

Strip Surface Defect Diagnosis Method Based on Extreme Learning Machine with Different Excitation Functions

Jia-Ning Hou, Jie-Sheng Wang *, Tian-Cheng Li

Abstract—Strip surface defects seriously impress the appearance of products, also lead to a decrease in wear resistance, corrosion resistance and the increase of fatigue strength of strip products, thus greatly reducing the service life of strip products. Therefore, a strip surface defect diagnosis algorithm by using on extreme learning machine (ELM) with variable excitation functions was proposed. Based on the neural network principle and the learning method of ELM, the seven excitation functions (Sigmoid, Sin, Hardlim, ReLu, Tanh, Morlet and Arctan) were used to establish the strip surface defect diagnosis model respectively. The neuron numbers of the hidden layer of ELM were analyzed experimentally. The simulation experiment results for the strip surface defect data set in UCI show that the ELM neural network can effectively diagnose the strip surface defect accurately.

Index Terms—strip surface defect diagnosis, extreme learning machine, excitation function

I. INTRODUCTION

THE iron and steel industry has long occupied the primary position in the field of economic development. Strip steel, as the major product in the steel manufacturing industry, is also an indispensable important material in the industrial production process, which is extensively adopted in various fields. In the process of strip production, due to the influence of many objective factors, the surface of the product will appear different types of defects. For example, the raw materials, rolling equipment and manufacturing process used in the manufacturing process will become the main reasons for the formation of surface defects. The types of surface defects of strip steel can be roughly divided into: needle eye, scar, scratch, scratch, bond, roll mark, hole, pitting, surface stratification, etc. [1]. These shortcomings have serious influence on the products surface and degrade the service life of the tape products, for example, reducing wear resistant and corrosion resistant, and increasing the

Manuscript received August 10, 2021; revised October 24, 2021. This research was supported by the Liaoning Provincial Basic Research Project for Higher Education Institutions (Approval No. 2019LNQN10) and the Liaoning Provincial Natural Science Foundation Project.(Grant No. 20180550700).

Jia-Ning Hou is a postgraduate student of School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan, 114051, P.R. China (e-mail: 19994905806@qq.com).

Jie-Sheng Wang is a professor of School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan, 114051, P. R. China. (Corresponding author, phone: 86-0412-2538246; fax: 86-0412-2538244; wang_jiesheng@126.com).

Tian-Cheng Li is an undergraduate student of School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan, 114051, P. R. China (e-mail: 1055800457@qq.com).

fatigue strength, so as to greatly reduce the service life of strip products [2]. Therefore, the automatic strip surface defects identification method is an essential research field in the steel industry.

A defect inspection algorithm on the basis of significant linear scanning morphology principle was recommended to detect the silicon steel strips defect under the interference of oil pollution, and a morphological edge processing strategy was proposed to remove oil pollution interference edge and reflection pseudo-defect edge [3]. The experiment results indicate the proposed algorithm has an excellent performance on detecting surface defects such as scratches, scratches and micro-defects. Based on the improved least-square twin support vector machine, a strip surface defect identification method was realized by adopting the extracted features [4]. A strip surface defect identification algorithm was proposed by combing second-order cone programming (SOCP) strategy and multi-kernel relevance vector machine (MKRVM) [5]. Based on the classification experiments, the identification accuracy and time efficiency are better than the existing strip surface defect identification algorithms. methods. To enhance the recognition of defects, an anti-noise diagnosis methodology was proposed to identify the surface defects of hot rolled strip based on complete localized binary model [6]. Experiments results illustrate the proposed strategy with higher recognition precision under the influence of defects on strip surface, such as in-class variation, illumination variation and gray scale variation.

ELM is a novel single-hidden layer feed-forward neural networks (SLFN) [7-10]. The ELM neural network has similar model structure with BP neural network, but ELM has a fixed number of layers, with input, hidden and output layers [11]. ELM has successfully realized face recognition [12], fault diagnosis [13], soft-sensing modeling [14], photovoltaic control system [15] and other areas. This manuscript proposed a strip surface defect diagnosis method based on ELM with different excitation functions. Simulation experiments are carried out to verify its effectiveness.

II. EXTREME LEARNING MACHINE NEURAL NETWORK AND EXCITATION FUNCTION

A. Extreme Learning Machine Neural Network

Among many artificial neural networks (ANN) models, the feed-forward neural network (FNN) has been studied by many researchers because of its good adaptive ability and nonlinear self-learning ability. Among them, the back

propagation neural network (BPNN) has been extensively used because of its good nonlinear mapping ability, adaptive ability, generalization and fault tolerance ability, but it also has some shortcomings, such as local minimization, slow converging velocity, too obvious sample dependence and long training time. ELM has the same structure as BPNN and is previously learned from the propagation mode [8]. ELM is not iteratively adjusted like BP neural network, but completed in turn by solving equations. The structural principle of ELM neural network is shown in Fig. 1 [11].

Suppose there is n neuron in the hidden layer, that is to say that there is an n input characteristic variable. The output layer has m neurons, that is, m characteristic variables. The hidden layer has l neurons. Suppose w is the connected weights matrix from input layer to hidden layer, on the other hand, w is a $(l \times n)$ matrix shown in Eq. (1), where each row represents the connection weights from the i -th neuron in the input layer to the n -th neuron in the hidden layer.

$$w = \begin{bmatrix} W_{11} & W_{12} & \dots & W_{1n} \\ W_{21} & W_{22} & \dots & W_{2n} \\ \dots & \dots & \dots & \dots \\ W_{l1} & W_{l2} & \dots & W_{ln} \end{bmatrix}_{l \times n} \quad (1)$$

where, the first line $(W_{11}, W_{12}, \dots, W_{1n})$ is the connection weight between neuron O_1 to n neurons, and there are altogether n neurons. The second line $(W_{21}, W_{22}, \dots, W_{2n})$ is the connection weight between neuron O_2 to n neurons, and the last line $(W_{l1}, W_{l2}, \dots, W_{ln})$ is the connection weight between the neuron O_l to n neurons.

Suppose that the connection weight from hidden layer to output layer is β , which is a $(l \times m)$ matrix shown in Eq. (2), whose each line describe the connection weight between the i -th hidden neuron and the m -th output neuron.

$$\beta = \begin{bmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1m} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2m} \\ \dots & \dots & \dots & \dots \\ \beta_{l1} & \beta_{l2} & \dots & \beta_{lm} \end{bmatrix}_{l \times m} \quad (2)$$

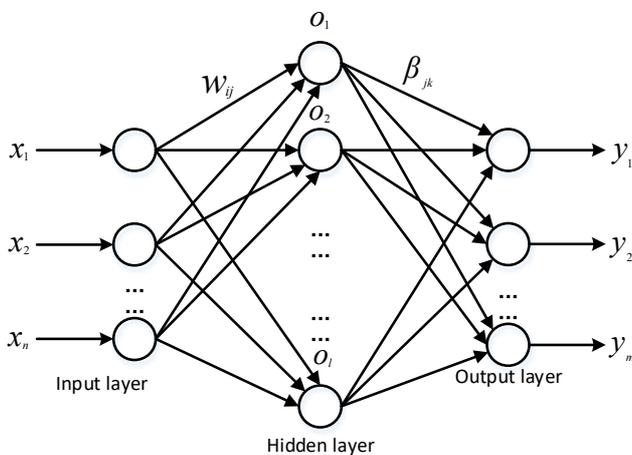


Fig. 1 ELM Structure.

For example, the second line $(\beta_{21}, \beta_{22}, \dots, \beta_{2m})$ in Eq. (3) describes the neuron connection weights from hidden layer to output layer.

$$b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_l \end{bmatrix}_{l \times 1} \quad (3)$$

where, b represents the neuron threshold of the hidden layer.

Assuming that there are Q samples in training data set, an input sample matrix can be written as:

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1Q} \\ x_{21} & x_{22} & \dots & x_{2Q} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{nQ} \end{bmatrix}_{n \times Q} \quad (4)$$

In Eq. (4), each row represents a feature and each column represents a sample. For example, the second subscript of the first column x is 1, the first subscript changes from 1 to n , which also represents the n characteristics of the first sample. Similarly, the output is a $(m \times Q)$ matrix (Q samples, m output variables), which can be expressed as follows.

$$Y = \begin{bmatrix} y_{11} & y_{12} & \dots & y_{1Q} \\ y_{21} & y_{22} & \dots & y_{2Q} \\ \dots & \dots & \dots & \dots \\ y_{m1} & y_{m2} & \dots & y_{mQ} \end{bmatrix}_{m \times Q} \quad (5)$$

After defining input samples, output samples, connection weights and hidden function neuron threshold among input, output and hidden layers, the forward propagating output T can be obtained, as shown in Eq. (6)-(7).

$$T = [t_1, t_2, \dots, t_Q]_{m \times Q} \quad (6)$$

$$t_j = \begin{bmatrix} t_{1j} \\ t_{2j} \\ \vdots \\ t_{mj} \end{bmatrix}_{m \times 1} = \begin{bmatrix} \sum_{i=1}^l \beta_{i1} g(w_i x_j + b_i) \\ \sum_{i=1}^l \beta_{i2} g(w_i x_j + b_i) \\ \dots \\ \sum_{i=1}^l \beta_{im} g(w_i x_j + b_i) \end{bmatrix}_{m \times 1} \quad (j = 1, 2, \dots, Q) \quad (7)$$

where, g is the activation function. The equation form in Eq. (6)-(7) can be converted into matrix form.

$$H\beta = T$$

$$H = \begin{bmatrix} g(w_1 \cdot x_1 + b_1)g(w_2 \cdot x_1 + b_2)g(w_l \cdot x_1 + b_l) \\ g(w_1 \cdot x_2 + b_1)g(w_2 \cdot x_2 + b_2)g(w_l \cdot x_2 + b_l) \\ \dots \\ g(w_1 \cdot x_Q + b_1)g(w_2 \cdot x_Q + b_2)g(w_l \cdot x_Q + b_l) \end{bmatrix}_{Q \times l} \quad (8)$$

where, H represents the output matrix.

B. Training Algorithm of ELM

According to the theory of SLFN and ELM [10], the desired optimal training of ELM will be transformed into a least squares solution β for solving $H\beta = T$, which is expressed in Eq. (9).

$$\|H\hat{\beta} - T\| = \min_{\beta} \|H\beta - T\| \quad (9)$$

When L is equal to N , the output is a reversible matrix H , which is also a square matrix. At this time, the ELM model learns samples in the way of zero error, and the output weight β can be expressed as:

$$\beta = H^{-1}T \quad (10)$$

Of course, in the real world, where most of the cases are $L \neq N$, the output is not a square matrix. Then the output weight β can be expressed by H^+ (generalized inverse of matrix H):

$$\beta = H^+T \quad (11)$$

It can also solve H^+ by orthogonal projection. Ridge regression theory is adopted to deal with the non-singular matrix. I/λ is added to the non-singular matrix ($H^T H$ and HH^T) in the solution to ensure the stability of the model, in which λ is a constant and I is a unit matrix.

$$\beta = H^+T = \begin{cases} H^T (\frac{I}{\lambda} + HH^T)^{-1}T & (L > N) \\ (\frac{I}{\lambda} + H^T H)^{-1}H^T T & (L < N) \end{cases} \quad (12)$$

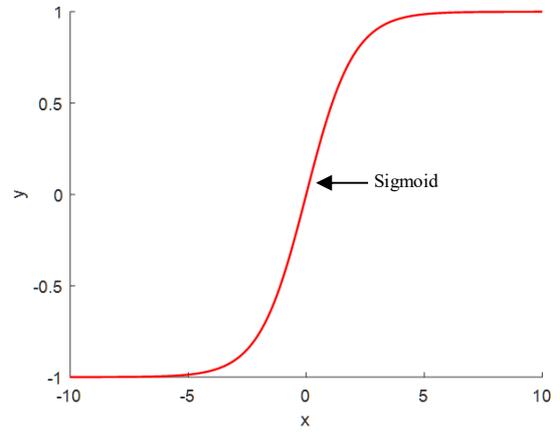
The output Y of ELM is calculated by:

$$y = h(x)\beta = \begin{cases} h(x)H^T (\frac{I}{\lambda} + HH^T)^{-1}T \\ h(x)(\frac{I}{\lambda} + H^T H)^{-1}H^T T \end{cases} \quad (13)$$

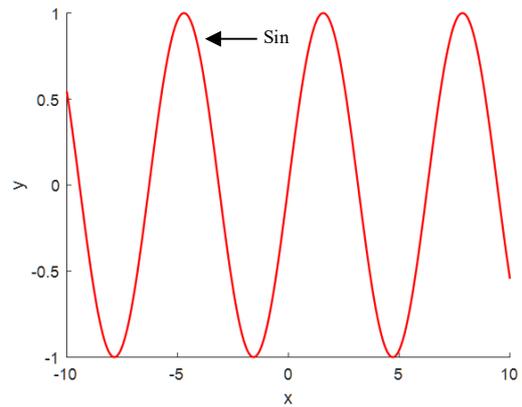
To sum up, the training algorithm of ELM includes three steps. (1) Generate hidden layer parameters randomly. (2) Obtain output matrix H of hidden neurons. (3) Obtain β for output weights.

C. Excitation Functions of ELM

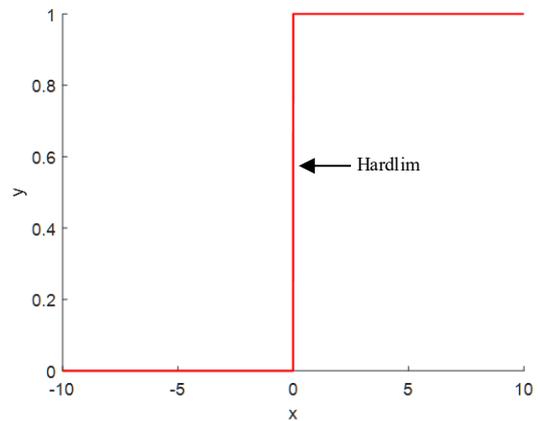
Based on analysis and experiment for ELM, from the difference theory perspective, if the excitation function g is indefinitely divisible in any given interval, the parameter of the excitation function g of the hidden neurons may be any value. For SLFN, ELM is able to learn the given input training set samples with any accuracy. Therefore, for studying the impact of excitation functions on the precision of the strip surface defect diagnosis, seven different activation functions are adopted in ELM neural network, which are Sigmoid, Sin, Hardlim, ReLu, Tanh, Morlet and Arctan [14]. Their function descriptions and images are listed in Fig. 2 and Table 1.



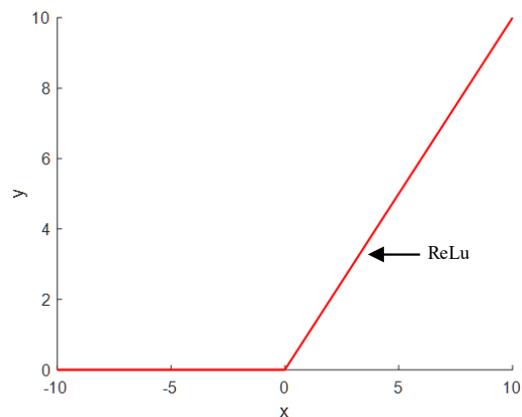
(a) Sigmoid



(b) Sin



(c) Hardlim



(d) ReLu

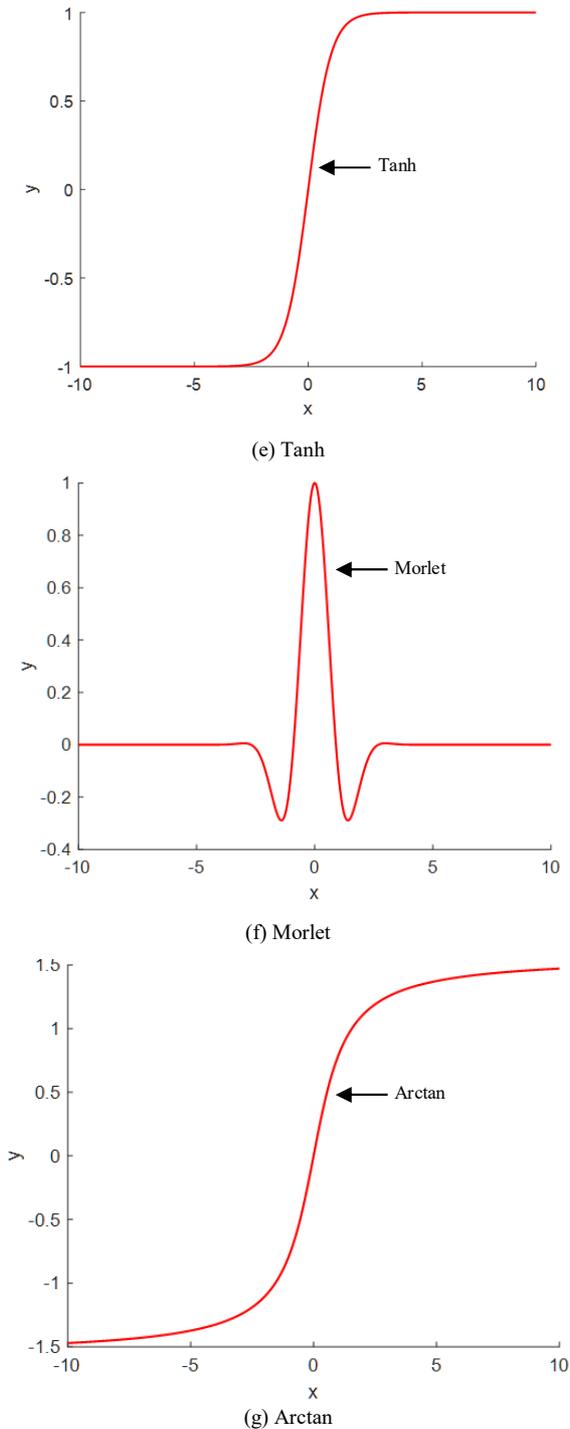


Fig. 2 Image of excitation functions.

TABLE 1. EXPRESSION OF EXCITATION FUNCTIONS

Function name	Function expression
Sigmoid	$f(x) = \frac{1}{1 + e^{-x}}$
Sin	$f(x) = \sin(x)$
Hardlim	$f(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases}$
ReLu	$f(x) = \begin{cases} 0 & x \leq 0 \\ x & x > 0 \end{cases}$
Tanh	$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$

Morlet	$f(x) = \cos(1.75x)e^{-x^2/2}$
Arctan	$f(x) = \arctan(x)$

III. CLASSIFICATION EXPERIMENTS AND SIMULATION RESULTS ANALYSIS

A. Classification Experiments Based on Iris Data Set

For verifying the effectiveness of ELM to solve the classification problems, the Iris data set in UCI database (as shown in Table 2) was first chosen to do the verification simulation experiments. Iris data set consists of three flower types (Setosa, versicolor, and Virginica). There are 150 samples in the database, with each flower accounting for one-third of the total. There are four characteristic values in Iris database samples, namely, calyx length, calyx width, petal length and petal width, in units of cm. Iris data set classification is to determine which category an item belongs to according to the four characteristic values. For the purpose of decreasing the simulation uncertainty, 150 data were randomly sorted, 120 data were picked to carry out the model training and 30 data were chosen to carry out the model testing. 20 hidden neurons were initialized and Sigmoid function was chosen as excitation function. The classification simulation results on Iris data set are described in Fig. 3. Seen from Fig.4, 28 test samples are classified accurately, the accuracy is 93.33% and the time is only 0.27 seconds.

B. Diagnosis Experiment of Strip Surface Defect

(1) Strip Surface Defect Data Set

The main objective of this experiment is to diagnose strip surface defects by using ELM. The chosen data set is the strip steel surface defects data in UCI database. There are a total of 1941 samples in this data set, and each sample has 27 characteristic values, such as X_Minimum, X_Maximum, Y_Minimum and Sigmoid of Areas. There are seven steel plate defects, which are Pastry, Z_Scratch, K_Scratch, Stains, Dirtiness, Bumps, and Other_Faults. Typical data samples in strip surface defect data set are listed in Table 3.

TABLE 2. IRIS DATA SET

Category	Serial number	Calyx length	Calyx width	Petal length	Petal width
Setosa	1	5.0	3.4	1.5	0.3
	2	4.9	3.5	1.4	0.2

Versicolor	50	5.1	3.6	1.4	0.2
	1	5.7	2.3	4.0	1.2
	2	5.6	2.6	4.2	1.5
Virginica
	50	5.9	2.9	4.5	1.3
	1	6.8	3.2	5.8	1.9
Virginica	2	6.9	3.1	5.5	2.1

	50	6.6	3.2	5.6	2.2

TABLE 3. DATA SET OF STRIP SURFACE DEFECTS

Eigenvalue	Steel plate defects						
	Pastry	Z_Scratch	K_scratch	Stains	Dirtiness	Bumps	Other_Faults
X minima	42	1166	464	1002	1325	38	1108
X maximum	50	1185	4724	1027	1339	49	1120
Y minima	270900	2258648	28542	155255	30207	735612	2497122
Y maximum	270944	2258662	28553	155262	30238	735624	2497138
Pixel area	267	123	72	62	268	113	137
Perimeter of X	17	33	13	31	29	11	13
Perimeter of Y	44	17	12	9	31	12	16
Brightness combined	24220	15858	13094	7479	25809	12652	15672
Minimum brightness	76	116	168	114	79	93	85
Maximum brightness	108	143	198	132	124	130	133
Transport captain	1678	1708	1387	1360	1353	1707	1373
300 steel	1	1	0	1	0	1	0
400 steel	0	0	1	0	1	0	1
Plate thickness	80	100	40	100	120	100	40
Edge index	0.0498	0.6124	0.6691	0.4897	0.0207	0.0445	0.3685
Empty index	0.2415	0.5376	0.3455	0.6457	0.3825	0.1439	0.2845
Square index	0.1818	0.7368	0.9091	0.28	0.4516	0.9167	0.75
External index	0.0047	0.0111	0.0072	0.0184	0.0104	0.0064	0.0087
X edge index	0.4706	0.5758	0.7692	0.8065	0.4828	1	0.9231
Y edge index	1	0.8235	0.9167	0.7778	1	1	1
External index	1	0	1	0	1	1	1
Logarithmic area	2.4265	2.0899	1.8573	1.7924	2.4281	2.0531	2.1367
Logarithmic index	0.9031	1.2787	1	1.3979	1.1461	1.0414	1.0792
Logarithmic index	1.6435	1.1461	1.0414	0.8451	1.4914	1.0792	1.2041
Directional index	0.8182	-0.2632	0.0909	-0.72	0.5484	0.0833	0.25
Photometric index	-0.2913	0.0072	0.4208	-0.0576	-0.2476	-0.1253	-0.1063
B region	0.5822	0.4399	0.2173	0.2998	0.7065	0.2432	0.3241

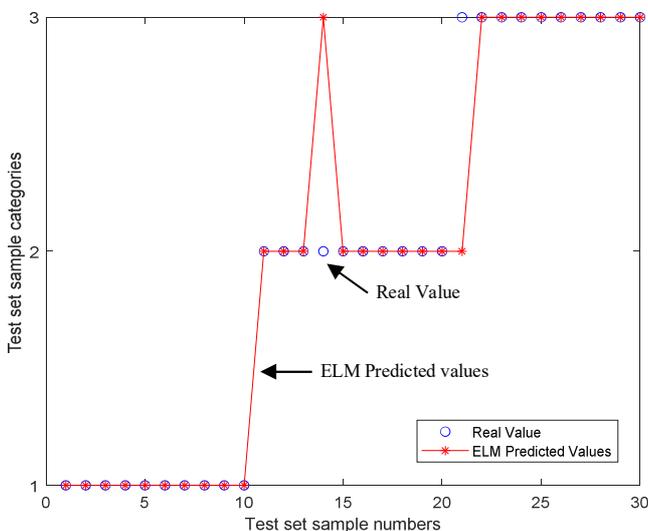


Fig. 3 Test results on iris data set.

(2) Diagnostic Results of Strip Surface Defects

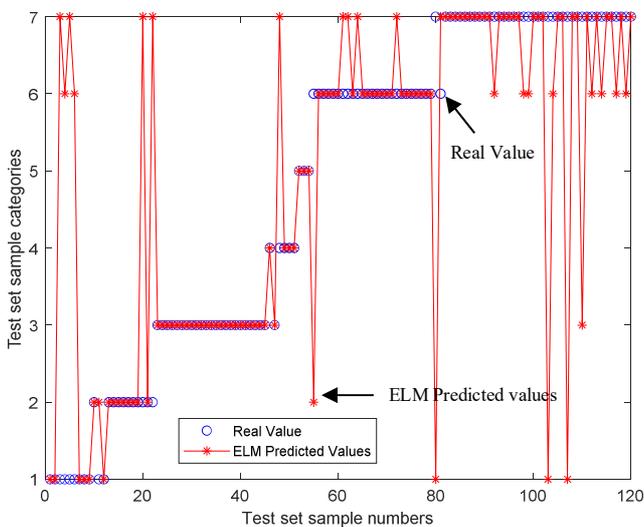
Before using the data set of strip surface defects in UCI database for diagnostic tests, the data need to be preprocessed. Since the data of these seven defects are not equal in quantity, the randomly scrambled order and grouping method is adopted to train and test these samples. 960 of the randomly distributed 1941 samples are adopted as the training data set and 120 as the testing data set. In addition, the dimension of each characteristic value in the strip surface defect data set is different, and the data normalization method is used to eliminate the influence of the dimension to improve the diagnosis accuracy.

In this experiment, a total of 120 samples were selected as test samples, the number of hidden neurons was initialized as 100 and 300, and adopted excitation function was Sigmoid function. The simulation results were described in Fig. 4. When the number of hidden neurons is 100, the

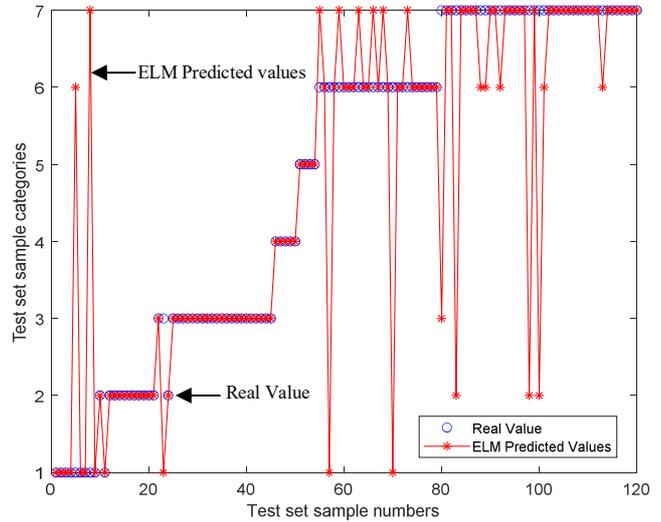
accuracy for training dataset was 79.48% and that of the testing dataset was 78.3%. When the number of hidden neurons is 300, the accuracy for training dataset was 90% and that of the testing set was 83.3%. In conclusion, when adopting the same activation functions and different number of hidden neurons, the diagnostic accuracy on testing set is also different. Therefore, in order to further study the relationship between the number of hidden neurons and the classification accuracy on the testing data set, further experimental verification is needed.

C. Optimal Number of Hidden Neurons for ELM

For further verifying the relationship between the number of hidden neurons and diagnostic precision, seven functions, namely Sigmoid function, Sin function, Hardlim function, ReLu function, Tanh function, Morlet function and Arctan function, were picked as the excitation functions of ELM respectively. The number of hidden neurons is set to 1-500. Through the diagnostic test on strip surface defects, the diagnostic results of ELM with different activation functions are shown in Fig. 5, where the number of hidden neurons is varied. According to Fig. 5, when the excitation functions are Sin function, Sigmoid function and ReLu function, the diagnostic results have higher accuracy on the testing data set. Although diagnostic result of ELM with Hardlim function is better, its fluctuation is relatively large. When the excitation function is Tanh and Arctan, the diagnostic accuracy is relatively lower, but the fluctuation is smaller. When the excitation function is Morlet it can be seen that the data fluctuates too much, and its test diagnosis accuracy is the lowest among the seven excitation functions. the diagnostic results of testing data sets corresponding to the optimal diagnostic accuracy of each activation function are described in Fig. 6, where the number of hidden neurons changes. On the other hand, the number of neurons corresponding to highest diagnostic accuracy of strip surface defects is listed in Table 4.

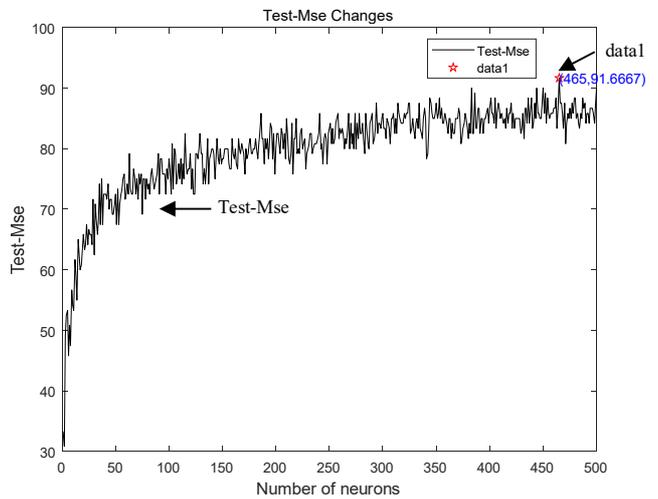


(a) Testing accuracy when the number of neurons in hidden layer is 100

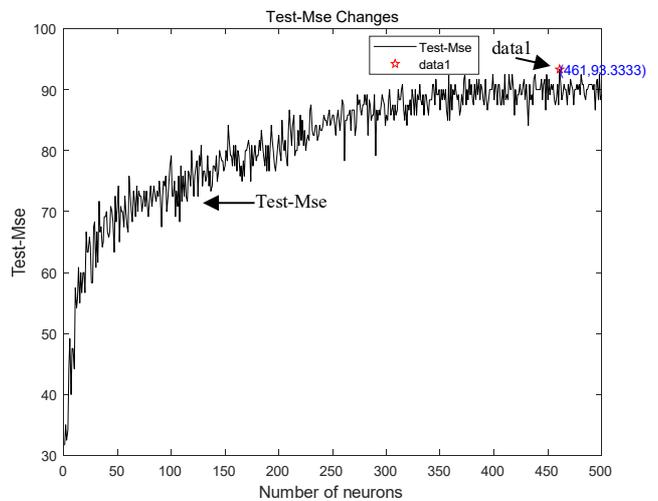


(b) Testing accuracy when the number of neurons in hidden layer is 300

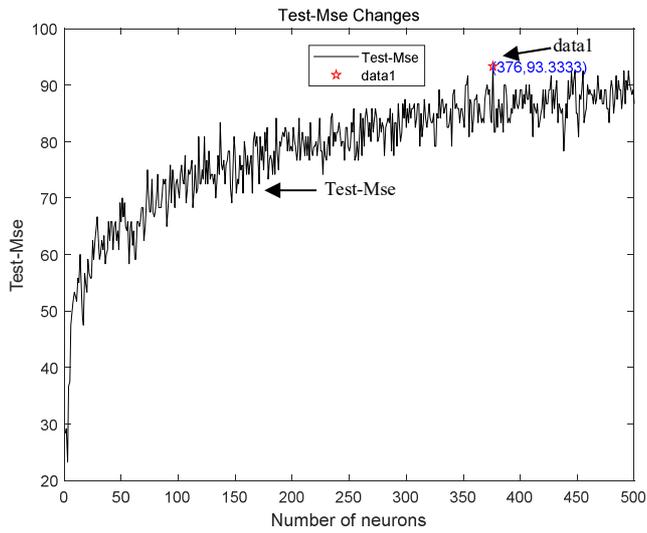
Fig. 4 Diagnosis test results of steel strip surface defect.



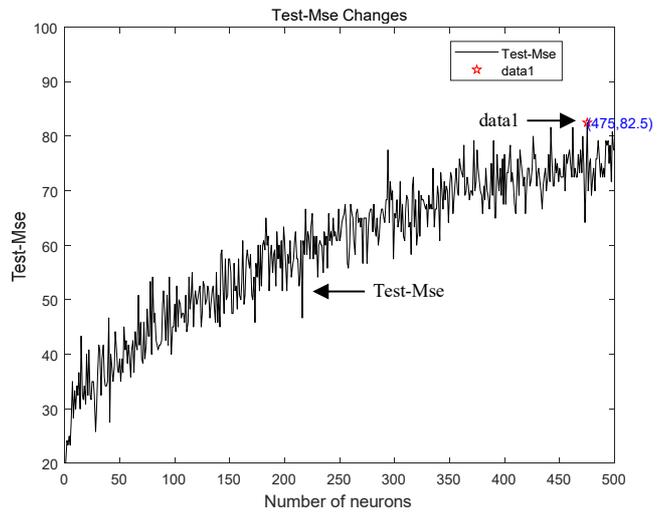
(a) Sigmoid excitation function



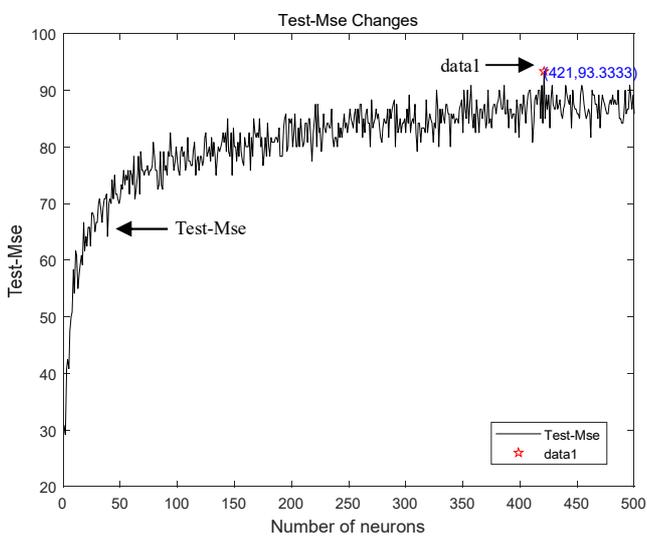
(b) Sine excitation function



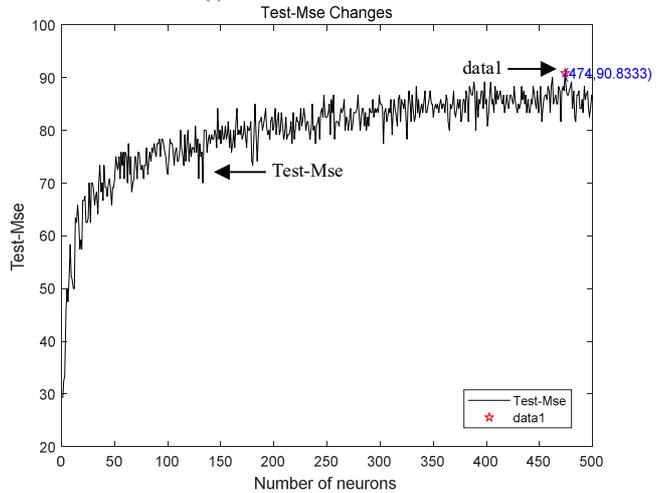
(c) Hardlim excitation function



(f) Morlet excitation function

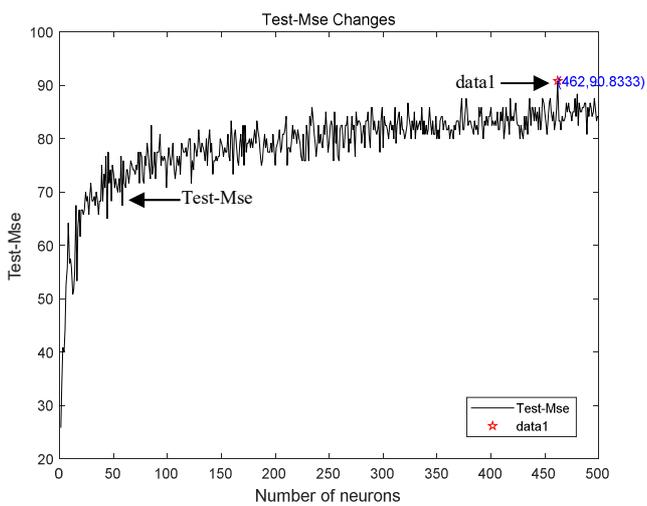


(d) ReLu excitation function

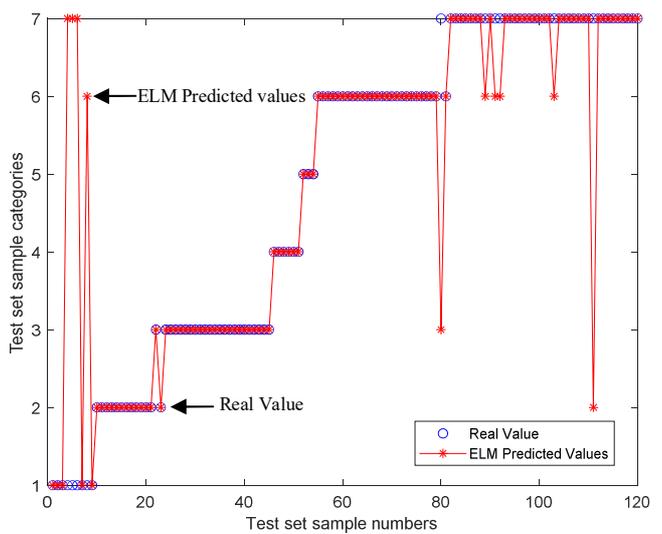


(g) Arctan excitation function

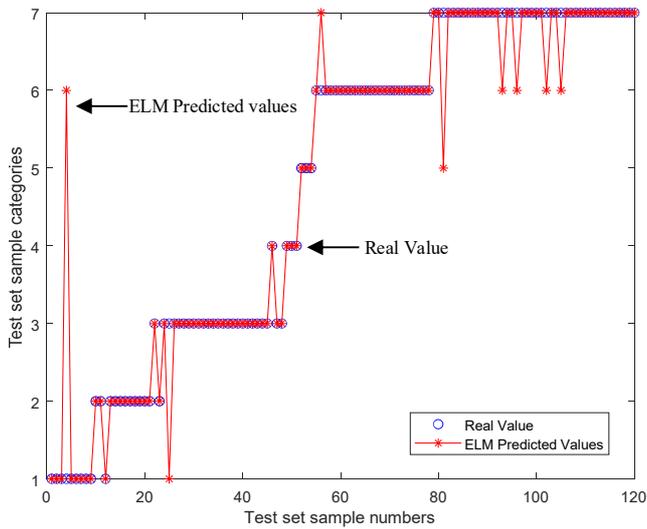
Fig. 5 Simulation results on different neuron number in hidden layer of ELM.



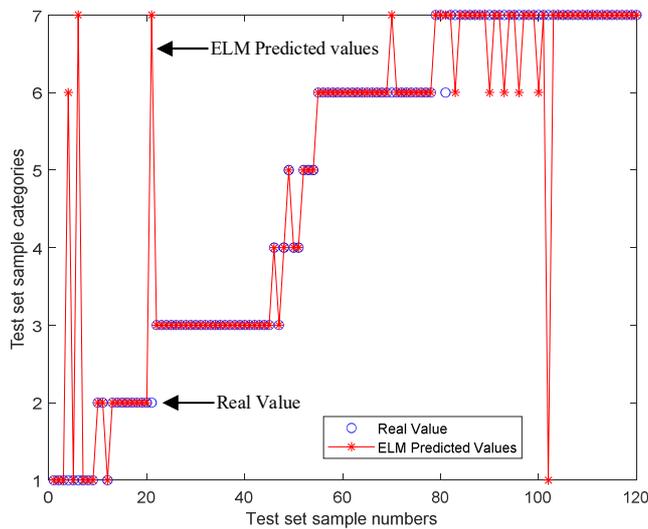
(e) Tanh excitation function



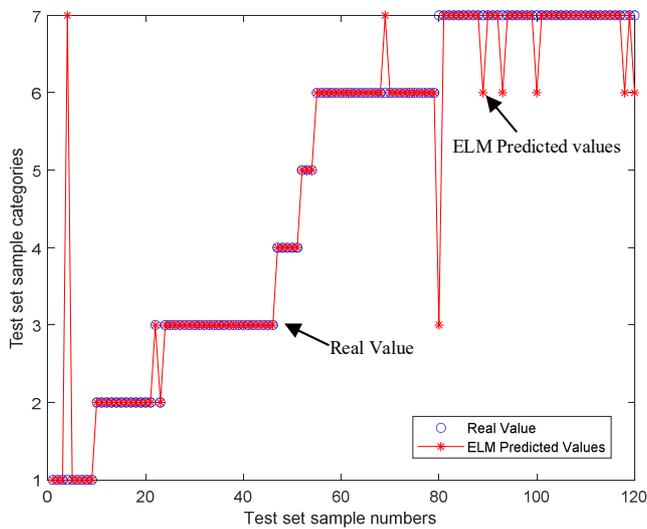
(a) Sigmoid excitation function



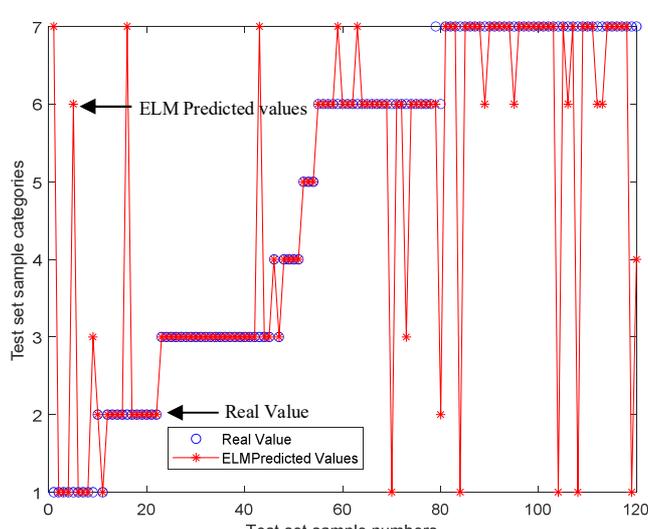
(b) Sine excitation function



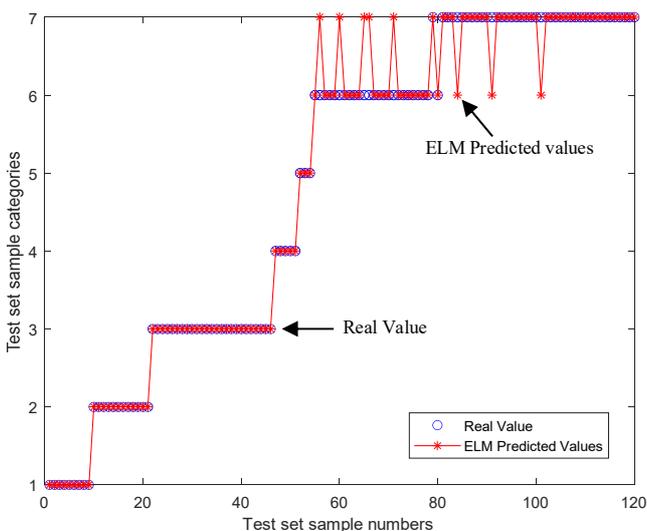
(e) Tanh excitation function



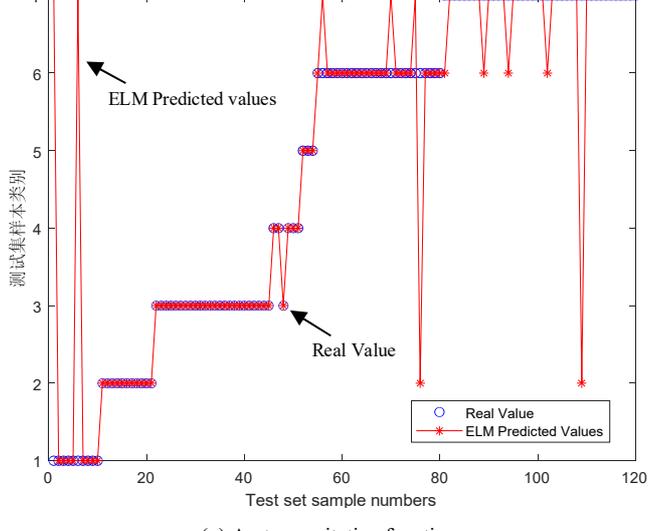
(c) Hardlim excitation function



(f) Morlet excitation function



(d) ReLu excitation function



(g) Arctan excitation function

Fig. 6 Diagnosis results of ELM with different excitation functions under optimal number of hidden layer neurons.

TABLE 4. DIAGNOSTIC RESULTS UNDER THE OPTIMAL NUMBER OF HIDDEN LAYER NEURONS FOR ELM WITH DIFFERENT ACTIVATION FUNCTIONS

Excitation function	Number of hidden layer neurons	Diagnostic accuracy
Sigmoid	465	91.16667%
Sin	461	93.3333%
Hardlim	376	93.3333%
ReLu	421	93.3333%
Tanh	462	90.8333%
Morlet	475	82.5%
arctan	474	90.8333%

IV. CONCLUSION

This paper uses ELM neural network to classify and diagnose strip surface defects. ELM is an algorithm with fast learning speed and good generalization performance. By adjusting the number of neurons in the hidden layer of ELM and selecting different activation functions, the accuracy of ELM neural network classification and diagnosis of strip surface defects is improved. But the connection weights in ELM are randomly initialized, the difference on recognition accuracy on the adopted data set is bigger in the ELM training process for each runtime, and the number of hidden neurons has no uniform standard, which makes it need to obtain the most appropriate number of neural nodes in the hidden layer. Thus, the optimization of ELM neural network need to carry out the conducted in-depth study.

REFERENCES

- [1] V. Arjun, B. Sasi, B. Rao, C. K. Mukhopadhyay, and T. Jayakumar, "Optimisation of Pulsed Eddy Current Probe for Detection of Sub-surface Defects in Stainless Steel Plates," *Sensors and Actuators A Physical*, vol. 226, pp. 69-75, 2015.
- [2] R. Gong, M. Chu, A. Wang, and Y. Yang, "A Fast Detection Method for Region of Defect on Strip Steel Surface," *ISIJ International*, vol. 55, no. 1, pp. 207-212, 2015.
- [3] K. C. Song, H. U. Shao-Peng, Y. H. Yan, and L. I. Jun, "Surface Defect Detection Method Using Saliency Linear Scanning Morphology for Silicon Steel Strip under Oil Pollution Interference," *ISIJ International*, vol. 54, no. 11, pp. 2598-2607, 2014.
- [4] M. Chu, A. Wang, R. Gong, and S. Mo, "Strip Steel Surface Defect Recognition Based on Novel Feature Extraction and Enhanced Least Squares Twin Support Vector Machine," *ISIJ international*, vol. 54, no. 7, pp. 1638-1645, 2014.
- [5] J. Hou, K. Xia, F. Yang, and B. Zu, "Strip Steel Surface Defects Recognition Based on SOCP Optimized Multiple Kernel RVM," *Mathematical Problems in Engineering*, vol. 2018, pp. 1-8, 2018.
- [6] K. Song, and Y. Yan, "A Noise Robust Method Based on Completed Local Binary Patterns for Hot-rolled Steel Strip Surface Defects," *Applied Surface Science*, vol. 285, no. 21, pp. 858-864, 2013.
- [7] G. B. Huang, Q. Y. Zhu, and C. K. Siew, "Extreme Learning Machine: Theory and Applications," *Neurocomputing*, vol. 70, no. 1-3, pp. 489-501, 2006.
- [8] G. B. Huang, D. H. Wang, and Y. Lan, "Extreme Learning Machines: a Survey," *International Journal of Machine Learning and Cybernetics*, vol. 2, no. 2, pp. 107-122, 2011.
- [9] G. Huang, G. B. Huang, S. Song, and K. You, "Trends in Extreme Learning Machines: a Review," *Neural Networks*, vol. 61, pp. 32-48, 2015.
- [10] G. B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme Learning Machine for Regression and Multiclass Classification," *IEEE Transactions on Systems Man & Cybernetics Part B*, vol. 42, no. 2, pp. 513-529, 2012.
- [11] S. Zhang, Z. Liu, X. Huang, and W. Xiao, "A Modified Residual Extreme Learning Machine Algorithm and its Application," *IEEE Access*, vol. 6, pp. 62215-62223, 2018.

- [12] J. W. Zhao, Z. H. Zhou, and F. L. Cao, "Human Face Recognition Based on Ensemble of Polyharmonic Extreme Learning Machine," *Neural Computing & Applications*, vol. 24, no. 6, pp. 1317-1326, 2014.
- [13] F. Lu, J. Wu, J. Huang, and X. Qiu, "Restricted-boltzmann-based Extreme Learning Machine for Gas Path Fault Diagnosis of Turbofan Engine," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 2, pp. 959-968, 2020.
- [14] W. Xie, J. S. Wang, C. Xing, S. S. Guo, M. W. Guo, and L. F. Zhu, "Extreme Learning Machine Soft-sensor Model with Different Activation Functions on Grinding Process Optimized by Improved Black Hole Algorithm," *IEEE Access*, vol. 8, pp. 25084-25110, 2020.
- [15] Q. An, R. Tang, H. Su, J. Zhang, and X. Li, "Robust Configuration and Intelligent MPPT Control for Building Integrated Photovoltaic System Based on Extreme Learning Machine," *Journal of Intelligent and Fuzzy Systems*, vol. 17, pp. 1-18, 2021.
- [16] C. Xing, J. S. Wang, L. Zhang, and W. Xie, "Neural Network Soft-sensor Modeling of PVC Polymerization Process Based on Data Dimensionality Reduction Strategy," *Engineering Letters*, vol. 28, no. 3, pp. 762-776, 2020.