data outside the training set is typically unsatisfactory. Therefore, the training set is commonly separated into two parts, one for training and the other for testing. The model generated by the training set is tested in a relatively objective manner to determine how consistent these parameters are with the data outside the training set. The result is a more accurate model [29].

In the time domain, the data waveform of the vibration signals collected by 16 faults after natural logarithm transformation is shown in Figure 5. The waveform of the original data after the activation function is shown in Figure 6. After comparing the original data with the processed data, we found that the data feature after activation function processing is more obvious.

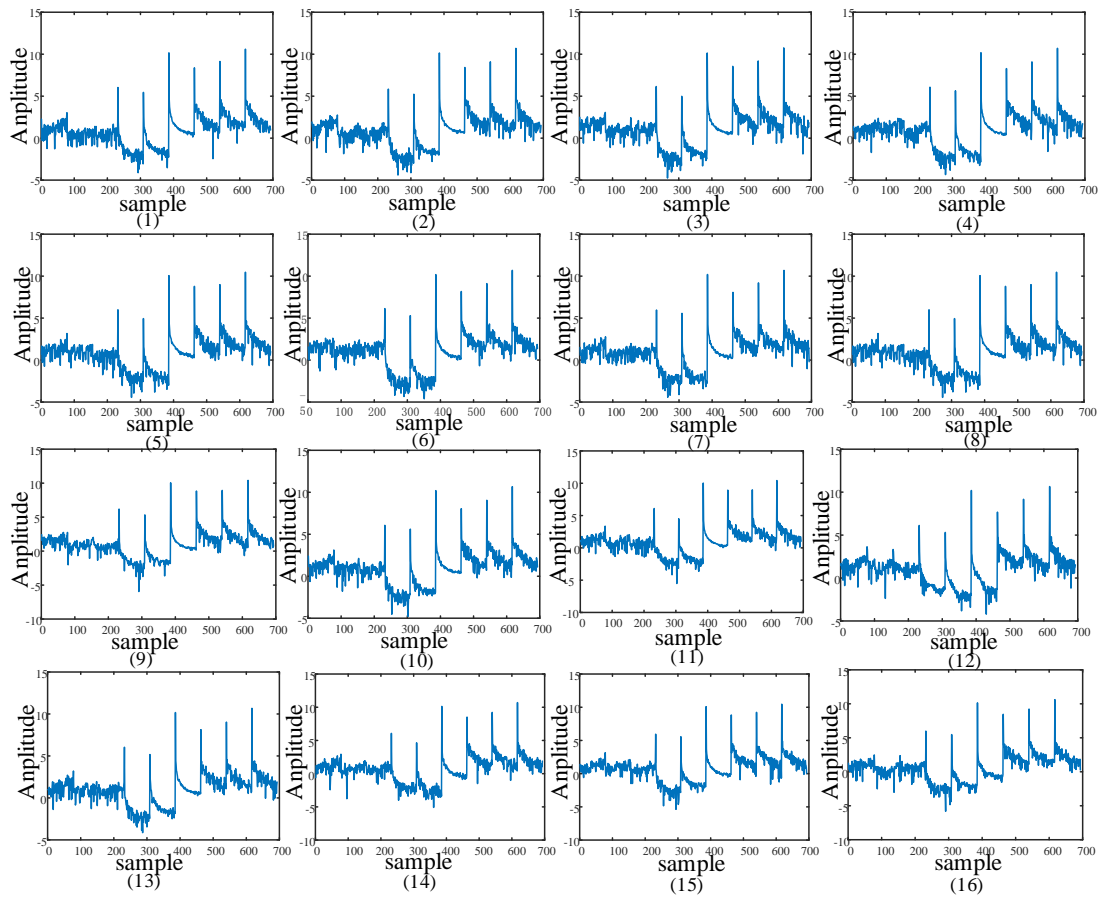


Figure 5. Different fault data of structure, where (1)-(16) represent the vibration data sequentially collected by the attitude sensor in Table 1

In Figure 6, the abscissa is the sample value, and the value range is [0, 700]. The ordinate is the amplitude value, and the value range is [-1, 1]. The six figures represent the sig, sin, hardlim, tribas, radbas, and mexihat functions. The vibration data is more obvious when the function is activated.

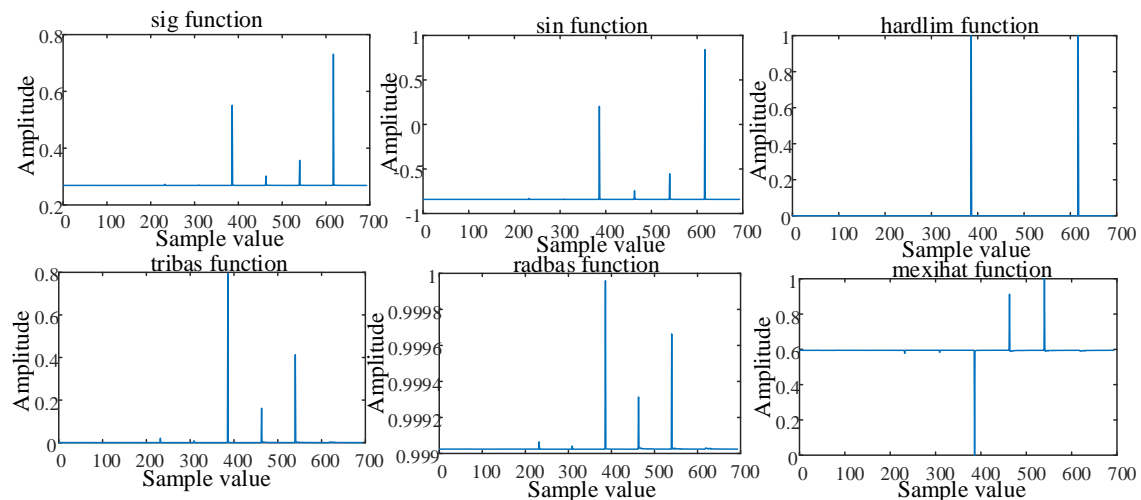


Figure 6. Data structure of the time domain signal after activation of the function



In this training class, 4/5 of the training specimens were used as training specimens and 1/5 was used as verification samples. In other words, 80% of the training samples were trained in the ELM model, the remaining 20% was used as a verification set to test the trained model, and then the ELM was evaluated as a performance indicator of the classifier.

When the number of samples is 600 and the corresponding quantity of training samples is 50, the fault categories are 8, 12, and 16, respectively. The results of the training parameters obtained by cross-validation are shown in Figure 7.

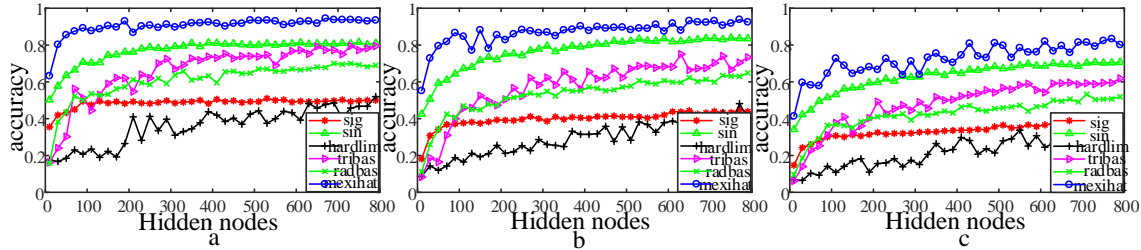


Figure 7. Comparison of recognition accuracy under different activation functions: the fault categories of a, b, and c are 8, 12, and 16, respectively

As can be seen from Figure 6, the superiority of the activation function is improved. The correct recognition rate of the mexihat activation function under different fault categories is much higher than those of other activation functions. When the failure mode is 8 and 12, the highest correct recognition rate of the mexihat activation function ELM model is 93%. When the failure mode is 16, the correct recognition rate is still higher than those of the other five activation functions, i.e., 83.4%. Among the other five activation functions, the sin activation function is the best, but its highest correct recognition rate is 84%. In the case of eight types of failure modes, the correct recognition rate improves the sin activation function. The failure mode is 8%, which is higher than the 6% of the sin activation function. When improving the ELM model under the sin activation function in the three types fault categories, the lowest fault recognition rate is 40%, the fault is 16 classes, and the correct recognition rate is at least 32%. For the eight types of faults, the lowest fault category of the improved sin is 67%, and the lowest correct recognition rate of the model under the sin activation function is 51%. In both cases, the activation of the sin activation function is 8% lower. The tribas activation function has a higher correct recognition rate than the radbas, sig, and hardlim activation functions, but its curve fluctuates greatly. The model curve under the sig activation function is smooth, but its recognition accuracy is too low.

When the amount of specimens is 600, the quantity of training specimens is 100, and the fault categories are 8, 12, and 16. The results of the training parameters obtained by cross-validation are shown in Figure 8.

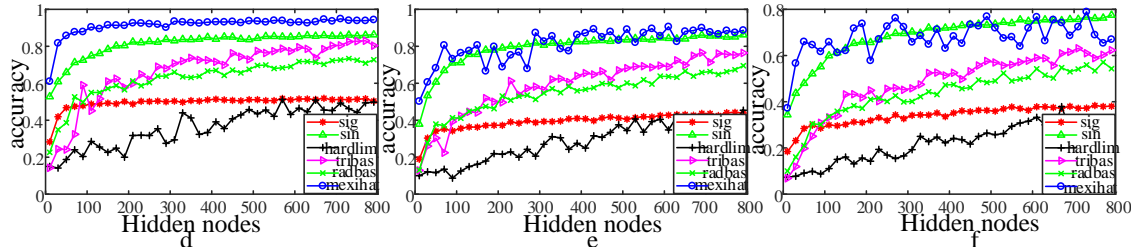


Figure 8. Comparison of recognition accuracy under different activation functions: the fault categories of d, e, and f are 8, 12, and 16, respectively

Figure 8 shows the correct recognition rate of the ELM model for different activation functions of the training sample of 100. From the three figures d, e, and f, the correct recognition rate of the model can be seen. When the fault category is eight categories, the highest fault recognition rate of the mexihat function is 94%. When the fault category is 12 categories, the highest fault recognition rate of the sin function is almost the same as that of the mexihat function, but the mexihat function recognizes that the curve is not smooth. The highest correct recognition rate of the model under the sin activation function is about 86%. When the failure is 16 categories, the recognition rate is reduced to 77%. The tribas, radbas, sig, and hardlim activation functions have lower recognition accuracies that vary greatly and are unstable under different fault categories.

When the amount of samples is 600, the quantity of training samples is 150, and the fault categories are 8, 12, and 16. The results of the training parameters obtained by cross-validation are shown in Figure 9.

In different failure modes, the activation functions have different correct recognition rates and vary across a wide range. When the fault category is eight categories, the mexihat function has the highest recognition accuracy rate of 96%, followed

by the sin function (88%). The two functions have the fastest convergence rate. When the fault categories are 12 and 16, the fault recognition rates of the sin function are 87% and 79%, respectively. Although the accuracy of the meixihat function is still high, the curve is cluttered, and there is no convergence trend. In the three fault categories, the convergence speed of the sig function is consistent, but its recognition precision is too low.

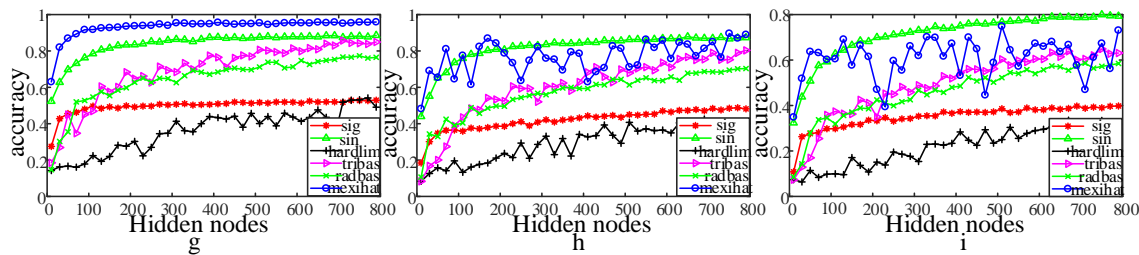


Figure 9. Comparison of recognition accuracy under different activation functions: the fault categories of g, h, and i are 8, 12, and 16, respectively

When the amount of samples is 600, the quantity of training samples is 200, and the fault categories are 8, 12, and 16. The results of the training parameters obtained by cross-validation are shown in Figure 10.

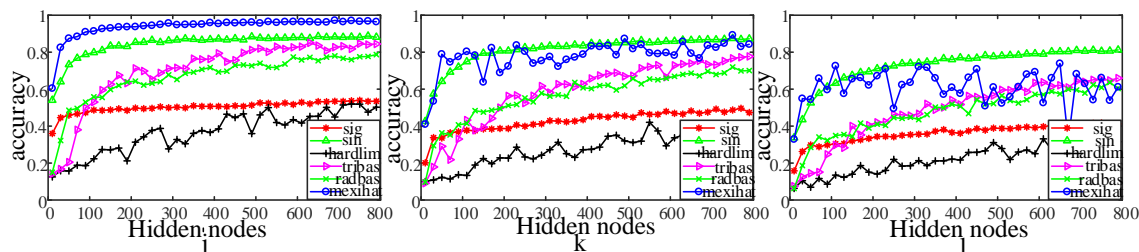


Figure 10. Comparison of recognition accuracy under different activation functions: the fault categories of j, k, and l are 8, 12, and 16, respectively

Comparing the cases with the different quantities of training samples, it is found that the recognition rates of the same type of activation function are different under different training samples. When the failure mode is eight categories, the training specimens are 150 and 200, and the recognition accuracies of the mexihat function and the sin function are correct. The highest and fastest convergence is achieved. When the fault categories are 12 and 16, the training specimens are 150 or 200, and the recognition accuracy of the sin activation function is superior to those of the other five activation functions, with a maximum of 87%.

It is clearly shown in Figures 7-10 that by comparing the correct recognition rates of different activation functions, the correct recognition rate of the ELM model will be different. The figures also verify the effectiveness and superiority of the improved activation function for the correct recognition rate of ELM.

The optimal ELM model under different activation functions is trained by cross-validation. Next, the trained model is verified by the test data. Using the mapping method, the hidden layer number of the ELM model with the highest recognition accuracy under different activation functions can be seen from Figures 7-10. The hidden layer is tested by the hidden layer to obtain the recognition accuracy rate.

Figure 11 shows the highest accuracy of the ELM model at different activation functions for training samples at 50, 100, 150, and 200, as well as failure modes for categories 8, 12, and 16, respectively.

In Figure 11, the abscissa indicates different activation functions; ordinate indicates the recognition accuracy; a, b, and c, indicate that the training specimen is 50 and the failure categories are 8, 12, and 16; d, e, and f indicate that the training sample is 100 and the failure categories are 8, 12, and 16; g, h, and i indicate that the training specimen is 150 and the failure categories are 8, 12, and 16; and j, k, and l indicate that the training sample is 200 and the failure categories are 8, 12, and 16.

As can be observed from Figure 11, when the failure mode has eight categories, the training specimens of the sig activation function are cross-validated at 50, 100, 150, and 200, and the accuracy and test accuracy are around 0.5. When the number of training specimen is 50, the cross-validation of the activation function accuracy is 0.83. The training specimen of 100, 150, and 200 with training accuracy and test accuracy is around 0.88. The training specimens of the hardlim activation function are cross-validated at 50, 100, and 150, which is 0.56. The cross-validation accuracy of the

training sample is 0.50 at 200. When the training specimens are 50 and 200, the test accuracy is 0.50. When the training specimens number are 100 and 150, the training sample is 0.45. When the training sample is 50, the tribas activation function is cross-validated with an accuracy and test accuracy of 0.78. When the training sample is 100, the cross-validation accuracy and test accuracy are 0.82. When the training sample is 100, the cross-validation accuracy and test accuracy are 0.85. When the training sample is 50, the radbas activation function is 50%, and the measurement accuracies are 0.68 and 0.65. The training specimens have 100 and 200 cross-validation accuracy with a test accuracy of 0.76. The training sample is 150, and the cross-validation accuracy and test accuracy are 0.80 and 0.77, respectively. The mexihat function has cross training specimens of 50, 100, 150, and 200, a fault category of 8, and a verification test with an accuracy and precision of 0.96. When the training specimens are 50, 100, 150, and 200, the fault category is 12, and the accuracy and precision of the test are 0.90 and 0.88, respectively. When the fault category is 16, the accuracy and precision of the test are reduced.

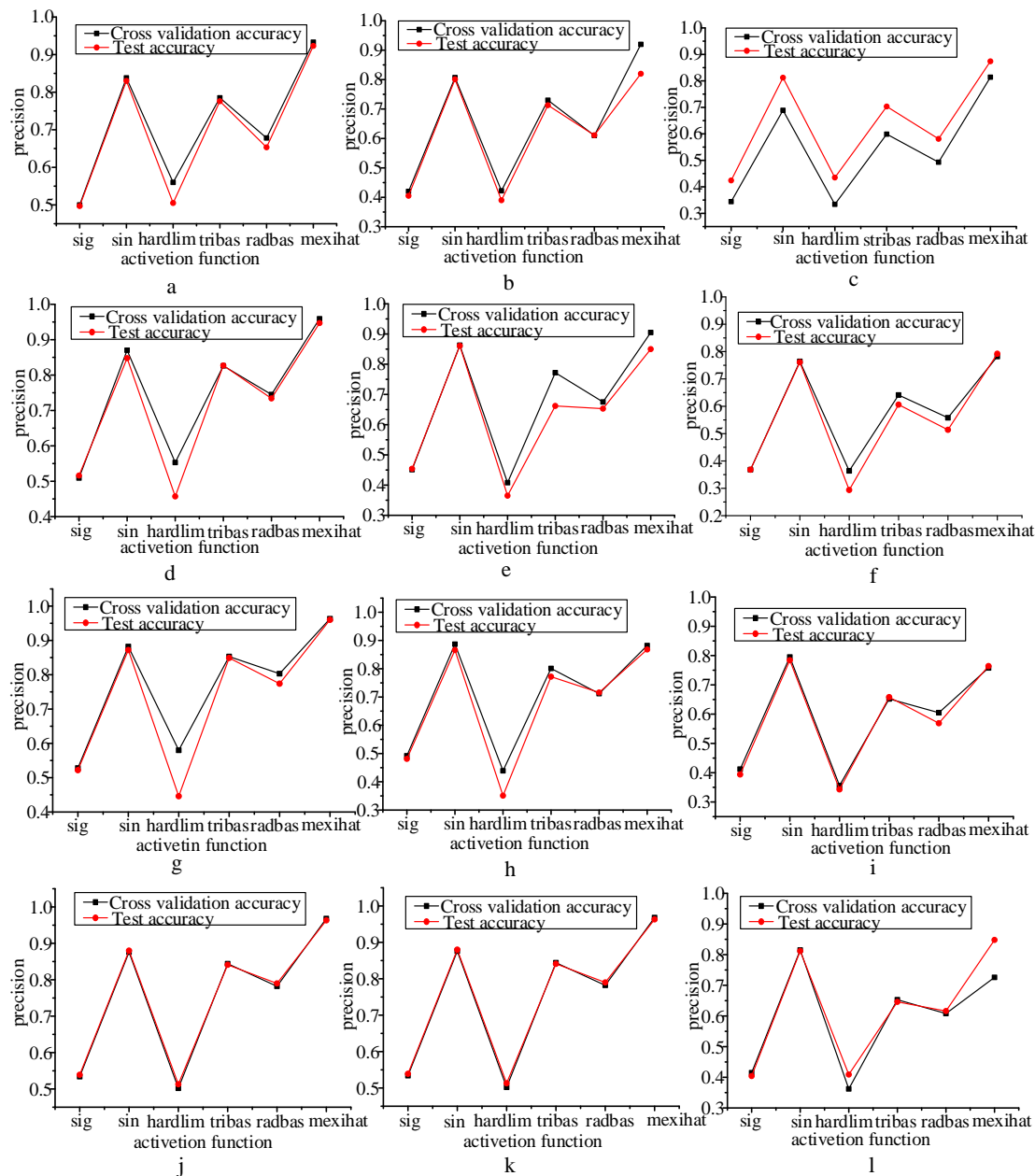


Figure 11. Training accuracy and test accuracy at different activation functions

5. Conclusions

Aiming at the problem of intelligent fault diagnosis of 3D printers, we proposed an intelligent fault diagnosis solution based on extreme learning machine (ELM) using low-cost and low-precision attitude sensors. In different failure modes, sample

data was collected, the ELM model was trained through cross-validation, the optimal model was found, and the test data was used for testing. During the experiment, ELM demonstrated high scalability and low computational complexity, avoiding signal preprocessing and feature extraction during the modeling process. The simulation results showed that the method has fast running speed and high recognition accuracy. During the training process, the correct recognition rate of the ELM model was trained. When the training specimens and the fault categories were different, the recognition efficiency of the same activation function was different, and there was also a relatively large individual activation function. In future studies, we will improve the activation function to increase its stability while enhancing the recognition accuracy.

## Acknowledgements

This work is supported partly by the National Natural Science Foundation of China (No. 51775112, 51605406), Natural Sciences Foundation of Guangdong in China (No. 2017A030313690), Department of Education of Guangdong in China (No. 2016KTSCX137, 2017KZDXM082, 2018KTSCX224), International Science and Technology Cooperation (Including Hong Kong, Macao, and Taiwan) Project of Dongguan (No. 2016508102010), Industry-University-Research Cooperation Base of DGUT & Heli Town, and Research Start-up Funds of DGUT (No. GC300501-12, GC300501-26).

## References

1. S. For and M. Despeisse, "Additive Manufacturing and Sustainability: An Exploratory Study of the Advantages and Challenges," *Journal of Cleaner Production*, Vol. 137, No. 20, pp. 1573-1587, November 2016
2. Z. Chen, S. Deng, X. Chen, C. Li, R. V. Sanchez, and H. Qin, "Deep Neural Networks-based Rolling Bearing Fault Diagnosis," *Microelectronics Reliability*, Vol. 75, pp. 327-333, August 2017
3. H. Shao, H. Jiang, X. Li, and S. Wu, "Intelligent Fault Diagnosis of Rolling Bearing using Deep Wavelet Auto-Encoder with Extreme Learning Machine," *Knowledge-based Systems*, Vol. 140, No. 15, pp. 1-14, January 2018
4. H. Shao, H. Jiang, Y. Lin, and X. Li, "A Novel Method for Intelligent Fault Diagnosis of Rolling Bearings using Ensemble Deep Auto-Encoders," *Mechanical Systems and Signal Processing*, Vol. 102, No. 1, pp. 278-297, March 2018
5. W. Qian, S. Li, J. Wang, and Q. Wu, "A Novel Supervised Sparse Feature Extraction Method and its Application on Rotating Machine Fault Diagnosis," *Neurocomputing*, Vol. 320, No. 3, pp. 129-140, December 2018
6. A. D. Rodríguez-Ramos, A. J. da Silva Neto, and O. Llanes-Santiago, "An Approach to Fault Diagnosis with Online Detection of Novel Faults using Fuzzy Clustering Tools," *Expert Systems with Applications*, Vol. 113, No. 15, pp. 200-212, December 2018
7. Y. Zhang, X. Li, L. Gao, and P. Lim, "A New Subset based Deep Feature Learning Method for Intelligent Fault Diagnosis of Bearing," *Expert Systems with Applications*, Vol. 110, No. 15, pp. 125-142, November 2018
8. X. Yan, M. Jia, and Z. Zhao, "A Novel Intelligent Detection Method for Rolling Bearing based on IVMD and Instantaneous Energy Distribution-Permutation Entropy," *Measurement*, Vol. 130, pp. 435-447, December 2018
9. L. Chen, G. Xu, Y. Wang, and J. Wang, "Detection of Weak Transient Signals based on Unsupervised Learning for Bearing Fault Diagnosis," *Neurocomputing*, Vol. 314, No. 7, pp. 445-457, November 2018
10. Y. Long, Z. Sun, P. M. Pardalos, Y. Hong, S. Zhang, and C. Li, "A Hybrid Multi-Objective Genetic Local Search Algorithm for the Prize-Collecting Vehicle Routing Problem," *Information Sciences*, Vol. 478, pp. 40-61, April 2019
11. Y. Ma and X. Wu, "Discriminant Sparse and Collaborative Preserving Embedding for Bearing Fault Diagnosis," *Neurocomputing*, Vol. 313, No. 3, pp. 259-270, November 2018
12. S. Lu and X. Wang, "Fault Diagnosis of Motor Bearing with Speed Fluctuation via Angular Resampling of Transient Sound Signals," *Journal of Sound and Vibration*, Vol. 385, No. 22, pp. 16-32, December 2016
13. X. Zhao and M. Jia, "Fault Diagnosis of Rolling Bearing based on Feature Reduction with Global-Local Margin Fisher Analysis," *Neurocomputing*, Vol. 315, No. 13, pp. 447-464, November 2018
14. C. Li, R. V. Sánchez, G. Zurita, M. Cerrada, and D. Cabrera, "Fault Diagnosis for Rotating Machinery using Vibration Measurement Deep Statistical Feature Learning," *Sensors*, Vol. 16, No. 6, pp. 895, June 2016
15. C. Li, Y. Tao, W. Ao, S. Yang, and Y. Bai, "Improving Forecasting Accuracy of Daily Enterprise Electricity Consumption using a Random Forest based on Ensemble Empirical Mode Decomposition," *Energy*, Vol. 165, Part B, No. 15, pp. 1220-1227, December 2018
16. T. D. Ngo, A. Kashani, G. Imbalzano, K. T. Q. Nguyen, and D. Hui, "Additive Manufacturing (3D printing): A Review of Materials, Methods, Applications and Challenges," *Composites Part B Engineering*, Vol. 143, No. 15, pp. 172-196, June 2018
17. J. Witold, K. Mateusz, C. Anna, S. Joanna, G. Karolina, and J. Renata, "3D Printing of Tablets Containing Amorphous Aripiprazole by Filaments Co-Extrusion," *European Journal of Pharmaceutics and Biopharmaceutics*, Vol. 131, pp. 44-47, July 2018
18. B. Berman, "3-D Printing: The New Industrial Revolution," *Business Horizons*, Vol. 55, No. 2, pp. 155-162, March-April 2012
19. H. Kun, Z. Yang, B. Yun, J. Long, and C. Li, "Intelligent Fault Diagnosis of Delta 3D Printers using Attitude Sensors based on Support Vector Machines," *Sensors*, Vol. 18, No. 4, April 2018
20. S. Zhang, Z. Sun, J. Long, C. Li, and Y. Bai, "Dynamic Condition Monitoring for 3D Printers by using Error Fusion of Multiple Sparse Auto-Encoders," *Computers in Industry*, Vol. 105, pp. 164-176, February 2019
21. Q. Hu, A. Qin, Q. Zhang, H. Jun, and G. Sun, "Fault Diagnosis based on Weighted Extreme Learning Machine with Wavelet Packet Decomposition and KPCA," *IEEE Sensors Journal*, Vol. 18, No. 20, pp. 8472-8483, October 2018

22. J. Liu, Q. Li, W. Chen, Y. Yan, and X. Wang, "A Fast Fault Diagnosis Method of the PEMFC System based on Extreme Learning Machine and Dempster–Shafer Evidence Theory," *IEEE Transactions on Transportation Electrification*, Vol. 5, No. 1, pp. 271–248, March 2019
23. Z. Yang, X. Wang, and P. Wong, "Single and Simultaneous Fault Diagnosis with Application to a Multistage Gearbox: A Versatile Dual-ELM Network Approach," *IEEE Transactions on Industrial Informatics*, Vol. 14, No. 12, December 2018
24. F. Du, J. Zhang, N. Ji, G. Shi, and C. Zhang, "An Effective Hierarchical Extreme Learning Machine based Multimodal Fusion Framework," *Neurocomputing*, Vol. 322, No. 17, pp. 141–150, December 2018
25. G. B. Huang, X. Ding, and H. Zhou, "Optimization Method based Extreme Learning Machine for Classification," *Neurocomputing*, Vol. 74, No. 1–3, pp. 155–163, December 2010
26. S. Dong, H. Jiang, X. Li, and S. Wu, "Intelligent Fault Diagnosis of Rolling Bearing using Deep Wavelet Auto-Encoder with Extreme Learning Machine," *Knowledge-based Systems*, Vol. 140, No. 15, pp. 1–14, January 2018
27. D. Serre, "Matrices: Theory and Applications," Springer Verlag, New York, 2002
28. C. R. Rao and S. K. Mitra, "Generalized Inverse of Matrices and its Applications," Wiley, New York, 1971
29. J. D. Rodriguez, A. Perez, and J. A. Lozano, "Sensitivity Analysis of k-Fold Cross Validation in Prediction Error Estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 32, No. 3, pp. 569–575, March 2010